



Pedagogische studiën 1997 : Tijdschrift voor onderwijskunde en opvoedkunde

<https://hdl.handle.net/1874/205429>

Pedagogische Studiën

Tijdschrift voor onderwijskunde en opvoedkunde

Inhoud Jaargang 74 1997

BIBLIOTHEEK UNIVERSITEIT UTRECHT



2801 326 0

Pedagogische Studiën

Tijdschrift voor onderwijskunde
en opvoedkunde

Redactie

Prof. Dr. N. Verloop (voorzitter)
Dr. P.J.J. Stijnen (secretaris)
Dr. K.P. van den Bos
Prof. Dr. J.J. Hox
Prof. Dr. W. Koops
Prof. Dr. G.W. Meijnen
Prof. Dr. P.R.J. Simons
Dr. J.N. Streumer
Prof. Dr. L. Verschaffel
Dr. J.M. Wijnstra

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

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

Inhoudsopgave

Jaargang 74, 1997

Auteursregister

Artikelen (nummers 1 t/m 5; voor nummer 6: zie laatste pagina)

Aarnoutse, C.

zie: Manders, D.

Admiraal, W.F., Korthagen, F.A.J., & Wubbels, Th.

Korte-termijneffecten van coping-gedrag van docenten-in-opleiding 165

Bakermans, J.

zie: Veenman, S.

Bakermans-Kranenburg, M.J.

zie: Schuengel, C.

Beijaard, D.

zie: Jongmans, C.T.

Bleichrodt, N.

zie: Resing, W.C.M.

Blok, H.

Voorlezen op school aan jonge kinderen; een review 233

Blom, M.

zie: Schuengel, C.

Bock, D. De, Verschaffel, L., & Janssens, D.

De lineariteitsillusie bij het oplossen van meetkundeproblemen door leerlingen van het secundair onderwijs 261

Bok, I.A., & Taris, T.W.

Steun van de ouders, sekse, en het verloop van de schoolloopbaan: een retrospectieve studie 33

Dekkers, H., & Smeets, E.

Van vakkenkeuze naar eindexamen: sekseverschillen? 108

Elsäcker, W. van, & Verhoeven, L.

Kleuters leren meer van voorlezen in kleine groepjes 117

Franzen, Y.

zie: Veenman, S.

Galesloot, L.J., Koetsier, C.P., & Wubbels, Th.

Handelingsaspecten bij het wederzijds leren van ervaren docenten 249

Hoof, M. van

zie: Veenman, S.

IJzendoorn, M.H. van

zie: Schuengel, C.

Janssens, D.

zie: Bock, D. De

Jongmans, C.T., & Beijaard, D.

De professionele oriëntatie van leraren en hun betrokkenheid bij het schoolbeleid .. 97

Koetsier, C.P.

zie: Galesloot, L.J.

Korthagen, F.A.J.

zie: Admiraal, W.F.

Lange, A., & Ven, J.-P. van de

Systeemtheoretische benaderingen bij de behandeling van klachten van jongeren 271

Lodewijks, J.G.L.C.

zie: Ponti, E.C.M. De

Manders, D., & Aarnoutse, C.

Wat doen kleuters en leerkrachten in groep 2 van het basisonderwijs? 338

Ponti, E.C.M. De, Lodewijks, J.G.L.C., & Rögels, N.

Studiemotieven en leerstijlen van studenten in het hoger onderwijs voor ouderen 197

Redactie Pedagogische Studiën

Het belang van Nederlandstalige wetenschappelijke tijdschriften 153

- Reitsma, P., Wesseling, R., & Stiva, F.*
Het effect van computer-ondersteunde oefeningen in klanksynthese bij kleuters 1
- Resing, W.C.M., & Bleichrodt, N.*
Schoolgedrag opnieuw gemeten 313
- Resing, W.C.M.*
zie: Slenders, A.P.A.C.
- Rögels, N.*
zie: Ponti, E.C.M. De
- Schuengel, C., IJzendoorn, M.H. van, Bakermans-Kranenburg, M.J., & Blom, M.*
Onverwerkt verlies, gedestruïeerdheid en beangstigend opvoedersgedrag: een vooronderzoek 355
- Slenders, A.P.A.C., & Resing, W.C.M.*
Leerpotentieelonderzoek bij allochtone leerlingen: een verkenning met de Leerpotentieeltest voor Inductief Redeneren ... 183
- Smeets, E.*
zie: Dekkers, H.
- Staphorsius, G., & Verhelst, N.D.*
Indexering van de leesteknik 154
- Stiva, F.*
zie: Reitsma, P.
- Taris, T.W.*
zie: Bok, I.A.
- Uiterwijk, H., & Vallen, T.*
Onderzoek naar bias voor allochtone leerlingen in de Cito-Eindtoets Basisonderwijs 21
- Vallen, T.*
zie: Uiterwijk, H.
- Veenman, S.*
Leeruitkomsten van leerlingen in combinatieklassen: een 'best-evidence'-synthese 73
- Veenman, S., Bakermans, J., Franzen, Y., & Hoof, M. van*
Leren lesgeven met het directe instructiemodel op de tweedegraads lerarenopleiding 328
- Ven, J.-P. van den*
zie: Lange, A.
- Verhelst, N.D.*
zie: Staphorsius, G.
- Verhoeven, L.*
zie: Elsäcker, W. van
- Verschaffel, L.*
zie: Bock, D. De
- Voogt, J.M.*
Courseware in onderzoekend natuurkundeonderwijs: een implementatie perspectief 130
- Wesseling, R.*
zie: Reitsma, P.
- Wubbels, Th.*
zie: Admiraal, W.F.
zie: Galesloot, L.J.

Discussiebijdrage

Bosker, R.J.

Het einde van de klassenstrijd? Achtergronden bij de discussie over klassengrootte in het basisonderwijs..... 210

Kronieken

Bos, K.P. van den & Peer, D.R. van

Congres Dyslexie 1996..... 142

Mooij, T.

Wordt het studiehuis volgens tekening gebouwd?..... 64

Peer, D.R. van

zie: Bos, K.P. van den

Streumer, J.N. (red.)

De Onderwijsresearchdagen 1996 in Tilburg, 4-6 juni 1996, met bijdragen van: *H. Agterbos, A. Booms, S. Doolaard, R. Maslowski, N. Nieveen, E.A.M. van der Schoot, C.G. Sikkema, J.N. Streumer, A. Zanting* 46

Streumer, J.N. (red.)

De Onderwijsresearchdagen 1997 in Leuven, 21-23 mei 1997, met bijdragen van: *P. den Brok, J.F.M. de Fraiture, R. Maslowski, E.R. Savelsberg, J.N. Streumer, C. Terlouw, A.E. van der Valk, J.J.A. Visscher, M.J. Vollebregt, H.J. Vos, S.J. van Zolingen*..... 367

Verloop, N. (red.)

Het AERA-Congres 1997, Chicago, met bijdragen van: *I. Bakkenes, D. Beijaard, B.P.M. Creemers, H.P.J.M. Dekkers, M.L. Krüger, R.R. Meijer, B. van Oers, T. Terwel, N. Verloop*..... 285

Boekbesprekingen

Braster, J.F.A.

De identiteit van het openbaar onderwijs
(*S. Karsten*) 228

Castelijns, J.H.M.

Beelden van bekwaamheid. Een onderzoek
naar de implementatie en effecten van het
programma 'Responsieve instructie' voor
leerkrachten in groep 2 van de basisschool
(*E. Elbers*) 393

*Hair Jr., J.F., Anderson, R.E., Tatham, R.L., &
Black W.C.*

Multivariate data analysis with readings
(4th ed.) (*J.J. Hox*) 390

Luna Rubio, M.

Leerlingbegeleiding: wens en werkelijkheid
(*D.R. Veenstra*) 308

Nossent, S.P.M.

Een beweeglijke psyche. Over epigenese bij
baby's (*A.J. Beekman*) 306

Pols, W., Miedema, S., & Levering, B. (red.)

Opvoeding zoals het is (*F.K. Kieviet*) 69

Stevens, J.

Applied multivariate statistics for the social
sciences (2nd ed.) (*J.J. Hox*) 390

Tabachnick, B.A., & Fidell, L.S.

Using multivariate statistics (3rd ed.) (*J.J.
Hox*) 390

Vygotskij, L.

Cultuur en ontwikkeling (Samenstelling,
inleiding, vertaling en annotaties door R.
van der Veer) (*B.van Oers*) 146

Zolingen, S.J. van

Gevraagd sleutelkwalificaties. Een studie
naar sleutelkwalificaties voor het middelbaar
beroepsonderwijs (*F.T.L. Blokhuis*) 395

Nummer 6 van jaargang 74 is als boekeditie met de titel *Onderwijs, opvoeding en sociale integratie* verschenen en bevat de volgende bijdragen:

Broeder, P., & Extra, G.

Minderheidsgroepen en minderheidstalen in vergelijkend perspectief..... 47

Dekkers, H.

Onderwijs en vrouwen: van achterstand naar differentiatie..... 77

Extra, G.

zie: Broeder, P.

Fase, W.

Intercultureel onderwijs in het licht van sociale integratie..... 67

Gennep, A.Th.G. van

Personen met een verstandelijke handicap en integratie in de samenleving..... 131

Gerris, J.R.M.

Gezin en gezinsopvoeding: moreel fundament van de samenleving?..... 95

Hermanns, J.M.A.

Opvoedingsondersteuning als alternatief voor 'early intervention'..... 165

Heyting, G.F., & Meijnen, G.W.

Sociale integratie: omstrepen in theorie en praktijk..... 9

IJzendoorn, M. H. van

Gehechtheid, ontluikende moraliteit en agressie: naar een sociaal-emotioneel model van antisociaal gedrag..... 113

Kohnstamm, G.A.

Dertig jaar onderzoek naar effecten van pogingen de cognitieve ontwikkeling van jonge kinderen te stimuleren..... 145

Leune, J.M.G.

De normatieve taak van de school en culturele verscheidenheid..... 33

Meijnen, G.W.

zie: Heyting G.F.

Pels, T.

Opvoedingsdiversiteit en de rol van opvoedingsdeskundigen..... 179

Peschar, J.L.

Weer Samen Naar School: een passende structuur voor sociale integratie?..... 195

AB-LOP: 15 OCT 1991. AB-10

Het effect van computer-ondersteunde oefeningen in klanksynthese bij kleuters*

P. Reitsma, R. Wesseling en F. Stiva

Samenvatting

In een onderzoek zijn we nagegaan of een experimenteel ontwikkeld pakket aan computer-ondersteunde oefeningen voor klanksynthese voor kleuters in groep 2 een gunstig effect heeft op de synthese-vaardigheid. In deze ook door kleuters gemakkelijk uit te voeren computer-oefeningen wordt gebruik gemaakt van goed verstaanbare spraak zowel voor de opdrachten als voor commentaar en aanwijzingen. Met de specifiek op auditieve synthese gerichte oefeningen hebben 39 oudste kleuters gedurende ongeveer 4 maanden geoefend, terwijl 45 groepsgenoten op dezelfde computers naar benadering even vaak oefeningen deden in het leggen van associaties van gesproken woorden en plaatjes. Bij deze 84 kleuters en ook bij 69 andere kleuters is zowel voor als na de oefenperiode een synthesesetoets afgenomen.

Uit de resultaten blijkt dat de groep kinderen die de specifieke computer-oefeningen hebben gedaan significant meer vooruitgaan in de vaardigheid om van losse spraaksegmenten één woordklank te maken dan kinderen die de wat deze vaardigheid betreft neutrale computer-oefeningen hebben gedaan. Ook gaan ze significant meer vooruit dan de kleuters die alleen de in hun groepen gebruikelijke activiteiten op school hebben uitgevoerd.

Nadere analyses wijzen verder uit dat de synthesevaardigheid van de kleuters positief beïnvloed wordt door activiteiten van de leerkrachten die gericht zijn op het auditief analyseren en synthetiseren van woorden, maar dat de com-

puteroefeningen een significant additioneel effect hebben. De variatie tussen proefpersonen in de grootte van het effect van de computer-ondersteunde oefeningen bleek gerelateerd aan het aantal daadwerkelijk uitgevoerde opdrachten tijdens de oefeningen. Bovendien zijn de effecten van de training na ongeveer 10 en 18 weken leesonderwijs in groep 3 nog terug te vinden in synthesevaardigheid en decodeerbaarheid. De leerlingen die in groep 2 de klanksynthese-oefeningen hebben gedaan haalden gemiddeld de hoogste scores. Ten slotte worden enkele voordelen en beperkingen van dergelijke oefeningen besproken.

1 Inleiding

Gebruikers van gesproken taal beschikken over buitengewoon efficiënte mechanismen voor de verwerking van gesproken taal. Deze stellen hen in normale situaties in staat tamelijk complexe fonologische structuren functioneel te verwerken. Na een lange evolutie beschikken mensen over dergelijke aangeboren mogelijkheden gesproken taal vrijwel automatisch te verwerken, zowel expressief als receptief. Bij voldoende aanbod van de taal in de omgeving zal een kind in enkele jaren kennis van de taal opdoen en vaardigheden verwerven waardoor het deze taalmogelijkheden in het dagelijks leven volop kan benutten. Een kleuter van een jaar of vijf, hoewel bepaald nog niet volleerd, kan meestal al heel goed zijn mondje roeren en het oor te luisteren leggen. Maar deze reeds betrekkelijk grote vaardigheid in de gesproken taal is niet voldoende om zonder al te veel verdere moeite ook de geschreven taal te leren gebruiken.

Uit de resultaten van veel onderzoek dat in de laatste jaren naar beginnend lezen is verricht blijkt dat voor het leren gebruiken van een alfabetisch geschreven taal een zekere mate van fonologisch bewustzijn nodig is (Mattingly,

* Met dank aan de leerlingen en leerkrachten van De Ark, De Regenboog, De Grote Beer en de A. Bekemaschool voor hun medewerking. We danken eveneens B. Wessels, S. Stroombergen en A. Benschop voor hun enthousiaste hulp in de begeleiding van de leerlingen en leerkrachten en in de verzameling van gegevens.

1972). Onder fonologisch bewustzijn wordt doorgaans verstaan dat iemand beseft heeft van de klankcomponenten van een gesproken woord en in staat is hiermee te manipuleren (fonologische vaardigheden). Er is veel empirische evidentie voor een relatie tussen voorschools fonologisch bewustzijn en vorderingen in het aanvankelijk lezen. In verschillende (longitudinale) onderzoeken zijn hoge correlatie's gevonden tussen scores op toetsen voor fonologische vaardigheid gemeten op een moment voordat formeel leesonderwijs is gestart en leesprestaties na een zekere periode van instructie. Dit geldt niet alleen voor het Nederlandse taalgebied (o.a. Van Dongen, 1984; Mommers, 1990; Bast, 1995), maar ook voor het Engels (Blachman, 1984; Juel, 1991; Stanovich, 1986), Zweeds (Lundberg, Olofsson & Wall, 1980; Olofsson, 1993), Frans (Alegria, Pignot & Morais, 1982), Duits (Schneider & Näsland, 1993) en Russisch (Elkonin, 1973). Het blijkt dus een tamelijk universeel gegeven dat hogere scores op toetsen voor fonologische vaardigheid die afgenomen zijn in de kleutertijd corresponderen met hogere scores op leestoetsen die na een periode van leesonderwijs worden voorgelegd.

De verklaring voor deze relatie is niet zo lastig. Omdat in een alfabetisch schrift de (meeste) lettertekens corresponderen met fonemen, is het voor een beginnende gebruiker van geschreven taal nuttig om enig idee te hebben waar het bij een foneem om gaat. Met fonemen worden de klankelementen in de gesproken taal bedoeld die van belang zijn om verschillende woorden te onderscheiden. Hoewel veel kinderen graag met spraakgeluiden experimenteren, zijn de meeste kleuters van een jaar of vijf niet gewend aandacht aan de klankvorm en interne structuur van woorden te besteden. Ze richten zich vooral op de betekenis, op de communicatieve waarde van woorden. Maar voor leren lezen en spellen is uitdrukkelijk aandacht nodig voor de afzonderlijke klankelementen, de fonemen in een woord. Voor leren lezen en spellen is enig fonologisch bewustzijn nodig. Het aanleren van de relatie tussen de letter *m* en de klank /m/ is voor het lezen en spellen van een woord als *maan* vooral zinvol als begrepen wordt dat deze klank correspondeert met het begin van het woord. Het is uiteraard ook zo dat door instructie en de interactie met schrift-

tekens tijdens het aanvankelijk leesonderwijs een (verdere) ontwikkeling van fonologische vaardigheden sterk wordt gestimuleerd. Er is dus waarschijnlijk een wederzijds beïnvloedende (reciproke) relatie tussen leren lezen en fonologische vaardigheden.

Omdat veel kinderen voordat ze leesonderwijs krijgen niet of nauwelijks over fonologische vaardigheden beschikken en omdat fonologische vaardigheden van belang blijken voor succes in aanvankelijk lezen, is in verschillende studies gepoogd om door middel van instructie en trainingsprogramma's voorafgaand aan het onderwijs in lezen het fonologische bewustzijn bij kinderen te wekken en te verhogen (Ball & Blachman, 1988; Bradley & Bryant, 1983; Bus, 1985; Cunningham, 1990; Lundberg, Frost & Petersen, 1988; Olofsson & Lundberg, 1983, 1985; Wallach & Wallach, 1979; Williams, 1980). In het algemeen hebben deze onderzoeken aangetoond dat training van fonologische vaardigheden in de kleuterperiode, hoewel tamelijk arbeidsintensief, zeker mogelijk is. De trainingen bestaan uit allerlei oefeningen en spelletjes, zoals het vergelijken van gesproken woorden op rijm of overeenkomende beginklank, ontleding van woorden in lettergrepen of fonemen, synthese van losse klanken tot een woord, en dergelijke. Als gevolg van dergelijke oefeningen kunnen fonologische vaardigheden gunstig worden beïnvloed. Bovendien blijkt een significant effect op latere vorderingen in leren lezen en spellen. De kinderen die door middel van dit soort oefeningen geleerd hebben stil te staan bij de vormaspecten van woorden, vertonen na enige tijd van leesonderwijs beduidend betere prestaties op lees- en spellingstaken.

Dus enige voorbereiding van kinderen voordat daadwerkelijk met instructie in lezen wordt begonnen lijkt hun een gunstige uitgangspositie te verschaffen om te profiteren van het aanvankelijk leesonderwijs. In veel Nederlandse kleutergroepen (groep 2 van de basisschool) is het min of meer gebruikelijk om in de groepen enige tijd uit te trekken voor voorbereidende activiteiten met het oog op het latere leesonderwijs. Er wordt dikwijls geen vastomlijnd programma voor gebruikt, maar naar vrije keuze en zonder strikte regelmaat worden in allerlei didactisch georiënteerde spelvormen geprobeerd onder andere de taalontwikkeling van

kinderen te stimuleren. Hierbij komt soms ook het objectiveren van gesproken taal, het aandacht schenken aan de vormaspecten van gesproken woorden aan de orde. Maar het is onduidelijk of leerkrachten in de praktijk dergelijke activiteiten op systematische wijze uitvoeren en of deze activiteiten met voldoende aandacht en frequentie worden uitgevoerd om een zekere vooruitgang in fonologische vaardigheden te bewerkstelligen. In de bovengenoemde trainingstudies kwam in de meeste gevallen het succes tot stand na tamelijk intensief en langdurig oefenen. Hoewel er in principe geen beletsel is deze werkwijze over te nemen en te integreren in het werkplan, vormt een volledig identieke aanpak toch een forse belasting in tijd en moeite voor leerkrachten in kleutergroepen.

Een belangrijk punt van aandacht is ook of met klassikale activiteiten bereikt wordt dat alle leerlingen in een groep voldoende oefening krijgen. Het is bepaald niet uitgesloten dat leerlingen die totaal nog geen idee hebben van klankvormelijke aspecten van de taal, bij gezamenlijke activiteiten in de groep maar moeilijk in de gaten krijgen waar het om draait. Het zou wenselijk kunnen zijn voor deze leerlingen een meer individueel toegesneden oefensituatie te creëren. De individuele aandacht die een leerkracht in een kleutergroep aan de leerlingen kan schenken is maar beperkt, terwijl er toch een stimulerende invloed van uit kan gaan (Schonewille & Van der Leij, 1995). Het is op grond van dergelijke overwegingen dat we naar een alternatief hebben gezocht. We menen die te hebben gevonden in de vorm van computer-ondersteunde oefeningen.

Voor het stimuleren van het fonologische bewustzijn van kleuters zou goed gebruik gemaakt kunnen worden van de verworvenheden van de hedendaagse informatietechnologie. Mits aan een aantal voorwaarden wordt voldaan zou een computer op speelse doch indringende wijze aan individuele leerlingen oefeningen kunnen aanbieden, waardoor fonologische vaardigheden tot ontwikkeling kunnen komen. Het voordeel van een computer is dat zonder al te veel investering van de leerkracht de leerlingen tamelijk intensief en aan het individu aangepaste oefeningen kunnen doen. In een computerprogramma kan de leerling voortdurend worden uitgenodigd actief te

blijven en kan op alle acties van de leerling of het uitblijven daarvan door de computer worden gereageerd. Als voorwaarden voor een vruchtbaar gebruik van computeroefeningen kunnen worden genoemd dat de programma's didactisch verantwoord zijn samengesteld en zeer eenvoudig door kleuters te gebruiken zijn. Voorts is een belangrijk aspect met betrekking tot het gebruik van de programma's door kleuters vooral de toepassing van goed verstaanbare gedigitaliseerde spraak. Kleuters kunnen immers nog niet lezen en typen en tenzij vooraf allerlei afspraken over symbolen worden gemaakt, bieden gesproken boodschappen de enige mogelijkheid om gerichte aanwijzingen en commentaar tijdens de oefeningen te verstrekken. Daarnaast wordt ook door de aard van de oefeningen de eis gesteld dat er zeer goed verstaanbare spraak wordt aangeboden. Als de oefeningen gericht zijn op het bevorderen van fonologische vaardigheden, is het uiteraard nodig om te kunnen beschikken over goed onderscheidbare spraakklanken. Onlangs is door ons een dergelijk computerprogramma ontwikkeld en beproefd (zie voor een verdere beschrijving onder 3. Methode) en dit programma is gebruikt om de onderhavige onderzoeksvraag te beantwoorden.

2 Vraagstelling van onderzoek en opzet

De hoofdvraag van het onderzoek betreft de vraag of met behulp van computer-ondersteunde oefeningen de auditieve synthesevaardigheid van kleuters (verder) kan worden gestimuleerd. Zowel de leeftijd van de leerlingen (rond de 6 jaar) als de aard van de aan te leren vaardigheid zouden kunnen worden beschouwd als redenen om te pleiten voor een aanpak die de persoonlijke inzet van een leerkracht vergt. Omdat een leerkracht vertrouwd is met de jonge leerlingen in de groep, zou vooral goed rekening gehouden kunnen worden met de verschillende ontwikkelingsniveaus en kan nauw worden aangesloten op bijvoorbeeld hun belangstelling en aandachtspanne. Ook is een leerkracht beter in staat om signalen van begrip of onbegrip, enthousiasme of desinteresse op te vangen en om de didactische aanpak aan te passen aan de vorderingen of stagnaties in het leer-

proces. De vraag is daarom of een computerprogramma, waarin dergelijke mogelijkheden uiteraard veel beperkter zijn, toch ook leerzame oefeningen kan bieden. Want aan de andere kant is het wellicht ook zo dat niet alle leerkrachten altijd optimaal van de mogelijkheden gebruik maken om kinderen metalinguïstisch of fonologisch bewustzijn bij te brengen. Hoewel leerkrachten doorgaans zeker werken aan het bevorderen van taalvaardigheden en taalkennis, wordt er in ieder geval vaak beduidend minder tijd en aandacht besteed aan de ontwikkeling van fonologische vaardigheden dan in de eerder genoemde trainingstudies gebruikelijk is. Wellicht kan een computerprogramma hierin een bijdrage leveren.

Van allerlei fonologische vaardigheden als rijmbeoordelingen, beginklanken isoleren, foneemsynthese, woordsegmentatie (auditieve analyse), enzovoorts, is als specifiek doel van de oefeningen gekozen voor auditieve synthese. Deze keuze voor synthesevaardigheid berust op twee overwegingen. In de eerste plaats is auditieve synthese een element van de zogenaamde elementaire leeshandeling. Bij het lezen van een woord door beginnende lezers kunnen drie deelprocessen worden onderscheiden. De eerste is de analyse van het geschreven of gedrukte woord in de samenstellende letters of grafemen, waaraan vervolgens de correcte foneemklanken worden gekoppeld. Door het samenvoegen of de synthese van de afzonderlijke fonemen tot één woordklank wordt ten slotte de identiteit van het woord bepaald. Voor het proces van aanvankelijk lezen is auditieve synthese dus een cruciale en taakrelevante vaardigheid. Het is daarom alleszins de moeite waard na te gaan of deze vaardigheid in de kleuterperiode kan worden geoefend. Een tweede reden waarom auditieve synthese als doel van het trainingsprogramma is gekozen is nogal praktisch van aard. Het bleek niet onoverkomelijk om computeroefeningen te ontwerpen die een beroep op de synthesevaardigheid van een leerling doen. Maar andere oefeningen, zoals segmentatie, doen vaak een beroep op het uitdrukkelijk articuleren van de leerling. Omdat computers vooralsnog niet kunnen horen en zeker niet zo genuanceerd kunnen luisteren als vereist is voor een goede beoordeling van de uitingen in een segmentatie-oefening, is besloten de aandacht voorals-

nog te richten op synthesevaardigheid.

Een groep oudste kleuters, in het laatste halfjaar van groep 2, krijgt gedurende ongeveer vijf maanden de gelegenheid diverse computer-ondersteunde oefeningen te verrichten. In de oefeningen worden eerst voor een beperkte verzameling plaatjes de associatie tussen deze plaatjes en gesproken woorden geoefend. Bij een door de computer uitgesproken woord dient de leerling het juiste plaatje op het scherm te kiezen. Nadat op deze wijze de koppeling tussen plaatje en woord bekend is geraakt, wordt vervolgens de oefening gewijzigd. De computer spreekt nu niet meer een geheel woord uit, maar het woord wordt gesegmenteerd in afzonderlijke foneemklanken uitgesproken. Dezelfde verzameling woorden en plaatjes wordt hiervoor gebruikt. In een volgende oefening wordt de leerling gevraagd een reeks foneemklanken te vergelijken met een hele woordklank zonder dat er nog ondersteuning wordt geboden door presentatie van plaatjes. Een laatste oefenvorm bestaat uit het zoeken van het juiste woord uit een aantal alternatieven op basis van een gesegmenteerd aangeboden woord.

Voordat met de oefeningen wordt begonnen (januari) en vlak voor de zomervakantie (juni) worden de leerlingen diverse taken voorgelegd, waaronder een auditieve syntheseset. De vraag is nu of de prestaties op deze taak tussen de voor- en nameting in belangrijke mate toe zijn genomen. Om na te gaan of een eventuele toename vooral toe te schrijven is aan het uitvoeren van de oefeningen, zijn de taken op dezelfde momenten eveneens afgenomen bij twee controlegroepen. Een eerste controlegroep wordt gevormd door enkele groepen kleuters in een andere school, die het gebruikelijke onderwijs in groep 2 blijven volgen. In zowel de experimentele groepen als in deze controlegroepen zal worden nagegaan welke activiteiten de leerkrachten ontplooiën om het fonologisch inzicht of de fonologische synthesevaardigheid van hun leerlingen te bevorderen. Als de leerkrachten van deze groepen in dezelfde mate (groepsgewijs) aandacht schenken aan het stimuleren van de betreffende leesvoorwaarden, zal indien de experimentele groep een grotere ontwikkeling in synthesevaardigheid laat zien, deze vooral zijn veroorzaakt door het uitvoeren van de oefeningen aan de computer.

Om twee redenen is echter een tweede controlegroep toegevoegd. Ten eerste is het onzeker of de door de leerkracht geïnitieerde leer- en ontwikkelingsactiviteiten met betrekking tot synthesevaardigheid in de verschillende groepen vergelijkbaar zijn. Indien er verschillen zijn tussen de groepen, dan is het maar de vraag of er een onderscheid gemaakt kan worden tussen het effect van de leerkracht c.q. de ontplooide activiteiten in de groep en het effect van de computerondersteunde oefeningen. Ten tweede kan de introductie van de computer al allerlei gevolgen hebben. Het mogen werken aan de computer en het bezig zijn met woordklanken kan zowel bij de leerkracht als bij de leerlingen de aandacht in het bijzonder richten op het bezig zijn met taal en klankoefeningen. Het effect van de specifieke synthese-oefeningen met behulp van de computer kan daardoor versluierd raken. Daarom is de tweede controlegroep gekozen uit dezelfde kleutergroepen als waaruit de experimentele groep leerlingen afkomstig is. In deze groepen kregen alle leerlingen ongeveer even frequent computer-ondersteunde oefeningen, maar de controlegroep ontving geen oefeningen voor klanksynthese.

Nadat het bovenbeschreven hoofdonderzoek naar het effect van een computerondersteund trainingsprogramma voor klanksynthese was opgezet en uitgevoerd, kwam als tweede onderzoeksvraag naar voren of een mogelijk effect van de training in klanksynthese ook doorwerkt op de vorderingen in het aanvankelijk lezen. Als kleuters althans één van de deeltaalvaardigheden van het technisch lezen alvast met succes hebben geoefend, dan zou verwacht kunnen worden dat dit een gunstige invloed heeft op het lees-leerproces. Gezien het in de inleiding reeds gememoreerde verband tussen fonologische vaardigheden in de kleuterperiode en de leesresultaten in groep 3 mag een positief effect inderdaad worden verwacht. Fonologische vaardigheden kunnen in zekere zin als voorwaarde gelden voor leren lezen en spellen en enige oefening in dergelijke voorwaarden zal een gunstig effect moeten hebben op vorderingen in leren lezen. Maar tevens is opgemerkt dat het onderricht in lezen en spellen een sterk effect op de ontwikkeling van fonologische vaardigheden kan hebben. Inderdaad blijkt uit diverse onderzoeken (o.a. Morrison, Smith &

Dow-Ehrensberger, 1995; Bentin, Hammer & Cahan, 1991) dat de toename in fonologische vaardigheden gedurende het jaar voorafgaand aan leesinstructie vele malen geringer is dan de toename in deze vaardigheden tijdens de periode van het aanvankelijk lezen. De reden is uiteraard dat tijdens de instructie in het aanvankelijk lezen de fonologische vaardigheden zoals analyse en synthese heel nadrukkelijk worden geoefend, ook met behulp van geschreven letters en woorden. In de eerste paar maanden van de instructie in het lezen staat in Nederland het leren decoderen centraal en wordt doorgaans veel aandacht besteed aan het auditief analyseren en synthetiseren van gesproken woorden. In feite maken deze vaardigheden vaak een onderscheiden deel uit van de onderwijsleermethode. De fonologische vaardigheden van kinderen in het eerste halfjaar van groep 3 nemen vermoedelijk veel sterker toe dan in het voorafgaande kleuterjaar. Door de intensieve instructie in groep 3 kunnen aanvankelijke verschillen in fonologische vaardigheden spoedig vervagen. Terwijl de vaardigheden in de kleuterperiode het gecombineerde resultaat kunnen zijn van betrekkelijk toevallige persoonlijke talenten en interesses en een niet voor alle kinderen even stimulerende thuis- en schoolomgeving, worden ten gevolge van klassikale instructie in groep 3 veel van deze aanvankelijke verschillen ingewisseld door verschillen in ontvankelijkheid voor dergelijke instructie.

In het huidige onderzoek is evenwel bij een deel van de leerlingen expliciet de klanksynthese geoefend. De vraag is nu of de tot stand gebrachte individuele verschillen gerelateerd is aan de mate waarin de leerlingen gedurende de eerste maanden in een aanvankelijk leesonderwijs de elementaire leeshandeling, het decoderen van woorden met een eenvoudige structuur, onder de knie krijgen. Om de effecten van de oefeningen in klanksynthese op aanvankelijk lezen te bepalen is bij zoveel mogelijk kleuters, die in één van de eerder beschreven onderzoekscondities zaten, na ongeveer 10 weken onderwijs in groep 3 een test voor decoderen en auditieve synthese afgenomen. Na 5 maanden in groep 3 werd nogmaals de decodeervaardigheid en de synthesevaardigheid getoetst.

3 Methode

Proefgroepen

Kleuters uit 12 groepen van vier verschillende basisscholen hebben aan het onderzoek meegewerkt (na verkregen toestemming / instemming van hun ouders). In totaal zijn de gegevens van 153 kleuters in het onderzoek betrokken. In Tabel 1 staat de verdeling van de aantallen over de verschillende onderzoekscondities: de groep die de experimentele synthese oefeningen op de computer kreeg (experimentele computergroep - EC), de controlegroep die controleoefeningen op de computer kreeg (computer controlegroep - CC), en als gewone controlegroep (GC) de kleuters die in het kader van dit onderzoek op de voor- en natoets na verder ongemoeid werden gelaten. De kinderen bezochten basisscholen zeer nabij Amsterdam en hadden verschillende etnische achtergronden (Nederlands - 68%, Surinaams of Antilliaans - 18%, of overige, zoals Marokkaans of Turks - 14%). De gemiddelde leeftijd was op de peildatum van 1 januari 5 jaar en 10 maanden; de jongste was 5 jaar en 2 maanden en de oudste 7 jaar en 1 maand). Het waren allen kinderen waarvan in januari werd aangenomen dat ze het volgend cursusjaar naar groep 3 zouden gaan om onder andere onderwijs in aanvankelijk lezen te krijgen. De gemiddelde leeftijd (in aantal maanden) en het aantal jongens en meisjes per onderzoeksgroep staat in Tabel 1 weergegeven.

Voor alle groepen en leerlingen is door middel van klasobservaties en interviews met leerkrachten en ouders gepoogd na te gaan of en hoe de onderwijsomgeving en de thuissituatie van invloed zijn op de fonologische synthesevaardigheid. Ook werd nagegaan of de onderzoeksgroepen hierin verschillen. Aan vrijwel alle leerkrachten is in mei/juni door middel van een gestructureerd interview informatie gevraagd over de dagindeling, frequentie en

inhoud van diverse activiteiten. Gelet werd op het voorkomen en de frequentie van activiteiten als kringgesprekken, voorlezen, vertellen, voorbereidend rekenen, taalactiviteiten en oriëntatie op schrift cq. voorbereidend lezen (zoals woordspelletjes, versjes, gebruik van werkbladen, stempelen, rijm oefeningen, auditieve discriminatie). Door plotselinge persoonlijke wisselingen is voor één (controle) groep dergelijke informatie niet beschikbaar gekomen. Aan de ouders of verzorgers zijn in een gestructureerd interview vragen gesteld over nationaliteit, thuistaal, opleiding, beroep, bezit van boeken, bibliotheekbezoek, frequentie van voorlezen, en dergelijke. Voor 114 van de 153 leerlingen zijn deze vragenlijstgegevens verkregen. Een weergave en bespreking van deze gegevens zou hier te veel ruimte vergen. Hoewel er verschillen gevonden werden in klas-situatie en in thuisachtergrond, die enigszins met vaardigheid van klanksynthese samenhangen, is het voor het huidige onderzoek vooral belangrijk te kunnen melden dat de hier beschreven onderzoeksgroepen geheel vergelijkbaar bleken. Er zijn, op één uitzondering na, geen belangrijke verschillen in achtergrondvariabelen gevonden die van invloed zijn op de hier gerapporteerde vergelijkingen tussen de onderzoeksgroepen. De uitzondering betreft het in de groep oefenen van auditieve synthese en analyse. In vier van de 12 groepen werd hieraan al vanaf het begin van het schooljaar met enige regelmaat aandacht besteed, terwijl dit in de overige groepen niet of nauwelijks aan bod kwam.

Na de zomervakantie gingen de meeste kleuters naar groep 3 van dezelfde basisschool. In begin november, na ongeveer 10 weken leesonderwijs, is de technische leesvaardigheid van alle leerlingen in de groepen 3 van de betrokken vier scholen bepaald en is nog een natoets voor auditieve synthese afgenomen.

Tabel 1

Enkele leerlinggegevens per onderzoeksconditie

	EC	CC	GC
totaal aantal	39	45	69
aantal meisjes	17	28	25
aantal jongens	22	17	44
gemiddelde leeftijd (maanden)	70.3	70.1	70.8
woordenschat (max. 98)	54.7	53.6	54.1
woorden herkennen (max. 10)	1.2	0.8	0.7
letterkennis (max. 27)	6.6	6.1	7.5

Door verhuizingen, en dergelijke, waren uit de oorspronkelijke onderzoekscondities nog 133 kinderen beschikbaar (voor EC, CC, en GC, respectievelijk, 32, 38 en 63). De vermindering in aantal proefpersonen was redelijk gelijk verdeel over de drie condities. Er is geen enkele controle geweest op de aard of kwaliteit van de instructie in aanvankelijk lezen in groep 3.

Voortoetsen en natoetsen

Tijdens de voortoets is allereerst bij alle kinderen de passieve woordenschat gepeild door het afnemen van de betreffende subtoets van de Taaltoets voor Allochtone kinderen (Verhoeven & Vermeer, 1986). De toets is individueel afgenomen. De gemiddelde scores van elke groep op deze woordenschattoets staan in Tabel 1. Vergelijking met de normtabellen van de toets geeft aan dat de woordenschatscores van de huidige wat betreft afkomst divers samengestelde groepen redelijk overeenkomt met hetgeen kan worden verwacht op basis van de normgegevens.

Tevens is tijdens de voortoets een lijst met 10 eenvoudige woorden (zoals *vuur*, *boom*) aan elke kleuter voorgelegd met de vraag of ze één of meer woorden konden lezen. Genoteerd werd het aantal woorden dat correct door de kleuter werd genoemd.

Kennis van letters, de naam of de klank waarnaar de letter verwijst, werd eveneens getoetst en wel op de volgende manier. Een blad met alle (onderkast) lettertekens in groot formaat en inclusief de *ij* werd aan elke leerling voorgelegd met de vraag of hij hiervan ook letters kende. Er werd doorgevraagd totdat geen enkele letter meer kon worden genoemd. De gemiddelde scores per groep voor het aantal herkende woorden en de letterkennis staan in Tabel 1.

Als belangrijkste toets fungeerde verder in dit onderzoek een test voor synthesevaardigheid. De toets bestaat uit 52 opgaven. In elke opgave werd een leerling verschillende klanksegmenten aangeboden met het verzoek te proberen hier één geheel van te maken. De aangeboden klanksegmenten variëren van lettergrepen, eenheden van een medeklinker (C - consonant) en een klinker (V - vocaal) tot afzonderlijke foneemklanken en werden via een speciaal vervaardigd computerprogramma aangeboden. Er kwamen 10 categorieën van

woorden aan de orde. Van elke categorie wordt hier een tweetal voorbeelden gegeven, waarbij een streepje aangeeft dat op die plaats in het spraakaanbod een pauze is ingelast van ongeveer 1 seconde. De categorieën zijn: (1) lettergrepen: ja-rig, za-ter-dag; (2) CV & VC: t-oe, a-f; (3) C-VC z-us, g-eit; (4) CV-C: ga-s, me-t; (5) CC-VC: kl-op, st-ip; (6) CV-CC: ta-nd, do-rp; (7) C-V-C: l-ie-f, r-oo-d; (8) C-C-V-C: t-r-e-k, k-l-i-m; (9) C-V-C-C: g-e-l-d, v-a-l-s; (10) C-C-V-C-C: b-l-o-n-d, k-r-a-n-s. Het totaal aantal opgaven per categorie was 5, uitgezonderd de eerste 2 waarvoor 6 opgaven waren opgenomen. Voor elke opgave werd een punt toegekend indien het bedoelde woord door de leerling als één geheel werd uitgesproken. De maximum score was 52. Er was een afbreekregel, waardoor de opgaven uit categorieën 7 tot en met 10 slechts aangeboden werden aan leerlingen die in de voorgaande 32 opgaven niet meer dan 10 fouten hadden gemaakt.

De opbouw van de voortoets was gelijk aan de natoets, dat wil zeggen, hetzelfde aantal woorden met hetzelfde type opsplitsingen werd gebruikt, de woorden zelf verschilden echter. De betrouwbaarheid van de twee toetsen werd door middel van item-analyse bepaald voor de gehele groep van 153 leerlingen. De interne consistentie bleek hoog: Cronbachs alpha was 0.95 in de voormeting en van de toets die in de nameting werd gebruikt 0.97.

De experimentele decodeertoets, die in begin november in groep 3 werd afgenomen, bestond uit twee lijsten van 35 woorden met een eenvoudige CVC structuur (bijvoorbeeld, *boot*, *lip*, *map*, *rok*, *bier*). Elke leerling werd individueel getoetst en kreeg gedurende één minuut de gelegenheid zoveel mogelijk woorden van een lijst hardop te lezen. Tussen de afname van de twee lijsten werden enkele andere testjes gedaan. Het aantal correct gelezen woorden in één minuut werd per lijst genoteerd. De correlatie tussen de 133 individuele scores op de twee lijsten bleek .89. Als decodeerscore werd de som van de twee scores genomen.

De in november bij groep 3 afgenomen klanksynthese-toets bestond uit 20 items (C-V-C, vergelijkbaar met categorie 7 uit de bovengenoemde toets). Het aantal correct en volledig gesynthetiseerde woorden werd genoteerd.

In januari werd tot slot nogmaals een deco-deertoets afgenomen (Eén-minuuttest, Brus & Voeten, 1973) en een synthesesetoets die vergelijkbaar was met de oorspronkelijke voortoets.

Het computerprogramma (AdWord)

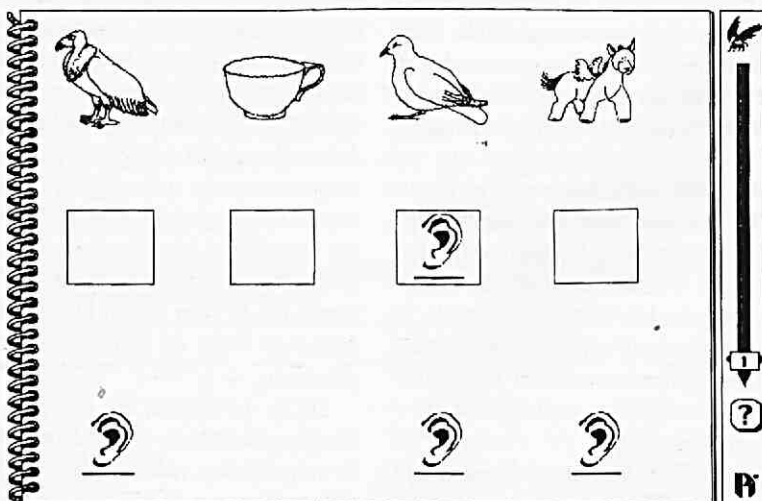
Voor dit onderzoek is gebruik gemaakt van een binnen het samenwerkingsverband tussen het Paedologisch Instituut en de Vrije Universiteit ontwikkeld computerprogramma voor aanvankelijk lezen en spellen (AdWord). Het is een pakket van oefeningen waarmee leerlingen individueel en intensief allerlei deelvaardigheden voor het aanvankelijk lezen en spellen op een computer (PC of Macintosh) kunnen oefenen. Een belangrijk kenmerk van het programma is dat bij de oefeningen veel gebruik gemaakt wordt van gesproken aanwijzingen en commentaar. De oefeningen kunnen hierdoor ook gedaan worden door leerlingen die nog niet of slecht kunnen lezen. Alle oefeningen hebben een sterk interactief karakter, waardoor de leerling voortdurend uitgenodigd wordt de oefeningen voort te zetten. De leerling wordt actief gehouden door op vrijwel elke actie van de leerling een gepaste (gesproken) reactie door het programma te laten volgen. De in deze programma's geproduceerde spraak is van uitstekende kwaliteit en heel goed verstaanbaar. De kunstmatig geproduceerde spraak berust op professionele geluidsopnames van een logopediste, welke gedigitaliseerd en na bewerking in de computer zijn opgeslagen (22 kHz / 8 bits) en via een geluidskaart (PC) of ingebouwde

DA converter weer ten gehore kunnen worden gebracht.

Het pakket AdWord bevat oefeningen voor de beginfasen van leren lezen en spellen. Het begint met oefeningen gericht op het vergroten van de woordenschat en het oefenen van fonologische vaardigheden (o.a. synthese en analyse). In deze oefeningen wordt leerlingen vooral geleerd op allerlei klankaspecten van woorden te letten. Er komen slechts plaatjes en gesproken woorden of foneemklanken voor en nog geen geschreven letters of woorden. Er zijn ook oefeningen waarin de verbinding tussen een gesproken en een geschreven woord, al of niet vergezeld van een afbeelding kan worden geoefend. Ook kunnen letters en klanken afzonderlijk worden geoefend. Voor de verder gevorderde leerlingen zijn er o.a. dictee-oefeningen om spelling te oefenen. De verschillende oefenvormen zijn in moeilijkheidsgraad te variëren door de keuze van diverse varianten van de oefeningen en verschillende materialen. De oefeningen sluiten allemaal aan – of kunnen zodanig eenvoudig worden aangepast – bij vele gangbare leesmethoden, zowel in materiaal als in vormgeving.

Ten behoeve van het huidige onderzoek is gebruik gemaakt van een viertal oefenvormen uit het AdWord pakket. De volgende korte beschrijving van deze oefeningen en varianten is van toepassing voor de leerlingen in de experimentele groep (EC).

a. *Woordenschat.* Het oogmerk van deze oefening is het leren luisteren naar gesproken



Figuur 1. Een voorbeeldscherm van de derde variant van de woordenschatoefening

woorden via de hoofdtelefoon, het leren kennen van de plaatjes en het koppelen van de plaatjes aan gesproken woorden en omgekeerd. Op het scherm staan onderaan een aantal (getekende) oren en bovenaan een aantal plaatjes met lege vakjes eronder. Wordt er op een oor geklikt dan wordt er een woord gezegd. Wordt er op een plaatje geklikt dan wordt het woord uitgesproken dat bij het plaatje hoort. Het is de bedoeling dat de leerling het oor sleept naar het lege vakje onder het plaatje waar het bij hoort. Wordt een oor in het goede vakje gelegd dan blijft het daar liggen en vertelt de computer dat het goed is. Bij het maken van een fout, plaatst de computer het oor weer terug onderaan het scherm en kan er een nieuwe poging worden gedaan. Maakt een leerling teveel fouten achter elkaar (voor dit onderzoek ingesteld op drie fouten achtereen), dan doet de computer het voor door de oren naar het goede vakje te slepen. Al deze gebeurtenissen op het scherm worden begeleid door gesproken commentaar en/of uitleg.

De eerste variant van deze oefening heeft de vorm van één woord (en oor) met 4 plaatjes, de tweede variant heeft meer woorden (oren) en één plaatje en een derde variant heeft 4 oren en 4 plaatjes (zie Figuur 1). Met een verzameling van 20 verschillende CVC woorden en bijbehorende plaatjes hebben de leerlingen in deze 3 varianten de associatie tussen woord en plaatje geoefend. Hierbij waren de klanken van de verschillende woorden duidelijk van elkaar te onderscheiden. Vervolgens zijn de twee laatste varianten nogmaals herhaald, maar nu zo gegroepeerd dat woorden die per scherm beschikbaar waren qua klank erg op elkaar lijken en slechts in één foneem van elkaar verschillen (bijvoorbeeld, *dik, fik, lik, sik* of *gaas, haas, kaas, vaas*). In de laatste twee varianten van deze oefenvorm is het dus noodzakelijk goed te luisteren naar relatief kleinere verschillen tussen woorden.

b. Synthese. Evenals in de laatste variant van de vorige oefening kwamen in deze oefening per opgave 4 oren en 4 plaatjes tegelijk op het scherm. Het verschil is dat na het klikken op een oor niet meer de gehele woordklank wordt geproduceerd, maar dat de woorden nu gesegmenteerd in foneemklanken (ongeveer 1 per sec) ten gehore worden gebracht. In vergelijking met Figuur 1 is nu in deze oefening de

streep onder een oor onderbroken in plaats van doorlopend. De leerling moet na het afluisteren zelf de klanken synthetiseren en het bijbehorende plaatje kiezen. Terwijl in de eerste variant van deze oefening de woorden onderling in foneemstructuur duidelijk verschillen, wijken de alternatieven in een tweede variant slechts in één foneem van elkaar af.

c. Auditieve woordvergelijking. In deze oefening worden twee gesproken woorden aangeboden, geheel of gesegmenteerd, en de leerling moet bepalen of deze twee woorden wel of niet hetzelfde klinken. Het eerste woord wordt altijd geheel en het tweede woord geheel of gesegmenteerd uitgesproken. De klank van elk woord is opnieuw oproepbaar door op een oor te klikken. Tegelijk met het uitspreken van het gehele woord wordt één vak op het scherm getekend, bij het uitspreken van het gesegmenteerde woord worden net zoveel vakjes getekend als er losse klanken in het woord zijn. De vakken voor de twee woorden komen onder elkaar. Naast elk vak komt een oor met een continue of onderbroken streep waarmee aangegeven wordt of er een gehele of gesegmenteerde woordklank te horen is. Daaronder komen twee keuzeblokken (één blanco en één met een 'is-gelijk' teken) waarmee de leerling door erop te klikken kan aangeven of de woorden gelijk zijn of ongelijk. Het computerprogramma legt dit van te voren eerst duidelijk uit door middel van een voorbeeld. Er zijn vier varianten van deze oefening: (a) het tweede woord wordt als geheel uitgesproken en verschilt in meer dan één klanksegment van het eerste woord, (b) als (a) maar met slechts één foneem verschil, (c) en (d) zijn identiek aan (a) en (b) met dien verstande dat het tweede woord steeds in stukjes (segmenten) wordt aangeboden.

d. Auditieve woordvergelijking (meerkeuze). Op het scherm ziet de leerling een oor met daaronder een gehele of onderbroken streep. Wanneer de leerling op het oor klikt, hoort het respectievelijk een geheel of gesegmenteerd woord. Rechts daarvan ziet de leerling drie tot negen lege vakjes. Door op één van deze vakjes te klikken wordt of een geheel of een gesegmenteerd woord uitgesproken. De bedoeling is dat de leerling het oor naar het vakje sleept met het corresponderende woord. Opnieuw zijn er vier varianten: (a) oor geeft hele woordklank

en vakjes geven gesegmenteerde woorden, (b) oor geeft gesegmenteerde woordklank en vakjes geven hele woorden, (c) en (d) zijn identiek aan (a) en (b) wat betreft vorm en taakstelling, maar de (meeste) woorden zijn nu onderling zeer gelijkkluidend (meestal verschillen ze slechts in één foneem).

De leerlingen in de computer-controle groep (CC) kregen eveneens een reeks computer-ondersteunde oefeningen toegewezen. Maar het betrof hier slechts de eerste drie varianten van de woordenschatoefening (a) en het auditief vergelijken van gehele woorden (de eerste variant van oefening c). Dezelfde oefening werd dus veel vaker herhaald. In plaats van een beperkte verzameling van 20 woorden, werd aan deze groep echter een steeds nieuwe verzameling woorden aangereikt, in totaal waren hiervoor 420 verschillende plaatjes en woorden beschikbaar gesteld.

Procedure van de training

Nadat in de maand januari bij alle leerlingen de voortoetsen waren afgenomen, zijn binnen 7 groepen van drie basisscholen de computer-ondersteunde oefeningen begonnen. Binnen elke groep waren ongeveer evenveel kinderen in de experimentele condities opgenomen als in de controle-conditie. De toewijzing van leerlingen aan een conditie gebeurde volstrekt aselektief, met dien verstande dat gepoogd werd de groepen qua beginniveau vergelijkbaar te houden. Daartoe zijn per groep de leerlingen gerangordend in prestaties op de voortoetsgegevens van de synthesetoets. Vervolgens is per opeenvolgend tweetal één van de twee aselekt toegewezen aan de experimentele conditie, de andere leerling kreeg de controleoefening op de computer. De 5 groepen van de vierde basisschool fungeerde als normale controlegroep.

Begin februari werd begonnen met de computeroefeningen. Nadat eerst groepsgewijs enige introductie over de computerprogramma's was gegeven, is bij de eerste oefening in groepjes van drie leerlingen enige begeleiding verzorgd. Hierbij werd uitgelegd, gedemonstreerd en geoefend hoe de leerling de oefeningen zelf kon opstarten, hoe de muis werkte, wat de functie van de 'oren' was, de hulpknop, en dergelijke. Vrijwel alle leerlingen waren daarna in staat de oefeningen zelfstandig uit te voeren. Voor slechts enkele leerlingen is nog een

korte periode van nazorg nodig gebleken.

Iedere groep had een eigen computer beschikbaar voor dit onderzoek. Nadat een computer door de leerkracht was aangezet (het programma startte vervolgens automatisch op), konden de leerlingen beurtelings aan de computer aan het werk. Op het beginscherm stond een lijst met alle (voor)namen van de leerlingen in de groep. Een leerling kon met behulp van de muis zijn eigen naam aanwijzen en klikken. Vervolgens begon er een soort wachtwoord procedure voordat de oefening daadwerkelijk werd opgestart. Iedere leerling had vooraf uit een verzameling van 64 symbolen en plaatjes een eigen en 'geheim' plaatje kunnen kiezen. Nadat de eigen naam was aangewezen moest ook het eigen plaatje worden aangewezen om de oefeningen echt te kunnen laten beginnen. De leerkrachten beschikte over de lijst van namen met bijbehorende plaatjes voor het geval een leerling zijn 'geheimplaatje' was vergeten. Deze procedure was bedacht omdat in het kader van het onderzoek het wenselijk was te kunnen beschikken over de individuele oefengegevens. Het was dus van belang dat een leerling uitsluitend de hem toegewezen oefeningen zou verrichten en niet die van een andere. Alle acties, aangeboden woorden, keuzes van de kinderen, gebruikte tijd, en dergelijke, van alle oefensessies werden door het programma opgeslagen, waardoor achteraf controle mogelijk was op de verrichte oefeningen en de bereikte resultaten. Uit observaties, reacties van leerlingen en rapportage van leerkrachten bleek dat deze wachtwoord-procedure zeer bevredigend werd toegepast.

Gedurende ongeveer 4 maanden (van begin februari tot begin juni) is door de leerlingen op deze wijze met de computer geoefend met een streeffrequentie van 2 à 3 keer per week. Vakantieperiodes, en dergelijke, maakten dat effectief er een 12-tal weken kon worden geoefend. Om iedereen zoveel mogelijk gelijkkelijk de gelegenheid tot oefenen te geven was het programma zo ingesteld dat iedere leerling per dag maximaal 10 minuten kon werken aan de computer. Indien op een dag de oefentijd verstreken was, werden de gegevens zorgvuldig bewaard en kon de leerling een volgende dag verder gaan met de oefening die eerder was afgebroken. Ook voor dit doel was een opslag van individuele oefengegevens en dus een juis-

te toepassing van de wachtwoord-procedure vereist.

Begin juni is aan alle betrokken leerlingen in de drie verschillende onderzoekscondities opnieuw een synthese-toets met hetzelfde aantal en type items als de voortoets voorgelegd. De woorden die in deze toets voorkwamen kwamen overigens niet tijdens de computeroefeningen aan bod.

4 Resultaten

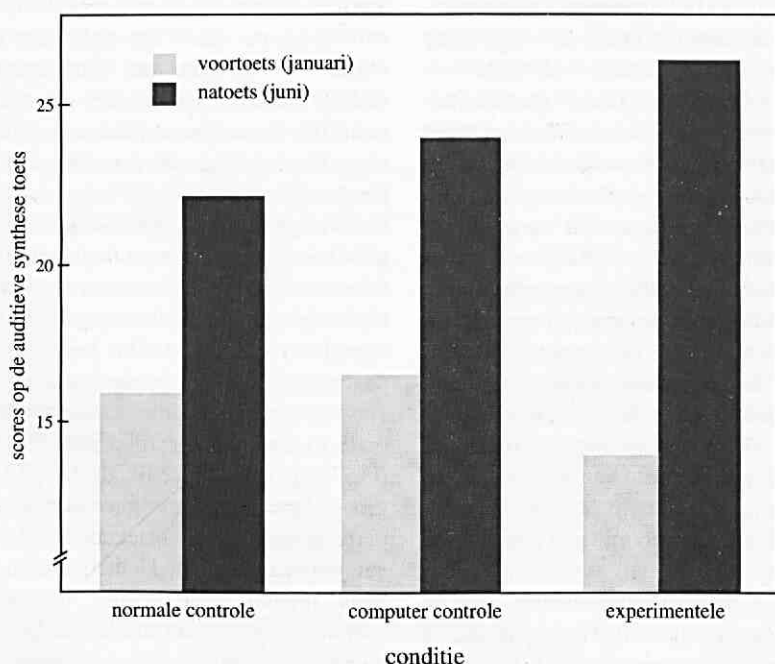
Oefengegevens

In totaal hebben de leerlingen in de experimentele conditie (EC) gemiddeld 19.9 maal geoefend. Omdat de sessietijd op 10 minuten is gesteld, is er dus gemiddeld 3 uur en 20 minuten aan de oefeningen aan de computer besteed. Voor de CC-groep die de controle-oefeningen op de computer uitvoerden was het gemiddeld aantal sessies 17.9 (bijna 3 uur).

Uiteraard werkten niet alle kleuters even snel. De voortgang in het vastgestelde pakket van oefeningen varieerde. Hoewel het aanbod in principe voor alle leerlingen hetzelfde is geweest, zijn er door omstandigheden, zoals tijdelijke afwezigheid wegens ziekte, snelheid

van werken, afleidbaarheid tijdens de oefeningen, en dergelijke, verschillen tussen leerlingen opgetreden in het aantal oefeningen dat ze werkelijk hebben verricht. Sommigen deden langer over de beginoefeningen, zodat zij ook pas op een later moment in aanraking kwamen met de specifieke oefeningen. Er is op grond van de automatisch opgeslagen voortgangsgegevens per leerling bepaald hoe vaak elke leerling dergelijke specifieke leermomenten heeft meegemaakt. Per leerling is geteld hoe vaak er een opdracht is verwerkt waarin een gesegmenteerd woord gesynthetiseerd moest worden of waarin woorden moesten worden vergeleken die slechts in één foneem verschilden. Dus de niet specifiek op synthese gerichte oefeningen (bijvoorbeeld de eerste drie varianten van de woordenschatoefening) werden hierbij buiten beschouwing gelaten. Het totaal aantal vervulde specifieke opdrachten is voor de EC groep gemiddeld 210 ($sd = 124$). Ter vergelijking, het daadwerkelijk aantal gelegde associaties tussen een woordklank en een plaatje en het aantal woordvergelijkingen in de CC groep was over de hele periode gemiddeld 489 ($sd = 204$).

In het algemeen waren de reacties van alle leerlingen uitermate positief over het gebruik van het programma, zowel van leerlingen in de



Figuur 2. De gemiddelde scores op de synthesesoetsen tijdens de voortoets en de natoets voor de drie verschillende onderzoeksgroepen

experimentele groep als van de leerlingen in de controle-conditie. Het niveau van de oefeningen en de eisen die aan de leerlingen werden gesteld bleken niet te hoog en het gemiddeld aantal direct goede responsen tijdens de oefeningen was meestal ruim boven de 80%. De motivatie bleek voor vrijwel alle leerlingen op een hoog peil en dit bleef zo tot het moment dat de oefeningen werden afgebroken en de natoets werd afgenomen. Alleen enkele kinderen uit de computer-controle groep gaven aan dat ze graag ook andere oefenvormen zouden willen doen, vooral omdat de leerlingen uit de experimentele groep van hun enthousiasme blij gaven en de controle-leerlingen deze andere oefenvormen bij hun klasgenoten zagen.

Effect van de training en verschil tussen groepen

De resultaten op de voor- en natoets voor auditieve synthese geven aan dat de scores gemiddeld flink vooruit zijn gegaan (zie Tabel 3). Het verschil is statistisch significant ($F(1,150) = 122.57, p < 0.001$). Terwijl het gemiddelde op de voortoets in januari 15.9 ($sd = 9.1$) was, behaalden deze oudste kleuters op de parallel toets voor auditieve synthese in juni een gemiddelde van 23.9 ($sd = 13.5$).

De vooruitgang is echter niet voor alle drie condities gelijk. Er blijkt een significante interactie te bestaan tussen de mate van vooruitgang en de experimentele condities ($F(2,150) = 4.84, p < .01$). De effectgrootte η^2 (partiële etakwadraat) is 0.06 en kan volgens Cohen (1988, p. 26) als heel behoorlijk worden beschouwd. De gemiddelde scores voor de drie groepen op de voortoets en de natoets staan weergegeven in Figuur 2. In de figuur is duidelijk te zien dat het verschil tussen de toetsen het grootst is voor de experimentele trainingsgroep. Voor de twee andere groepen is er ook een verschil, maar de vooruitgang is beduidend kleiner. Contrast analyses tonen aan dat de normale controle-groep en de computer controlegroep onderling niet verschillen in de mate van vooruitgang in synthesevaardigheid, terwijl de experimentele groep significant meer vooruitgaat dan de twee controlegroepen.

De Pearson correlatiecoëfficiënten voor de samenhang tussen enkele relevante variabelen staan in Tabel 2 weergegeven. Uit deze tabel blijkt dat de onderlinge verschillen in leeftijd slechts een lichte samenhang vertonen met

woordherkenning, letterkennis en synthesevaardigheid tijdens de voortoets. De scores op de toets voor woordenschat corresponderen in iets sterkere mate met woordherkenning, letterkennis en synthesevaardigheid zowel tijdens de voortoets als de natoets. In deze tabel valt verder op dat vooral woordherkenning, letterkennis en de beide synthese-toetsen onderling samenhangen. De leerlingen binnen deze groep van 153 kleuters die enkele woorden herkennen en letters kennen, behalen veelal ook hogere scores op de synthesesoetsen. Belangrijk voor dit onderzoek is echter dat de mate van vooruitgang in synthesevaardigheid niet samenhangt met het beginniveau op de synthesesoets, of met leeftijd, woordenschat en woordherkenning. Slechts het aantal letters dat een leerling (in januari) kent vertoont een lichte positieve samenhang met de vooruitgang in synthesevaardigheid: kinderen die relatief meer letters kennen, gaan iets meer vooruit in synthesevaardigheid. De gevonden verschillen tussen condities blijven gelijk of worden enigszins sterker als analyses worden verricht met enkele voortoetsgegevens als covariabele. Hierdoor blijkt vooral een gedeelte van de binnengroep variantie te worden verklaard. Bijvoorbeeld, een analyse met leeftijd (in maanden), de scores op de woordherkenningsoets en letterkennis, geeft een significant conditie-effect voor de mate van vooruitgang van $F(2,147) = 6.46, p < .01$ en een effectgrootte η^2 van 0.08. Vooral letterkennis en woordherkenning dragen bij aan de vermindering van binnengroepsvariantie.

Overigens is in Tabel 2 ook te zien dat er geen samenhang is tussen het beginniveau van de scores op de synthesesoets en de mate van vooruitgang; de correlatie van 0.05 wijkt niet significant van nul af. Ook berekeningen van de correlatie tussen beginniveau en de mate van vooruitgang voor de kleuters binnen elke onderzoeksconditie afzonderlijk (0.13, -0.07, 0.15 voor respectievelijk de GC, CC en EC groep) laten zien dat er geen noemenswaardig verband bestaat. Dat betekent dus dat de mate van vooruitgang voor kleuters die op de voortoets relatief laag scoren bij benadering evenveel vooruitgaan als kinderen die op de voormeting al blijf geven over enige synthesevaardigheid te beschikken.

Tabel 2

Correlatiecoëfficiënten (Pearson) voor de samenhang tussen enkele voortoetsgegevens, score op de synthese-natoets en de mate van vooruitgang

	Leeft	WrdSchat	WrdHerk	LetKen	Synth-voor	Synth-na
WrdSchat	-.01					
WrdHerk	.21**	.26**				
LetKen	.18*	.36**	.68**			
Synth-voor	.18*	.40**	.69**	.63**		
Synth-na	.15	.37**	.54**	.59**	.71**	
Vooruitgang	.04	.14	.10	.24**	.05	.74**

* - Signif. <.05 **-Signif.<.01 (twee-zijdig)

Effect van leerkrachtactiviteiten

Zoals eerder gemeld (onder 2. Methode), bleek uit de analyses van de gegevens over de thuisachtergrond en de klasse-situatie dat er tussen de drie thans onderscheiden onderzoeksgroepen geen belangwekkende verschillen waren ten aanzien van deze gegevens. Eén variabele uit de klasse-situatie bleek echter met synthesevaardigheid van de leerlingen samen te hangen en dit betrof het moment waarop in de groep oudste kleuters gestructureerd oefeningen werden begonnen in auditieve analyse en synthese. De leerkrachten werd gevraagd of en op welk moment ze systematisch aan gezamenlijke activiteiten in het ontleden van woordklanken en het samenstellen van woordgehelen waren begonnen. Van de 12 groepen bleken 4 groepen al vanaf september met enige regelmaat met deze zogenaamde 'hak-en-plak' oefeningen bezig te zijn, de anderen hebben hieraan gedurende het voorjaar wel eens gewerkt of zijn er niet aan toe gekomen. De kleuters in de groepen die reeds vanaf het begin van het schooljaar hieraan aandacht hadden besteed bleken hoger te scoren op de synthese voortoets (16.3) dan de kleuters uit groepen waar deze activiteiten niet of nauwelijks werden ontplooid (15.7),

maar dit verschil was statistisch niet significant. Van belang is echter dat er wel een verschil tussen deze groepen is in de mate van vooruitgang in synthesevaardigheid. De kleuters in de groepen waarin door de leerkracht auditieve structureringsoefeningen werden uitgevoerd, gingen gemiddeld 13.3 punten vooruit op de synthesetoets, terwijl de anderen gemiddeld 6.3 punten vooruitgingen, een significant verschil ($F(1,151) = 16.75, p < .01$). Overigens behoorden de 4 groepen, waarin de leerkracht op gezette tijden de analyse en synthese van klanken oefende, tot de (7) groepen waarin ook de computerprogramma's werden gebruikt. Het is daarom goed om hier nogmaals op te merken dat de twee computercondities (experimentele training en controle) uit dezelfde groepen afkomstig waren. De invloed van de verschillen tussen groepsactiviteiten zal daarom naar verwachting tamelijk beperkt zijn ten aanzien van de effecten van de verschillende oefencondities op de computer.

In het tweede deel van Tabel 3 zijn de gegevens van de CC en EC condities opgesplitst naar het wel of niet voorkomen van leerkrachtactiviteiten met betrekking tot auditieve structureringsoefeningen. De resultaten van een

Tabel 3

Gemiddelde scores (en standaard deviaties tussen haakjes) op de voor- en natoets voor synthesevaardigheid in de verschillende condities

conditie	n	voortoets	natoets	verschil
GC	69	16.1 (8.5)	22.3 (13.1)	6.2 (8.9)
CC	45	16.6 (10.4)	24.1 (13.5)	7.5 (9.4)
EC	39	14.5 (8.7)	26.4 (14.0)	11.9 (9.7)
+/- wel of niet auditieve structureringsoefeningen door leerkracht				
CC -	20	14.8 (9.3)	19.2 (11.5)	4.5 (8.5)
EC -	22	15.5 (9.1)	24.5 (12.9)	8.9 (8.3)
CC +	25	18.9 (11.5)	30.2 (13.6)	11.3 (9.2)
EC +	17	13.2 (8.4)	28.9 (15.3)	15.7 (10.3)

variantie-analyse voor het effect van condities op de mate van vooruitgang in synthesevaardigheid binnen de CC en EC condities met het al of niet groepsgewijs uitvoeren van hak-en-plak activiteiten als covariable bleek geheel overeen te komen met de eerder gerapporteerde conditie-effecten. Naast het zeer duidelijke regressie-effect van de groepsactiviteiten ($F(1,81) = 11.93, p < .01$; effectgrootte $\eta^2 = .13$), bleek het effect van het al of niet ontvangen van synthese-oefeningen op de computer ook significant ($F(1, 81) = 5.16, p = .03$; effectgrootte $\eta^2 = .06$). Uit de gemiddelde verschillen in de laatste kolom van Tabel 3 is te zien dat de effecten van leerkrachtactiviteit en van de computeroefening additief zijn; het verschil tussen de EC en CC conditie is bij benadering even groot voor de groepen waarin wel of niet door de leerkracht analyse- en synthese-oefeningen worden aangeboden.

Verschillen in oefenintensiteit

De mate waarin de leerlingen daadwerkelijk aan de computer geoefend hebben was verschillend. Hoewel het aanbod in principe voor alle leerlingen hetzelfde is geweest, zijn er door omstandigheden, zoals tijdelijke afwezigheid wegens ziekte, snelheid van werken, afleidbaarheid tijdens de oefeningen, en dergelijke, verschillen tussen leerlingen opgetreden in het aantal oefeningen dat ze werkelijk hebben verricht. Als de synthese-oefeningen op de computer tot gevolg hebben dat er een grotere vooruitgang is in synthesevaardigheid zoals gemeten door het verschil op de voor- en natoets, dan kan worden verwacht dat naarmate een leerling meer geoefend heeft er ook meer vooruitgang is geboekt. De correlatie tussen het aantal specifieke opdrachten en de winst in synthesevaardigheid was binnen de experimentele groep .44. Binnen de computer-controlegroep bleek de correlatie tussen het aantal vervulde opdrachten en de vooruitgang in

synthesevaardigheid .18 te zijn. Deze secundaire analyse geeft dus aanvullende ondersteuning voor de gedachte dat de synthesevaardigheid binnen de experimentele groep inderdaad door de oefeningen aan de computer positief is beïnvloed.

Effecten na 3 maanden

Een inventarisatie van de scores op de toets voor technisch lezen in de eerste helft van november in groep 3 (zie Tabel 4) laat zien dat de EC groep gemiddeld de hoogste score haalt. Om twee redenen is het aantal betrokken leerlingen hier kleiner. Ten eerst bleek een tamelijk groot aantal leerlingen van school te zijn veranderd. Ten tweede is besloten om slechts de gegevens te verwerken van leerlingen die in de voortoets niet meer dan twee woorden correct konden herkennen en niet meer dan 12 letters bleken te kennen. De resultaten en uitkomsten van analyses betreffende de synthesevaardigheid van de thans geselecteerde groep bleken overigens geheel overeen te komen met de boven weergegeven gegevens van de totale groep.

De CC en GC groep halen gemiddeld lagere decodeerscores (zie Tabel 4). Maar uit de gegevens in de Tabel blijkt ook dat de spreiding van scores tamelijk hoog is, vooral in de EC groep. Een variantie-analyse geeft aan dat er geen significant verschil is tussen de EC groep en de andere twee groepen ($F(1, 95) = 2.48, p = .12, \eta^2 = .025$).

De gegevens op de toets voor klanksynthese laten eveneens zien dat de EC groep gemiddeld de hoogste score haalt, gevolgd door de CC groep en daarna met een duidelijk lagere score de GC groep. Een variantie-analyse levert een significant effect van conditie ($F(2,95) = 6.89, p < .01$). Maar het verschil is vooral tussen GC en de andere twee condities. Overigens, de gemiddelde score van ongeveer 13 goed gesynthetiseerde CVC woorden van de 20 aangebo-

Tabel 4
Toetsresultaten per onderzoeksconditie in groep 3 na 3 maanden leesonderwijs

		EC	CC	GC
aantal leerlingen		25	28	45
decodeervaardigheid	<i>gemiddeld</i>	22.2	17.3	18.3
	<i>(sd)</i>	(16.9)	(9.5)	(10.3)
klanksynthese-score	<i>gemiddeld</i>	13.9	12.5	10.1
	<i>(sd)</i>	(3.6)	(4.4)	(4.6)

Tabel 5

Toetsresultaten per onderzoeksconditie in groep 3 na 5 maanden leesonderwijs

		EC	CC	GC
decodeervaardigheid	<i>gemiddeld</i>	13.4	10.6	10.1
	<i>(sd)</i>	(6.1)	(4.8)	(4.8)
klanksynthese-score	<i>gemiddeld</i>	45.0	42.7	39.8
	<i>(sd)</i>	(5.3)	(8.3)	(11.0)

den levert in de EC groep een percentage correct van 65 op, terwijl bijvoorbeeld het percentage correct op de natoets in juni voor de 5 CVC woorden in dezelfde groep leerlingen gemiddeld 33% was. De synthesevaardigheid is in de tussentijd dus nog aanzienlijk toegenomen.

Effecten na 5 maanden

Tot slot, voor dezelfde leerlingen worden de gemiddelde leesprestaties op de in januari afgenomen Eén-minuuttest en de gemiddelde scores op de laatste synthesesoets in Tabel 5 vermeld. Een statistisch significant verschil tussen de groepen werd geconstateerd voor de leescores ($F(2,95) = 3.42, p = .04$). De experimentele groep verschilde zowel van de computer-controle groep ($F(1, 95) = 3.87, p = .05$; effectgrootte $\eta^2 = .039$) als van de gewone controlegroep ($F(1, 95) = 6.47, p = .01$; effectgrootte $\eta^2 = .064$).

De verschillen in synthesevaardigheid tussen de groepen zijn nog steeds aanwezig. Maar toetsing leverde geen statistisch effect op gebruikelijk significantieniveau ($F(2, 95) = 2.85, p = .06$). Het verschil tussen EC en GC was wel significant ($F(1, 95) = 5.43, p = .02$).

5 Discussie

De resultaten laten ten eerste zien dat kleuters gedurende het laatste halfjaar van groep 2 duidelijk vooruitgaan in auditieve synthesevaardigheid. Deze ontwikkeling kan te maken hebben met een algemene toename in linguïstische competentie. Het proces van taalverwerving, het toenemend inzicht in en kennis van de taal, is in de kleuterleeftijd nog in volle gang. Door een voortdurend en in de kleuterfase meestal geïntensiveerd en gevarieerd taalaanbod neemt geleidelijk de vaardigheid op allerlei terreinen van taal bij kleuters toe. Door middel van taal-

spelletjes, liedjes, werkbladen, en dergelijke wordt in toenemende mate ook aandacht gevraagd voor de vormaspecten van de taal. Ook in het huidige onderzoek bleek de stimulerende invloed van door leerkrachten geïnitieerde taaloefeningen tijdens kringactiviteiten in groep 2 onmiskenbaar aanwezig. Zo kwam in ons onderzoek naar voren dat de leerlingen in groepen waarin gedurende het gehele schooljaar aandacht werd besteed aan het analyseren van woorden en het synthetiseren van woorddelen (het zogenaamde 'hakken-en-plakken') ten opzichte van andere leerlingen (in andere groepen) het meest vooruitgaan in synthesevaardigheid.

Maar zoals blijkt uit de resultaten van dit onderzoek, het oefenen met een computerprogramma dat is gericht op de bevordering van de vaardigheid tot auditieve synthese heeft duidelijk een aanvullend effect op de gebruikelijke ontwikkeling van synthesevaardigheid. De leerlingen die in het kader van het onderzoek met enige regelmaat aan de computer enkele specifieke oefeningen hebben verricht gaan duidelijk meer vooruit in synthesevaardigheid dan de twee andere groepen. De groep leerlingen die ongeveer even frequent oefeningen doet in het leggen van associaties tussen plaatjes en gesproken woorden én de groep leerlingen die geen bijzondere taken aan de computer vervult en slechts het reguliere kleuteronderwijs in hun groepen volgt, gaan beide minder vooruit in synthesevaardigheid. Het positieve effect van de experimentele training doet zich voor in alle betrokken kleutergroepen, zowel in groepen waarin niet of nauwelijks door de leerkracht expliciet aandacht wordt gevraagd voor de klankstructuur van gesproken woorden als in groepen waarin met enige regelmaat hieraan aandacht wordt besteed door middel van allerlei spelletjes. Het oefenen met de computer heeft klaarblijkelijk een toegevoegde waarde.

De experimentele oefengroep vertoont een grotere vooruitgang in synthesevaardigheid.

De reden hiervoor kan gezocht worden in het feit dat de leerlingen tamelijk aandachtig bezig zijn geweest met het samenstellen van woordklanken op basis van aangeboden segmenten. In vergelijking met kringactiviteiten is het voordeel van het computerprogramma dat op individuele basis vrij intensief kan worden geoefend. Het programma bleek niet te moeilijk voor deze jonge leerlingen, er waren duidelijke opdrachten en bij elke actie kwam ogenblikkelijk een passend commentaar. Hiermee werd een betrekkelijk optimale leeromgeving aangeboden. De tijd die doorgebracht werd aan de computeroefeningen werd doorgaans goed en effectief besteed. Bovendien is er in vergelijking met groepsactiviteiten relatief weinig storing en kunnen zowel oplettende als snel afleidbare leerlingen de oefeningen geheel zelfstandig doorlopen. Omdat het een bijdrage vormt aan het vergroten van de effectieve leertijd, kan het gebruik van het programma daarom als één van de ingrediënten voor een effectieve school worden beschouwd (Creemers, 1991). De gegevens laten zien dat een op het terrein van klankstructureringsoefeningen actieve leerkracht minstens zoveel zoden aan de dijk zet.

Een onderlinge vergelijking tussen de effecten van leerkracht geïnitieerde oefeningen en de computeroefeningen geeft aan dat de leerkracht een ongeveer tweemaal zo groot effect lijkt te hebben. De conclusie dringt zich bovendien op dat een combinatie het meest vruchtbaar is. De gegevens zijn althans met een dergelijke gedachtengang in overeenstemming. Het is evenwel noodzakelijk te bedenken dat het huidige onderzoek niet expliciet was gericht op het evalueren van het effect van leerkrachtactiviteiten op de ontwikkeling van synthesevaardigheid bij kleuters. De huidige onderzoeksopzet geeft geen goede gelegenheid om deze vraag te beantwoorden. Het aantal betrokken leerkrachten is te klein en de methode die gebruikt is om na te gaan welke activiteiten werden ontplooid was duidelijk voor deze doeleinden te beperkt. Op grond hiervan kunnen uit het onderzoek daarom slechts zeer voorlopige conclusies worden getrokken over de rol van de leerkracht.

Is nu te verwachten dat de leerlingen die de syntheseoefeningen op de computer hebben gedaan hier profijt van zullen hebben bij het

aanvankelijk lezen? Het voordeel van de oefeningen gedurende de kleutertijd in groep 2 is dat alvast aan één van de essentiële leesvoorwaarden is gewerkt. De resultaten lijken inderdaad uit te wijzen dat syntheses training in de kleuterperiode zeker een faciliterend effect kan hebben op de eerste vorderingen in leesvaardigheid. De leerlingen die als experimentele groep de computeroefeningen in klanksynthese hebben verricht halen na ongeveer 10 weken leesinstructie gemiddeld de hoogste score op een klanksynthese toets én de hoogste leesscore. Hetzelfde kan worden geconstateerd na ongeveer 18 weken onderwijs. Het is opmerkelijk dat deze effecten, hoewel bepaald niet groot, reeds kunnen worden geconstateerd na ongeveer drie maanden leesonderwijs en tamelijk stabiel blijken. Want na ongeveer vijf maanden onderwijs in lezen blijken de verschillen tussen de groepen statistisch significant. Er kan dus worden gesproken van een tamelijk langdurige invloed van de computerprogramma's. Deze langdurige doorwerking was betrekkelijk onverwacht omdat in een kwalitatief goede instructie in aanvankelijk lezen onder andere volop aandacht wordt besteed aan de ontwikkeling van fonologische vaardigheden. De verwachting was dat directe instructie in fonologische vaardigheden ingebed in de instructies en beginoefeningen in lezen de voorgaande effecten van oefeningen in klanksynthese geheel zouden overschaduwden. De toename in synthesevaardigheid was voor zover we konden nagaan ook bijzonder groot.

Hoewel de gemiddelde scores op een positioneel effect lijken te wijzen, was er voor de decodeerscores in begin november geen statistisch significant verschil tussen de onderzoekscondities. Hoewel de verschillen in januari statistisch wel significant zijn, kan over de grootte van het effect verschillend worden gedacht. Het is in dit verband goed te bedenken dat decodeervaardigheid uiteraard door meer factoren wordt bepaald dan de mate waarin klanksynthese wordt beheerst. Kennis van letters, woordenschat, interesse in lezen en schrijven, maar ook de aard en kwaliteit van de instructie speelt een rol. In het huidige onderzoek is geen informatie ingewonnen omtrent dergelijke medebepalende factoren. Maar als de vorderingen in leesvaardigheid door zoveel verschillen-

de factoren bepaald kunnen worden, dan zal het ook praktisch uiterst lastig zijn om precies te bepalen hoe groot het voordeel van oefening in klanksynthese tijdens groep 2 is op de vorderingen tijdens de eerste paar maanden van het aanvankelijk leesonderwijs. Voor een echt goede vergelijking zullen voor zowel vooraf getrainde leerlingen als ongetrainde leerlingen in een longitudinaal onderzoek de beginfasen van het leerproces, de verstrekte leesinstructie en andere van belang zijnde omstandigheden heel gedetailleerd in kaart moeten worden gebracht.

Overigens gaat een vergelijking met onderzoeksresultaten in andere landen vaak gebukt onder het probleem dat het leesonderwijs niet op dezelfde manier is ingericht zoals dat zich in hoofdlijnen in ons land voltrekt. In veel Angelsaksische landen wordt bijvoorbeeld het leren decoderen veel minder benadrukt. Dit kan tot gevolg hebben dat de individuele verschillen in fonologische vaardigheid, voordat met het leesonderwijs is begonnen, nog veel langer doorwerken tijdens de eerste fasen van het leren lezen. Als de instructie in de beginfasen van leren lezen niet expliciet gericht is op het aankweken van fonologische vaardigheden, dan zullen de oorspronkelijke verschillen een langerdurende invloed hebben. Het is in deze context ook begrijpelijk dat voorschoolse oefening in fonologische vaardigheden een duidelijk effect heeft op de verwerving van schriftelijke taalvaardigheden. Het is nog maar de vraag of dat in de Nederlandse situatie ook zo algemeen gevonden zal worden.

De didactische vormgeving van de oefeningen voor fonologische vaardigheden behoeft ook nog nadere aandacht. In het hier gerapporteerde onderzoek is gebruik gemaakt van getekende oren als wijze van representatie van de klanken: grote oren verwijzen naar hele woorden en kleine oortjes naar segmenten of foneemklanken. Men kan twisten over de vraag of deze visuele representatie de meest vruchtbare is, dat wil zeggen, of jonge kinderen hiermee het meest optimaal leren omgaan met spraaksegmenten. Een belangrijk nadeel is dat de oren onderling niet verschillen, terwijl ze toch corresponderen met verschillende klanken. Deze uniforme representatie zou wellicht beter vervangen kunnen worden door een symboolgebruik waarin de unieke relatie tussen

symbool en klank beter tot uitdrukking komt. De keuze voor lettertekens ligt dan voor de hand. Onderzoeksresultaten lijken uit te wijzen dat het gebruik van lettertekens inderdaad voordelen biedt, hoewel er ook studies zijn verricht waaruit blijkt dat hierover het laatste woord nog niet is gezegd (Hohn & Ehri, 1983; Bus, 1985; Kerstholt, 1995).

Zoals in de inleiding beargumenteerd is in deze studie gekozen om de vaardigheid in auditieve synthese te oefenen. De synthesevaardigheid kan worden beschouwd als een elementaire deelvaardigheid voor beginnend lezen. Maar synthesevaardigheid is onmiskenbaar gerelateerd aan analysevaardigheid. Beide vaardigheden doen een beroep op het vermogen met onderdelen van gesproken woorden om te gaan. In het algemeen worden dan ook hoge correlaties gevonden tussen de scores op toetsen voor auditieve analyse en auditieve synthese. Inhoudelijk en didactisch lijkt er weliswaar een verschillende rol voor analyse en synthese te zijn weggelegd. De eerste is meer van belang voor spellen en de tweede voor lezen. Maar aanvankelijk leesonderwijs is doorgaans tegelijk onderwijs in beginnend spellen. Het leren lezen en spellen is in het eerste leerjaar geheel geïntegreerd. Vanuit dit perspectief lijkt het wenselijk om ook in de computerondersteunde oefeningen aan auditieve analyse aandacht te schenken. We hebben weliswaar inmiddels een aantal oefenvormen geïmplementeerd waarmee het beginnend spellen geoefend kan worden, maar een goede oefening in auditieve analyse is lastig te ontwerpen, vooral wanneer de oefening geschikt moet zijn voor kleuters die de lettertekens nog niet hebben geleerd. De reden is dat er (nog) geen mogelijkheid bestaat om een computer zover te brengen dat de uitingen van een leerling die bezig een woord in stukjes uit te spreken op juiste wijze te verstaan. Slechts het aanbod van goed verstaanbare losse foneemklanken is technisch te realiseren. Maar ook hier gelden beperkingen. Vooral bij de consonanten is verwarring bijna niet uit te sluiten. Het verschil tussen een /m/ en een /n/ bijvoorbeeld is, indien deze klanken geïsoleerd worden aangeboden, heel moeilijk te horen. In onze oefeningen hebben we dit opgelost door dergelijke verwarbare alternatieven nooit tegelijk te presenteren. Voor een goede discriminatie tussen sommige spraakklanken is enig zicht

op de articulatiebewegingen onontbeerlijk (het zogenaamde spraakafzien). Slechts aan de stand van de lippen is het onderscheid tussen losse /m/ en /n/ klanken mogelijk. Hoewel er wordt geëxperimenteerd met computerprogramma's waarin schetsmatig de articulatieorganen worden afgebeeld, lijkt een dergelijke aanpak vooralsnog onnodig veel nieuwe problemen op te roepen. Bij problemen in auditieve analyse en moeilijkheden om bepaalde spraakklanken goed te discrimineren zal de leerkracht een nog onvervangbare en niet te imiteren rol moeten vervullen.

Tot slot, dit onderzoek heeft nogmaals aangetoond dat computer-ondersteunde oefeningen een goede bijdrage kunnen leveren aan het arsenaal van didactische middelen ten behoeve van aanvankelijk leesonderwijs (zie o.a. ook Van Daal & Reitsma, 1993; Yap & Van der Leij, 1993). De mogelijkheid om voor elk individu op maat gesneden oefeningen aan te bieden en de intensieve interactie tussen leerling en computer waarin op elke actie van een leerling ogenblikkelijk passend commentaar volgt, geven de kans om met betrekkelijk weinig inspanningen van de kant van de leerkracht de leerlingen leerzame oefeningen te laten verrichten. Het huidige onderzoek heeft aangetoond dat ook voor jonge kinderen dergelijke oefeningen een gunstig resultaat kunnen hebben. Computer-ondersteunde oefeningen in auditieve synthese hebben bij kleuters een positief effect.

Literatuur

- Alegria, J., Pignot, E., & Morais, J. (1982). Phonetic analysis of speech and memory codes in beginning readers. *Memory & Cognition*, 10, 451-456.
- Ball, E. W., & Blachman, B.A. (1988). Phoneme segmentation training: Effect on reading readiness. *Annals of Dyslexia*, 38, 208-225.
- Bast, J.W. (1995). *The development of individual differences in reading ability*. Academisch proefschrift Vrije Universiteit Amsterdam. Duivenrecht: Paedologisch Instituut.
- Bentin, S., Hammer, R., & Cahan, S. (1991). The effects of aging and first grade schooling on the development of phonological awareness. *Psychological Science*, 2, 271-274.
- Blachman, B.A. (1984). Language analysis skills and early reading acquisition. In G. Wallach & K. Butler (Eds.), *Language learning disabilities in school-age children* (pp. 271-287). Baltimore: Williams & Williams.
- Bradley, L., & Bryant, P.E. (1983). Categorising sounds and learning to read: a causal connection. *Nature*, 301, 419-421.
- Brus, B.Th., & Voeten, M.J.M. (1973). *Eén-minuut-test, vorm A en B*. Nijmegen: Berkhout.
- Bus, A.G. (1985). Voorbereidend leesonderwijs aan kleuters: Een vergelijkend onderzoek naar een auditief en auditief-visueel oefenprogramma van auditieve analyse en synthese. *Pedagogische Studiën*, 62, 269-279.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Cunningham, A.E. (1990). Explicit versus implicit instruction in phonemic awareness. *Journal of Experimental Child Psychology*, 50, 429-444.
- Creemers, B.P.M. (1991). *Effectieve instructie. Een empirische bijdrage aan de verbetering van het onderwijs in de klas*. Den Haag: SVO.
- Daal, V.H.P. van, & Reitsma, P. (1993). The use of speech feedback by normal and disabled readers in computer-based reading practice. *Reading and Writing: An Interdisciplinary Journal*, 5, 243-259.
- Dongen, D. van (1984). *Leesmoeilijkheden*. Tilburg: Zwijzen.
- Elkonin, D.B. (1973). USSR. In J. Downing (Ed.), *Comparative reading: Cross-national studies of behaviour and processes in reading and writing*. New York: Macmillan.
- Hohn, W.E., & Ehri, L.C. (1983). Do alphabet letters help prereaders acquire phonemic segmentation skill? *Journal of Educational Psychology*, 75, 752-762.
- Juel, C. (1991). Beginning reading. In R. Barr, M.L. Kamil, P.B. Mosenthal & P.D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 759-788). New York: Longman.
- Kerstholt, M.T. (1995). *The phonological skills of blending and segmentation: effects of training types*. Academisch proefschrift, Katholieke Universiteit Nijmegen.
- Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in pre-school children. *Reading Research Quarterly*, 23, 263-284.

- Lundberg, I., Olofsson, A., & Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology*, 21, 159-173.
- Mattingly, I.G. (1972). Reading, the linguistic process, and linguistic awareness. In J.F. Kavanagh & I.G. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading* (pp.133-147). Cambridge, MA: MIT Press.
- Mommers, M. (1990). Metalinguistic awareness and learning to read. In P. Reitsma & L. Verhoeven (Eds.), *Acquisition of reading in Dutch* (pp. 29-42). Dordrecht: Foris Publications.
- Morrison, F.J., Smith, L., & Dow-Ehrensberger, M. (1995). Education and cognitive development: a natural experiment. *Developmental Psychology*, 31, 789-799.
- Olofsson, Å., & Lundberg, I. (1983). Can phonemic awareness be trained in kindergarten? *Scandinavian Journal of Psychology*, 24, 35-44.
- Olofsson, Å., & Lundberg, I. (1985). Evaluation of long-term effects of phonemic awareness training in kindergarten: Illustrations of some methodological problems in evaluation research. *Scandinavian Journal of Psychology*, 26, 21-34.
- Olofsson, Å. (1993). The relevance of phonological awareness in learning to read: scandinavian longitudinal and quasi-experimental studies. In R.M. Joshi & C.K. Leong (Eds.), *Reading disabilities: Diagnosis and component processes* (pp. 185-198). Dordrecht: Kluwer.
- Schneider, W., & Näslund, J.C. (1993). The impact of early metalinguistic competences and memory capacity on reading and spelling in elementary school: Results of the Munich Longitudinal Study on the Genesis of Individual Competencies (LOGIC). *European Journal of Psychology of Education*, 8, 273-287.
- Schonewille, B., & Leij, A. van der (1995). De rol van de leerkracht voor allochtone en autochtone leerlingen in het eerste jaar van het basisonderwijs. *Pedagogische Studiën*, 72, 242-257.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-406.
- Verhoeven, L., & Vermeer, A. (1986). *Taaltoets allochtone kinderen*. Tilburg: Zwijsen.
- Wallach, M.A., & Wallach, L. (1979). Helping disadvantaged children learn to read by teaching them phoneme identification skills. In L.A. Resnick & P.A. Weaver (Eds.), *Theory and practice of early reading* (Vol. 3, pp. 227-259). Hillsdale, NJ: Erlbaum.
- Williams, J.P. (1980). Teaching decoding with an emphasis on phoneme analysis and phoneme blending. *Journal of Educational Psychology*, 72, 1-15.
- Yap, R., & Leij, A. van der (1993). Computergestuurde remediëring van dyslexie door het opvoeren van de herkenningssnelheid van subwoordeenheden. *Pedagogische Studiën*, 70, 402-419.

Manuscript aanvaard 22-1-1996

Auteurs

P. Reitsma is hoofd van de afdeling Onderzoek Lees- en Leerproblemen van het Paedologisch Instituut en als bijzonder hoogleraar verbonden aan de Faculteit Psychologie en Pedagogiek van de Vrije Universiteit.

R. Wesseling is als onderzoeker-in-opleiding verbonden aan het Paedologisch Instituut.

F. Stiva is als multi-media programmeur werkzaam op het Paedologisch Instituut.

Correspondentieadres: Paedologisch Instituut, Postbus 303, 1115 ZG Duivendrecht

Abstract

Training of blending skills in kindergarten children by using computer-assisted exercises

P. Reitsma, R. Wesseling & F. Stiva. *Pedagogische Studiën*, 1997, 74, 1-20.

In an empirical study the effect of computer-assisted training in sound blending is evaluated in a group of kindergarten children. A package of experimental programs was used in which word materials, instructions and comments were provided in high quality digitised speech. Children in the experimental condition received specific training in blending separate letter sounds into words during a 4 month period, while other classmates received training in vocabulary using the same computers and program. Yet another control group from separate classes did not have access to the computer programs. Both

before and after the 4 month period a test for blending skill was administered. All children appeared to improve in blending skill. However, the results revealed a significant additional effect due to the use of the computer program. Positive long-term effects of these exercises on the progress in reading development in Grade 1 were obtained.

Onderzoek naar bias voor allochtone leerlingen in de Cito-Eindtoets Basisonderwijs

H. Uiterwijk en T. Vallen

Samenvatting

Er is nog relatief weinig onderzoek verricht naar de bruikbaarheid van toetsen voor subgroepen uit de populatie. Het Cito en het Werkverband Taal en Minderheden van de Letterenfaculteit van de KUB verrichten onderzoek naar toets- en itembias in de Cito-Eindtoets Basisonderwijs voor allochtone leerlingen. Uit het onderzoek naar toetsbias blijkt dat het voorspellen van schoolsucces in het voortgezet onderwijs voor allochtone leerlingen minder trefzeker gebeurt dan voor autochtone leerlingen. Dit geldt zowel voor de Eindtoets Basisonderwijs als voor een onderzochte intelligentietest. Het onderzoek naar itembias laat zien dat de statistische procedures voor het opsporen van itembias in de Eindtoets Basisonderwijs verschillende uitkomsten opleveren. De meeste items blijken nooit, andere soms en enkele items blijken altijd gebiast te zijn. Een zoektocht naar de inhoudelijke oorzaken van itembias levert vooral op linguïstisch terrein gegevens op waarmee het risico op itembias in toekomstige toetsen kan worden verkleind.

1 Inleiding en probleemstelling

Het is gebruikelijk om de schoolprestaties van leerlingen te beschrijven aan de hand van scores op een of andere schoolvorderingstoets. Daarbij wordt veelvuldig gerapporteerd over de verschillen tussen de scores van bepaalde groepen in de populatie. In Nederland is er tot nu toe relatief weinig onderzoek gedaan naar de vraag of een toets die bruikbaar geacht wordt voor de hele populatie ook bruikbaar is voor daarbinnen te onderscheiden subgroepen. Als er scoreverschillen tussen bepaalde subgroepen geconstateerd worden, kan dat veroorzaakt worden door het feit dat die subgroepen de vaardigheid die de toets beoogt te meten in uiteenlopende mate beheersen. Scoreverschil-

len tussen subgroepen zijn op zich geen reden om aan de constructvaliditeit van een toets te twifelen. Als bijvoorbeeld een taaltoets Nederlands voor allochtone leerlingen moeilijker is dan voor autochtone, dan wordt meestal voldaan aan de functie van de toetsitems en de toets als geheel: discrimineren tussen taalvaardige en minder taalvaardige leerlingen. Scoreverschillen kunnen echter ook geheel of gedeeltelijk een artefact zijn van de gevolgde meetprocedure.

1.1 Itembias

Scoreverschillen tussen subgroepen kunnen ook veroorzaakt worden door vaardigheidsverschillen die de toetsitems niet beogen te meten, maar die onbedoeld toch gemeten worden. Wanneer voor het correct beantwoorden van de toetsitems nog een andere vaardigheid nodig is dan de vaardigheid die de toets beoogt te meten, kan afbreuk gedaan worden aan de constructvaliditeit van de toets. Dit laatste is aan de orde wanneer die benodigde additionele vaardigheid bij de onderscheiden subgroepen niet in vergelijkbare mate aanwezig is. Voor kinderen die afkomstig zijn uit verschillende subgroepen en die even vaardig zijn in hetgeen het item beoogt te meten is de kans op een goed antwoord dan ongelijk. Wanneer dit zich voordoet is er sprake van *itembias* of *itempartijdigheid*: de toetsitems doen bij een of meer subgroepen onbedoeld een beroep op een multidimensionale vaardigheid (Mellenbergh, 1989; Hambleton & Rogers, 1989; Glas & Ouborg, 1993; Camilli & Shepard, 1994; Uiterwijk & Vallen, 1994). Itembias is bijvoorbeeld aan de orde wanneer het taalgebruik in rekenitems voor allochtone leerlingen dermate ingewikkeld is, dat deze daardoor in onvoldoende mate toekomen aan het uitvoeren van de beoogde rekenoperaties zelf. Bij dit voorbeeld gaat het om itempartijdigheid in het nadeel van allochtone leerlingen, maar het is ook mogelijk dat items partijdig zijn in hun voordeel.

Voor het correct beantwoorden van toets-items moeten leerlingen vaak een beroep doen op additionele vaardigheden. Zo moeten leerlingen voor het beantwoorden van items inzake begrijpend lezen vaak eerst een tekst lezen. Voor het kunnen begrijpen van die tekst wordt kennis verondersteld over het onderwerp dat in die tekst aan de orde wordt gesteld. Ook bij rekenitems en items inzake wereldoriëntatie wordt bij het meten van de beoogde vaardigheid veelal gebruik gemaakt van contexten. Met het hanteren van contexten wordt gepoogd vast te stellen of de verworven kennis en vaardigheid en het verkregen inzicht bij een leerling voldoende groot en flexibel zijn om in een zo reëel mogelijke context te kunnen toepassen. De toetsconstructeur moet er bij de keuze van het contextmateriaal rekening mee houden dat de vereiste voorkennis, bijvoorbeeld inzake de beschreven situaties, het gehanteerde taalgebruik en het gebruikte beeldmateriaal (foto's, grafieken e.d.) bij alle leerlingen in vergelijkbare mate aanwezig is. Uit onderzoek naar itembias moet blijken in welke mate de toetsconstructeur daarin ten aanzien van bepaalde subgroepen is geslaagd. Door de inhoud te analyseren van partijdige items, kan gepoogd worden vast te stellen wat de oorzaak van de bias is. Deze inhoudsanalyse kan de toetsconstructeur aanwijzingen opleveren ter vermindering van itembias in nog te ontwikkelen toetsen.

1.2 Toetsbias

Toetsen worden ook gebruikt om voorspellingen te doen over buiten de toetssituatie liggend gedrag. Op grond van een behaalde toetsscore kan dan een verwachting worden uitgesproken over iemands niveau op een bepaald criterium op grond van eerder verworven kennis over de relatie tussen toetsscore en criteriumgedrag. Zo worden in Nederland aan het einde van de basisschool de taal- en reken- en vooral de totaalscores van leerlingen gebruikt om een indicatie te geven van het naar verwachting te behalen niveau in het voortgezet onderwijs, omdat uit eerder onderzoek de relatie tussen toetsscores en behaald niveau in het voortgezet onderwijs empirisch is vastgesteld. De predictieve validiteit van de toetsscore kan ook voor onderscheiden subgroepen bepaald worden. Er is sprake van *toetsbias* of *toetspartijdigheid* als systematische schattingsfouten gemaakt wor-

den bij het voorspellen van de positie op het extern criterium als een functie van groepslidmaatschap. In feite is een toets onpartijdig wanneer de regressielijnen van toets op extern criterium van twee subgroepen (bijna) samenvallen (Jensen, 1980; Reynolds, 1982; Uiterwijk, 1994; Camilli & Shepard, 1994).

Om te kunnen beoordelen of een toets partijdig is in het voor- of nadeel van een bepaalde subgroep is het noodzakelijk om over een extern criterium te beschikken waarvan is aangetoond dat het zelf niet partijdig is voor de onderscheiden subgroepen (Jensen, 1980; Van de Vijver, Willemse & Van de Rijt, 1993; Uiterwijk, 1994). Bij de overgang van basisonderwijs naar voortgezet onderwijs kan het extern criterium 'succes in het voortgezet onderwijs' minder bruikbaar zijn in onderzoek naar toetsbias, wanneer bijvoorbeeld leden van de ene subgroep bij de toelating tot het voortgezet onderwijs eerder het voordeel van de twijfel krijgen dan gelijkpresterende leden van de andere subgroep. Ook bij de beslissingen die over de doorstroming van leerlingen in het voortgezet onderwijs worden genomen, is het niet zeker dat alleen hetgeen de toets meet (bijvoorbeeld: het algemene prestatieniveau van de leerling) als criterium wordt gehanteerd. Hoe verder een gehanteerd extern criterium in de tijd verwijderd ligt van het moment waarop de toets wordt afgenomen die object van toetsbiasonderzoek is, des te groter is de kans dat er variabelen zijn die een differentieel effect hebben op de schoolloopbanen van twee subgroepen. Hierdoor kan afbreuk gedaan worden aan de predictieve validiteit van een toets voor onderscheiden subgroepen. Deze predictiever verschillen zijn echter niet te interpreteren als eigenschappen van de toets, maar als kenmerken van het extern criterium.

Hoewel er meestal geen volledig adequaat extern criterium beschikbaar is om de partijdigheid van een toets te beoordelen, is het niettemin zinvol om het verband tussen een toets en een maat voor succes in het voortgezet onderwijs voor bijvoorbeeld allochtone en autochtone leerlingen te onderzoeken. De gevonden verbanden geven immers aan of de mate van trefzekerheid waarmee het succes in het voortgezet onderwijs van allochtone en autochtone leerlingen voorspeld wordt vergelijkbaar is. Om een zo adequaat mogelijke verklaring te

kunnen geven voor eventuele verschillen, zijn echter de gegevens nodig van een groot aantal relevante onafhankelijke en afhankelijke variabelen en de effecten van die variabelen op elkaar moeten in een longitudinaal model geschat worden.

1.3 Literatuur over onderzoek naar item- en toetsbias in de Cito-Eindtoets Basisonderwijs

De Jong en Vallen hebben in 1989 in dit tijdschrift op grond van een door hen verrichte literatuurstudie bericht over mogelijke bronnen van itembias voor allochtone leerlingen. Zij hebben door een overzicht te geven van linguïstische en culturele bronnen van itembias in feite materiaal geleverd voor het formuleren van (werk)hypothesen over oorzaken van itembias. Coenen en Vallen hebben in 1991 in aansluiting op het artikel van De Jong en Vallen (1989) in dit tijdschrift verslag gedaan van een experiment waarin is nagegaan hoe vaak allochtone leerlingen bij een aantal partijdige items uit de Eindtoets Basisonderwijs 1987 van het Instituut voor toetsontwikkeling (Cito) een fout antwoord geven tengevolge van een element in het item dat waarschijnlijk de itembias veroorzaakt. Verder hebben Coenen en Vallen onderzocht hoe vaak vervolgens na manipulatie van dezelfde items tengevolge van die itemmanipulatie een goed antwoord wordt gegeven. Gemanipuleerde items zijn in dit verband items waarbij het itemelement dat vermoedelijk de biasbron vormt, is vervangen door een equivalent itemelement waarvan verwacht wordt dat het geen bias veroorzaakt. Het onderzoek van Coenen en Vallen heeft laten zien dat talige biasbronnen voor allochtone leerlingen voor een groot deel op het gebied van woordgebruik, impliciete zins- en tekstverbanden en verwijswaarden gezocht moeten worden. Verder lijken ongebruikelijke uitdrukkingen en woordvormgelijkenissen tot problemen te leiden. De geringere taalvaardigheid Nederlands van allochtone leerlingen kan er toe leiden dat ze meer moeite met complexe items hebben. Complexe items bevatten meer context en voor het oplossen van dergelijke items moet de leerling vaak een aantal tussenstappen maken. Welke dat zijn moet meestal uit de talige context afgeleid worden. Verder blijft het uiteraard mogelijk dat de itemcontext voor allochtone leerlingen minder herkenbaar is dan voor autochtone.

In Uiterwijk (1994) is gerapporteerd over het onderzoek naar zowel toets- als itembias dat in het kader van de Cito-Eindtoets Basisonderwijs 1987 en 1989 is verricht; in Uiterwijk en Vallen (1994) is het accent gelegd op het onderzoek naar itembias in beide Cito-eindtoetsen. In het onderhavige artikel, dat gezien kan worden als een vervolg op De Jong en Vallen (1989) en op Coenen en Vallen (1991) worden in de paragrafen 3 en 4 de belangrijkste resultaten gegeven uit Uiterwijk (1994). Bovendien zijn eerste resultaten opgenomen van het onderzoek naar toets- en itembias dat in het kader van de Eindtoets Basisonderwijs 1993 is gehouden.

2 Opzet van het onderzoek

Om na te gaan of de Eindtoets Basisonderwijs van het Cito voor allochtone leerlingen even bruikbaar is als voor autochtone leerlingen, voeren medewerkers van het Werkverband Taal en Minderheden van de Letterenfaculteit van de KUB en medewerkers van het project Eindtoets Basisonderwijs van het Cito een taalkundig-onderwijskundig onderzoek uit. In de eerste plaats hebben KUB en Cito onderzocht welk verband er bestaat tussen de scores op Eindtoets Basisonderwijs en een maat voor schoolsucces in het voortgezet onderwijs. Ten tweede is onderzoek gedaan naar itembias. In het onderzoek naar itembias zijn twee complementaire fasen onderscheiden: de opsporings- en de verklaringsfase. In de opsporingsfase worden met statistische technieken partijdige items gedetecteerd. In de tweede fase wordt gepoogd vast te stellen wat de oorzaak kan zijn van de bias in een item of in een reeks items. Bij het opsporen van mogelijke oorzaken van itembias zijn naast de projectmedewerkers van KUB en Cito ook een aantal experts en leerlingen uit groep acht van het basisonderwijs betrokken. Uit het onderzoek naar itembias moet blijken met welke aanpassingen de bruikbaarheid van de Eindtoets Basisonderwijs en andere toetsen voor allochtone leerlingen eventueel vergroot kan worden.

Voor de dataverzameling zijn door Cito en KUB achtergrondgegevens verzameld van de leerlingen die deelnamen aan de Eindtoets Basisonderwijs 1987 en 1989. Hierbij is aan de

leerkrachten van de leerlingen onder meer gevraagd aan te geven wat het land van herkomst van de ouders van de leerling is. Een leerling wordt tot een bepaald herkomstland gerekend wanneer beide ouders/verzorgers uit een hetzelfde land afkomstig zijn. Bij één-ouder-gezinnen geldt het herkomstland van de ouder/verzorger bij wie het kind woont. Tevens is nagegaan tot welk type voortgezet onderwijs de leerling werd toegelaten en naar welk type tweede leerjaar de leerling doorstroomde of dat er sprake was van doublure. Verder is het bestand met de resultaten van de deelnemers aan de Eindtoets Basisonderwijs 1993 gekoppeld aan het bestand dat de projectgroep Landelijke Evaluatie Onderwijsvoorrangsbeleid (LEO) van groep acht schooljaar 1992/1993 heeft opgebouwd. Door deze koppeling zijn er van 4218 deelnemers aan de Eindtoets Basisonderwijs 1993 achtergrondgegevens, zoals het herkomstland van de ouders, beschikbaar. Ook van deze leerlingen is nagegaan tot welke typen voortgezet onderwijs zij werden toegelaten en welke onderwijsposities de leerlingen na een jaar voortgezet onderwijs innamen.

Er zijn van de leerlingen uit 1987, 1989 en 1993 telkens twee databestanden opgebouwd, waarbij het tweede bestand een deelverzameling is uit het eerste. Het eerste bestand bestaat onder meer uit Eindtoets- en herkomstlandgegevens en het tweede bestand, dat door non-respons kleiner is, bevat naast Eindtoets- en herkomstlandgegevens ook gegevens over toelating tot en doorstroming in het voortgezet onderwijs. Het eerste bestand is in verband met het grotere aantal waarnemingen gebruikt voor onderzoek naar itembias, het tweede voor onderzoek naar toetsbias. Interessant is de vraag of het verwijderen van partijdige items leidt tot verhoging of verlaging van toetsbias. Door de opzet van dit onderzoek is het echter niet mogelijk om deze relatie te onderzoeken. In vervolgonderzoek is het evenwel gewenst het effect van itembias op toetsbias na te gaan.

24 3 Toetsbias

In het onderhavige onderzoek is nagegaan of de predictieve validiteit van de Eindtoets Basisonderwijs voor allochtone en autochtone leerlingen vergelijkbaar is. Hiermee wordt een

indicatie verkregen van de mate van trefzekerheid waarmee de Eindtoets Basisonderwijs in de onderwijspraktijk het schoolsucces in het voortgezet onderwijs van beide groepen voorspelt. Een beoordeling van de predictieve validiteit als zodanig is, zoals eerder is betoogd, niet mogelijk omdat er op dit moment geen adequate maat voor schoolsucces in het voortgezet onderwijs bestaat waarvan aanneemelijk gemaakt is, dat die onpartijdig is voor allochtone en autochtone leerlingen. In schoolloopbaanonderzoek bestaat een traditie ten aanzien van het construeren van een variabele die kan dienen als maat voor schoolsucces. Uitgangspunt hierbij is dat de verschillende schoolloopbanen van leerlingen in het voortgezet onderwijs hiërarchisch te ordenen zijn. Bosker (1990) en Van der Velden (1991) kwantificeren de posities die leerlingen in het voortgezet onderwijs innemen door de onderwijsniveaus (van IVBO tot VWO/Beroepsopleiding) en leerjaren (van een tot zes leerjaren) tot een zogenaamde leerjarenladder te ordenen.

In het onderhavige onderzoek is de schoolloopbaan relatief kort: de positie die na één leerjaar (de brugklas) wordt ingenomen. Voor de kwantificering van deze posities zijn in verband met dit onderzoek de verschillen tussen de scholen met betrekking tot het algemene prestatieniveau van de leerlingen van belang. Toetsscores zijn immers gerelateerd aan het prestatieniveau van de leerlingen en vanuit dit perspectief wordt gekozen uit de schooltypen van het voortgezet onderwijs die onderling qua moeilijkheidsgraad verschillen. In dit onderzoek zijn de posities die de leerlingen in het voortgezet onderwijs innemen, geordend door aan elke onderwijspositie de waarde toe te kennen die overeenkomt met het gemiddelde prestatieniveau van alle leerlingen in dat type, zoals gemeten door de Eindtoets Basisonderwijs 1987, respectievelijk 1989 en 1993 (zie Uiterwijk, 1994). Om na te gaan hoe groot de trefzekerheid is waarmee het schoolsucces van allochtone en autochtone leerlingen in het voortgezet onderwijs door de Eindtoets Basisonderwijs wordt voorspeld is de correlatie berekend tussen toetsscore en schoolsucces. In Tabel 1 staan de correlaties vermeld van de leerlingen die in 1987, respectievelijk 1989 aan de Eindtoets Basisonderwijs deelnamen en in datzelfde jaar toegelaten werden tot het voort-

Tabel 1

Product-moment correlaties tussen Cito-score 1987 en 1989 en schoolsucces in het voortgezet onderwijs

Groep	n	Cito-score 1987	n	Cito-score 1989
Autochtonen	3274	.76	3405	.78
Alle allochtonen	2661	.76	3092	.74
Noordwest-Europa	84	.77	93	.74
China	95	.81	127	.63
Oost-Europa	23	.75	28	.69
Zuid-Europa	124	.70	147	.69
Molukken	139	.73	157	.68
Antillen	58	.71	74	.81
Suriname: Hindoestanen	184	.76	157	.68
Suriname: Creolen	202	.68	196	.71
Turkije	431	.69	534	.70
Marokko	375	.69	540	.68

gezet onderwijs. De autochtone leerlingen vormen een a-selecte steekproef uit de autochtone toetsdeelnemers. Alle in Tabel 1 vermelde correlaties zijn significant.

Uit Tabel 1 blijkt dat de voorspellende waarde van de Cito-Eindtoets 1987 bij allochtone en autochtone leerlingen even hoog is. In 1989 is de voorspellende waarde van de Eindtoets bij allochtone leerlingen lager dan bij autochtone. Verder blijkt dat de coëfficiënten per etnische groep in 1987 en 1989 soms aanzienlijk verschillen. Hierbij kan het geringe aantal waarnemingen een rol gespeeld hebben.

Van de leerlingen die zowel aan de Eindtoets Basisonderwijs 1993 als aan het LEO-onderzoek deelnamen, is nagegaan hoe de hoog de correlatie is tussen de Cito-score, respectievelijk de score op de voor het LEO-onderzoek ontwikkelde intelligentietest en schoolsucces in het voortgezet onderwijs. De score op de LEO-intelligentietest (zie Doddema-Winse-mius & Van der Werf, 1989) is de som van de behaalde scores op de verbale en non-verbale subtests. In Tabel 2 geldt een leerling als allochtone leerling wanneer een van de ouders/verzorgers niet uit Nederland afkomstig is, in Tabel 1 is het herkomstland van beide ouders gehanteerd. Het effect van deze twee verschillende operationalisaties van herkomst op de uitkomsten van de analyses is niet bekend.

Tabel 2

Product-moment correlaties tussen Cito-score 1993, LEO-intelligentie-score en schoolsucces in het voortgezet onderwijs

Groep	n	Cito-score 1993	LEO-IQ
Autochtonen	1847	.80	.61
Allochtonen	518	.75	.56

Uit Tabel 2 blijkt dat het schoolsucces van allochtone leerlingen iets minder trefzeker wordt voorspeld dan dat van autochtone leerlingen. Dit geldt zowel voor de Cito-Eindtoets als voor de LEO-intelligentietest. Het verschil tussen allochtone en autochtone leerlingen is zowel op de Cito-Eindtoets als op de intelligentietest significant ($p < .001$, resp. $p < .05$). De voorspellende waarde van de LEO-intelligentietest ligt over de hele linie lager dan die van de Cito-Eindtoets 1993. In paragraaf 1.2 is gesteld dat niet zeker is of hoge of lage predictie toegeschreven moet worden aan de eigenschappen van het instrument of aan de partijdigheid van het extern criterium voor allochtone en autochtone leerlingen. In dit verband moet bij de verschillen in predictie tussen Cito-Eindtoets en LEO-test opgemerkt worden dat de Cito-Eindtoets afgenomen is om in schoolkeuzeprocessen een rol te spelen en dat de LEO-test louter ten behoeve van onderzoeksdoeleinden is gehanteerd.

4 Itembias

In het onderzoek naar itembias zijn, zoals eerder is opgemerkt, een opsporings- en een verklaringsfase onderscheiden. In de eerste fase

zijn met statistische technieken partijdige items opgespoord, in de tweede fase wordt gepoogd te achterhalen wat de oorzaak van de partijdigheid is.

4.1 Opsporen van partijdige items

Een item is partijdig wanneer leerlingen uit verschillende subgroepen maar met hetzelfde vaardigheidsniveau een ongelijke kans hebben om het betreffende item goed te beantwoorden. De statistische technieken voor het opsporen van een partijdig item evalueren dan ook de verschillen tussen subgroepen in de kansen op een goed antwoord gegeven een maat voor de vaardigheid die de toets als geheel beoogt te meten. Er moet dus een gevalideerde maat voor de vaardigheid beschikbaar zijn om de leerlingen van de onderscheiden subgroepen in niveaugroepen in te delen alvorens nagegaan kan worden of de kans op een goed antwoord voor de subgroepen vergelijkbaar is. Voor het opsporen van partijdige items zijn de meest algemeen toegepaste technieken gebaseerd op de Mantel-Haenszel-toets of op een model uit de Itemresponsstheorie. Omdat gebleken is dat de resultaten van de verschillende detectieprocedures niet gelijk zijn, wordt algemeen geadviseerd om meer dan één techniek te hantieren. Voor uiteenzettingen over de kenmerken en de voor- en nadelen van de verschillende technieken kan verwezen worden naar Hambleton, Clauser, Mazor & Jones, 1993; Dorans & Holland, 1993; Thissen, Steinberg & Wainer, 1993; Glas & Ouborg, 1993; Camilli & Shepard, 1994.

Voor het opsporen van partijdige items in de Eindtoets Basisonderwijs 1987 en 1989 voor allochtone leerlingen is in de eerste plaats het computerprogramma One Parameter Logistic Model (OPLM) (Verhelst, 1992) gebruikt als procedure die gebaseerd is op de itemresponsstheorie (IRT), vervolgens is het Mantel-Haenszel-programma (Verhelst, 1988) gehanteerd. Om de stabiliteit van de itemdetectieprocedures te bepalen, zijn de analyses telkens uitgevoerd op twee a-selecte steekproeven uit de onderscheiden subgroepen. In verband met het aantal waarnemingen per subgroep zijn alleen de Turkse, de Marokkaanse en de autochtone leerlingen in de analyses betrokken.

Als eerste stap in de procedure voor het opsporen van partijdige items is bepaald welke

eendimensionale schalen elk toetsonderdeel (Taal, Rekenen en Informatieverwerking) van Eindtoets Basisonderwijs 1987 en 1989 bevat. Hierbij is eerst nagegaan welke items voor autochtone leerlingen op een schaal passen. Een schaal is als basis voor het onderzoek naar itembias geaccepteerd wanneer geen enkel item partijdig is ($p < .01$) bij de vergelijking a-selecte steekproef Autochtonen 1 versus a-selecte steekproef Autochtonen 2. Uit de uitgevoerde analyses blijkt dat de Eindtoets Basisonderwijs 1987 en 1989 voor autochtone leerlingen uit twee taal- en twee informatieverwerkingschalen bestaat. De ene taalschaal bestaat telkens voornamelijk uit spellingitems, terwijl de andere voornamelijk uit taalgebruikitems bestaat. De ene informatieverwerkingschaal wordt telkens voornamelijk gevormd door items afkomstig uit de rubriek Lezen van teksten, terwijl de andere vooral items bevat uit de rubrieken Hanteren van Informatiebronnen, Kaartlezen en Lezen van Tabellen en Grafieken. De Eindtoets Basisonderwijs 1987 bevat twee en die uit 1989 bevat drie rekenschalen. De rekenschalen zijn minder eenvoudig te karakteriseren.

Als tweede stap in de procedure is zowel met het computerprogramma OPLM als met het Mantel-Haenszel-programma voor elke steekproef autochtone versus Turkse, respectievelijk Marokkaanse leerlingen vastgesteld welke items partijdig zijn ($p < .01$). Een item is partijdig voor Turkse leerlingen wanneer het item zowel bij de vergelijking steekproef Autochtonen 1 versus steekproef Turken 1 als bij de vergelijking Autochtonen 2 versus Turken 2 partijdig is. Dezelfde procedure is ook bij de vergelijking autochtone versus Marokkaanse leerlingen gebruikt.

Uitgangspunt bij de gevolgde procedure is dat als een reeks items voor autochtone leerlingen een eendimensionale schaal blijkt te vormen, die items ook een eendimensionale schaal dienen te vormen voor Turkse en Marokkaanse leerlingen. Wanneer dezelfde items voor Turkse en Marokkaanse leerlingen niet passen dan is deze schaal voor de allochtone leerlingen multidimensionaal en de niet-passende items kunnen dan als partijdig beschouwd worden. Voor zowel de Mantel-Haenszel- als voor de OPLM-procedure moeten de twee subgroepen in niveaugroepen ingedeeld worden aan de

Tabel 3

Aantal items uit de Eindtoets Basisonderwijs 1987 en 1989 dat partijdig is voor Turkse en/of Marokkaanse leerlingen

Onderdeel	Mantel-Haenszel		IRT		IRT & MH	
	voor-deel	na-deel	voor-deel	na-deel	voor-deel	na-deel
Taal (k=120)	8	10	3	1	2	7
Rekenen (k=120)		4		1		
Informatie- verwerking (k=120)	4	6	1	1	1	3
Totaal (k=360)	12	20	4	3	3	10

hand van een meting die kan gelden als maat voor de vaardigheid die de toets als geheel beoogt te meten. Uitgangspunt bij de keuze van het criterium voor het matchen van subgroepen is geweest dat de somscore van de items, die voor autochtone leerlingen op eendimensionale schaal passen, ook geldig dient te zijn als somscore voor de Turkse en Marokkaanse leerlingen.

De resultaten van de analyses naar itembias maken duidelijk dat het moeilijk is om aan te geven hoeveel items van Eindtoets Basisonderwijs 1987 en 1989 partijdig zijn. De verschillende analyses laten een wisselend beeld zien. Uit de uitgevoerde analyses blijkt dat 13 van de 360 geanalyseerde items (= 4%) zowel met de IRT- als met de Mantel-Haenszelprocedure partijdig zijn voor Turkse en/of Marokkaanse leerlingen. Hiervan zijn 3 items partijdig in het voordeel van deze leerlingen, 10 items zijn partijdig in hun nadeel. Verder blijkt dat 7 items (= 2%) alleen partijdig zijn met de techniek volgens het IRT-model, hiervan zijn 4 items partijdig in het voordeel, 3 items zijn in het nadeel. Er zijn 32 items (= 9%) alleen partijdig met de Mantel-Haenszelprocedure, hiervan zijn er 12 in het voor- en 20 in het nadeel van Turkse en/of Marokkaanse leerlingen. Verder volgt uit Tabel 3 dat itembias het meest voorkomt bij Taal, gevolgd door respectievelijk Informatieverwerking en Rekenen.

Uit nadere analyse blijkt dat de IRT- en de Mantel-Haenszel-procedures in 87% van de gevallen overeenstemmen in het detecteren van (on)partijdige items. Dit beeld komt overeen met de resultaten van andere onderzoekers (Bügel & Glas, 1991; Hambleton & Jones, 1992; Camilli & Shepard, 1994). De stabiliteit van de IRT- en de Mantel-Haenszel-procedure

is vrijwel gelijk: in twee steekproeven wijst de Mantel-Haenszel-procedure bij 86% van de items en de IRT-procedure bij 89% van de items in beide steekproeven een item als partijdig aan. De IRT-procedure spoort minder partijdige items op dan de Mantel-Haenszel-techniek. Items blijken nooit in het voordeel te zijn voor Turkse leerlingen en tegelijkertijd in het nadeel van Marokkaanse leerlingen of omgekeerd.

Van de leerlingen die zowel aan de Eindtoets Basisonderwijs 1993 als aan het LEO-onderzoek deelnamen, is nagegaan in welke mate er in de Eindtoets Basisonderwijs 1993 sprake is van itembias in het voor- of nadeel van allochtone leerlingen. Een schaal is, evenals dat gebeurde met de data uit 1987 en 1989, in 1993 als basis voor onderzoek naar itembias geaccepteerd wanneer geen enkel item partijdig is in de vergelijking steekproef Autochtonen 1 versus steekproef Autochtonen 2 ($p < .01$). Een item is partijdig ($p < .01$) voor allochtone leerlingen wanneer het item bij de vergelijking steekproef Autochtonen 1 versus alle allochtone leerlingen partijdig is. Er is niet voor gekozen om beide steekproeven autochtone leerlingen ($n = 2 \times 772$) samen te voegen, omdat de te vergelijken groepen allochtone ($n = 423$) en autochtone leerlingen bij voorkeur ongeveer even groot moeten zijn (Hambleton et al., 1993; Zieky, 1993). De steekproef Autochtonen 1 onderscheidt zich van de steekproef Autochtonen 2, omdat van de eerste steekproef de passing op de eendimensionale schalen telkens het grootst is.

Uit de resultaten van de analyses naar itembias in de Eindtoets Basisonderwijs 1993 blijkt dat 16 van de in totaal 180 geanalyseerde items (= 9%) zowel met de IRT- als met de Mantel-

Tabel 4

Aantal items uit de Eindtoets Basisonderwijs 1993 dat partijdig is voor allochtone leerlingen

Onderdeel	Mantel-Haenszel		IRT		IRT & MH	
	voor-deel	na-deel	voor-deel	na-deel	voor-deel	na-deel
Taal (k=60)	4	2	3		3	7
Rekenen (k=60)	2		1	2	1	1
Informatieverwerking (k=60)	4	5	2	1	1	3
Totaal (k=180)	10	7	6	3	5	11

Haenszelprocedure partijdig zijn voor allochtone leerlingen. Van deze 16 items zijn er 5 partijdig in het voordeel en 11 in het nadeel van allochtone leerlingen. Verder blijkt dat 9 items (= 5%) alleen met de IRT-procedure partijdig zijn voor allochtone leerlingen. Hiervan zijn 6 items partijdig in het voordeel, 3 items zijn in het nadeel. Van de 180 geanalyseerde items zijn 17 items (= 9%) alleen partijdig met de Mantel-Haenszel-procedure. Van deze 17 items zijn er 10 in het voordeel van allochtone leerlingen, 7 items zijn in het nadeel. Ook hier spoort de IRT-procedure minder partijdige items op dan de Mantel-Haenszel-techniek. Verder volgt uit Tabel 4 dat bij Taal het aantal partijdige items het grootst is, gevolgd door respectievelijk Informatieverwerking en Rekenen.

Er is een aanzienlijke overlap tussen de Mantel-Haenszel- en de IRT-procedure, maar er zijn ook verschillen waardoor het niet altijd duidelijk is of een item partijdig is of niet. Over het algemeen wordt dan ook geadviseerd in de vorm van kruisvalidatie beide technieken te hanteren. Onderzoek naar itembias heeft als resultaat dat uitspraken over de partijdigheid van items gradueel zijn: partijdig bij alle analyses, bij een deel van de analyses, bij geen enkele analyse.

Bij alle partijdige items is in de volgende fase van het onderzoek nagegaan wat de oorzaak van de partijdigheid zou kunnen zijn. De stelligheid waarmee op dit punt conclusies getrokken worden is mede afhankelijk van het aantal keren dat er bij een item sprake is van bias.

4.2 Bronnen van itembias

Er is niet alleen in Nederland maar ook in andere landen weinig onderzoek gedaan naar inhoudelijke oorzaken van itembias voor allochtone leerlingen. Volgens Schmitt, Holland &

Dorans (1993) zijn hiervoor drie redenen aan te wijzen. In de eerste plaats is onderzoek naar itembias relatief nieuw. Tot nu toe is de meeste aandacht uitgegaan naar statistische procedures voor het detecteren van partijdige items. In de tweede plaats veronderstelt het achterhalen van oorzaken van itembias voor allochtone leerlingen een theorie over de vraag waarom items voor de onderscheiden etnische groepen partijdig zijn. Omdat etnische groepen in velerlei opzichten heterogeen zijn, kunnen de verschillen tussen etnische groepen moeilijk beschreven worden. In de derde plaats is het opsporen van oorzaken van itembias uitermate complex, omdat bij een bepaald item verschillende oorzaken een rol kunnen spelen.

Omdat een theoretisch kader betreffende inhoudelijke bronnen voor itembias voor allochtone leerlingen vooralsnog niet beschikbaar is, moeten we de conclusies die op basis van het onderhavige onderzoek in dit verband getrokken worden, als voorlopig beschouwen. Om bij de inhoudsanalyse van in statistisch opzicht partijdige items toch enigszins gericht te kunnen kijken, is eerst een literatuuroverzicht gemaakt van potentiële bronnen van itembias. Dit overzicht bestaat uit een reeks van (deel)vaardigheden waarvan uit onderzoek is gebleken of waarvan te verwachten is dat allochtone en autochtone leerlingen deze aan het einde van het basisonderwijs in verschillende mate beheersen. Het gaat hier dus om verschil in moeilijkheid en niet om itembias. Verschillen in moeilijkheidsgraad zijn in het kader van dit onderzoek van belang omdat ze de basis kunnen vormen voor hypothesen omtrent oorzaken van itembias (De Jong & Vallen, 1989; Coenen & Vallen, 1991; Uiterwijk, 1994).

Bij het zoeken naar mogelijke oorzaken van itembias in de Eindtoets Basisonderwijs 1987 en 1989 zijn niet alleen de medewerkers van

het onderzoeksproject (van KUB en Cito) betrokken geweest maar ook een aantal niet aan het project verbonden experts en leerlingen uit groep acht van het basisonderwijs.

De partijdige items in het voor- of nadeel van Turkse en/of Marokkaanse leerlingen zijn op grond van de in het geding zijnde vaardigheden eerst door drie projectmedewerkers afzonderlijk inhoudelijk geanalyseerd tegen de achtergrond van de vraag welke itemelementen mogelijk een bron van itembias vormen. Hierbij ging speciale aandacht uit naar elementen die voorkomen in items die partijdig zijn in het nadeel van beide groepen allochtone leerlingen en ontbreken in items die in het voordeel zijn van deze leerlingen en omgekeerd. Bij deze inhoudsanalyse hebben de vooraf geformuleerde potentiële bronnen van itembias voor allochtone leerlingen als hypothesen gefungeerd. Vervolgens zijn de gemeenschappelijkheden in de analyses van de afzonderlijke projectmedewerkers geïnventariseerd.

De inhoudelijke analyse leverde twee fundamentele problemen op. Enerzijds blijkt het moeilijk te zijn om met zekerheid aan te geven welk element van een item de biasbron is en anderzijds blijkt dat bij sterk vergelijkbare items de ene keer wel sprake is van itembias en de andere keer niet. Tevens werd vastgesteld dat er een aanzienlijk aantal partijdige items is met onderlinge overeenkomsten. Items die samen een cluster vormen, vertonen overeenkomsten ten aanzien van hetgeen de items beogen te meten of vertonen overeenkomsten met betrekking tot additionele vaardigheden die relevant zijn voor het juist beantwoorden van de items. De inhoudelijke analyses leveren voorlopige conclusies op (vgl. Uiterwijk & Vallen, 1994). Deze conclusies worden voorlopig genoemd, omdat de conclusies in een volgende fase van het onderzoek zijn vergeleken met de oordelen van 16 niet bij het onderzoeksproject betrokken experts, elf autochtone en vijf allochtone (Van de Waal, 1992; Uiterwijk, 1994). Verder is met een kleinschalig hardopdenken-experiment nagegaan hoe vaak allochtone en autochtone leerlingen bij partijdige items een fout antwoord geven door een itemelement dat voorlopig als bron voor itembias is aangewezen. Tevens is onderzocht hoe vaak bij gemanipuleerde items tengevolge van de itemmanipulatie een goed antwoord wordt gegeven

(Coenen & Vallen, 1991). Gemanipuleerde items zijn items waarbij het itemelement dat als potentiële biasbron is aangewezen (bijvoorbeeld: 'Hoeveel moet hij betalen inclusief B.T.W.'), is vervangen door een itemelement waarvan verwacht wordt dat het geen bias veroorzaakt (bijvoorbeeld: 'Hoeveel moet hij betalen met B.T.W.'). De allochtone en autochtone leerlingen moesten eerst de aan hen voorgelegde oorspronkelijke, respectievelijk gemanipuleerde items goed bestuderen, daarna konden ze het naar hun oordeel goede antwoord aankruisen. Tot slot moesten ze zo uitgebreid en nauwkeurig mogelijk mondeling toelichten hoe ze tot hun antwoord gekomen waren. In de redeneringen van de leerlingen is getracht aanwijzingen te vinden omtrent inhoudelijke oorzaken van itembias.

De partijdige items uit de Eindtoets Basisonderwijs 1993 zijn eveneens onderzocht op mogelijke oorzaken van itembias. De bevindingen uit het onderzoek met de Eindtoets Basisonderwijs 1987 en 1989 zijn in feite geverifieerd, waardoor stelligheid waarmee uitspraken gedaan worden kan toenemen. Uit de resultaten van 1993 blijkt dat het grootste deel van de resultaten uit 1987 en 1989 ondersteund wordt. In het volgende overzicht, wordt aangegeven welke items gezien het onderzoek met de Eindtoets Basisonderwijs uit 1987, 1989 en 1993 partijdig kunnen zijn in het nadeel van Turkse en Marokkaanse leerlingen.

- Items die vragen naar de betekenis van een woord, een woordcombinatie of een zin, waarbij om een min of meer woordelijke herhaling van expliciet in de tekst gegeven informatie wordt gevraagd.
- Items die vragen naar de betekenis van moeilijke woorden, woordcombinaties, functiewoorden en verwijswoorden, waarbij de betekenis niet of moeilijk uit de context kan worden afgeleid.
- Items die betrekking hebben op de kennis van de vorm van idiomatische, letterlijk of figuurlijk gebruikte woordcombinaties en/of conventies op het gebied van de zinsbouw (zinsvolgorde, inversie/vraagvormen).
- Items die veel contextmateriaal bevatten. Voor het oplossen van complexe items moet de leerling vaak een aantal tussenstappen maken. Uit de context moet dan afgeleid

worden welke tussenstappen gemaakt moeten worden en dit wordt voor allochtone leerlingen extra moeilijk wanneer er in dergelijke items (vooral op zinsniveau) veel verwijswaarden en impliciete zins- en tekstverbanden voorkomen.

- Items die betrekking hebben op contextmateriaal (bijvoorbeeld een tekst) dat specifieke Nederlandse cultuurkennis vereist. Zo kunnen allochtone leerlingen minder vertrouwd zijn met het onderwerp dat in een tekst aan de orde wordt gesteld. Dit speelt vooral wanneer dat contextmateriaal cruciaal is voor het oplossen van het item.

Verder zijn er aanwijzingen dat de volgende items partijdig kunnen zijn in het voordeel van allochtone leerlingen.

- Items die naar de hoofdgedachte van een tekst vragen.
- Items die vragen spellingfouten in werkwoorden en in woorden met een vast woordbeeld aan te geven.

5 Conclusies en discussie

Of een toets het schoolsucces in het voortgezet onderwijs van een bepaalde subgroep even goed voorspelt als van een andere subgroep, is in feite niet te beoordelen zolang er geen extern criterium is waarvan aangetoond is dat het zelf onpartijdig is. In paragraaf 1.2 is daarom gesteld dat het niet zeker is of hoge of lage predictie toegeschreven moet worden aan de eigenschappen van het instrument of aan de partijdigheid van het extern criterium voor allochtone en autochtone leerlingen. In het onderhavige onderzoek is wel gebleken dat de Cito-Eindtoets Basisonderwijs het schoolsucces in het voortgezet onderwijs van allochtone leerlingen minder trefzeker wordt voorspeld dan dat van autochtone leerlingen. Voor de verklaring van deze verschillen zou nauwkeurig vastgesteld moeten worden of bij allochtone en autochtone leerlingen de beschikbare gegevens, zoals advies basisschool en toetsscore, door de betrokkene(n) in het schoolkeuzeprocess op dezelfde wijze worden gewogen. Het lijkt ook zinvol om na te gaan welke aanvullende informatie naast het advies basisschool en de Cito-score de trefzekerheid van de schoolkeuze van allochtone leerlingen kan vergroten. Opge-

merkt moet worden dat allochtone leerlingen in het voortgezet onderwijs meer dan hun autochtone klasgenoten blijven zitten of doorstromen naar andere – meestal ‘lagere’ – schooltypen (Uiterwijk, 1994). De relatief lage trefzekerheid waarmee het schoolsucces van allochtone leerlingen wordt voorspeld, is voor een deel te verklaren door de grote uitval in het voortgezet onderwijs van allochtone leerlingen. Het is van groot belang om na te gaan hoe de beslissingen over toelating in het voortgezet onderwijs verbeterd kunnen worden en hoe het onderwijsaanbod in het voortgezet onderwijs beter op de allochtone leerlingen kan worden afgestemd.

Ten aanzien van het onderzoek naar itembias kan gezegd worden dat dit met veel onzekerheden is omgeven. In de eerste plaats kan niet duidelijk worden aangegeven of een item partijdig is of niet (4.1). Dit wordt veroorzaakt door het ontbreken van volledige overeenstemming tussen de resultaten van de verschillende itembiasdetectieprocedures, waardoor het in dit soort onderzoek gebruikelijk is om meer dan één analysetechniek te hanteren.

In de tweede plaats kan niet duidelijk aangegeven worden welk element van een item verantwoordelijk is voor de itembias (4.2). De bron van itembias kan bijvoorbeeld in het contextmateriaal, in de vraag, de antwoordmogelijkheden en in de te meten (deel)vaardigheid zitten. Naarmate het contextmateriaal omvangrijker wordt, is het moeilijker om met zekerheid de inhoudelijke biasbron aan te wijzen. Zo kan de voorkennis omtrent hetgeen in teksten aan de orde wordt gesteld van invloed zijn op de itemantwoorden.

Via verschillende wegen (inhoudsanalyse, expertbevraging, experimenten) zal informatie over bronnen van itembias verzameld moeten worden.

Als uiteindelijk – zoals in de onderhavige studie – met een zekere mate van waarschijnlijkheid een aantal biasbronnen opgespoord is, dan rest nog de vraag of de biasbronnen afbreuk doen aan de constructvaliditeit van de toets of niet. Wanneer een bron van itembias behoort tot de te meten vaardigheid, dan meet het item wat het beoogt te meten. Als voorbeeld geldt het volgende taalitem. Het toetsonderdeel Taal beoogt onder meer de hier aan de orde zijnde vaardigheid ‘het kennen van de betekenis van woorden’ te meten.

- 43 Ik hoop echt dat jullie komen. Jullie
 44 moeten wel uiterlijk half zeven bij ons
 45 thuis zijn. Jullie moeten echt zo vroeg
 46 mogelijk komen hoor. Anders zijn er
 47 vast geen kaartjes meer te krijgen.

Wat had de briefschrijver in plaats van *uiterlijk* (regel 44) ook kunnen schrijven?

- A op zijn best
 B op zijn laatst
 C op zijn mooiste
 D op zijn vroegst

Dit item dat naar de betekenis van het specifieke woord 'uiterlijk' vraagt, is multidimensionaal voor Marokkaanse leerlingen. Toch kan niet gezegd worden dat het item iets anders zou meten dan het moet meten. Als aangenomen wordt dat leerlingen het woord 'uiterlijk' in deze betekenis aan het einde van het basisonderwijs moeten kennen, dan doet in dit geval itembias geen afbreuk aan de constructvaliditeit. Wanneer een biasbron niet tot de te meten vaardigheid behoort dan doet het item wel afbreuk aan de constructvaliditeit van de toets. Zo moeten bronnen van itembias op het gebied van woordenschat geen rol spelen bij rekenitems. Itembias als zodanig beperkt de constructvaliditeit van een toets dus niet in alle gevallen.

Gebleken is dat het aantal bronnen van itembias dat met een grote mate van zekerheid bias veroorzaakt, in feite gering is. Het detecteren van partijdige items en het verklaren van de oorzaken ervan is een lange weg met een bescheiden opbrengst. Informatie over potentiële bronnen van itembias komt ook beschikbaar uit elk empirisch onderzoek dat duidelijk maakt bij welke kennis- en vaardigheidsaspecten allochtone leerlingen significant lager scoren dan autochtone leerlingen. Wanneer de toetsconstructeur over gedetailleerde informatie op dit gebied beschikt, kan hij het risico verkleinen dat hij inzake allochtone leerlingen items construeert die afbreuk doen aan de constructvaliditeit van een toets. Hiervoor is het van belang dat in empirisch onderzoek kennis- en vaardigheidsaspecten op een zo gedetailleerd mogelijk niveau gemeten worden.

- Bosker, R.J. (1990). *Extra kansen dankzij de school?* Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen.
- Bügel, K., & Glas, C. (1991). Item-specifieke verschillen in prestaties van jongens en meisjes bij tekstbegripexamens moderne vreemde talen. *Tijdschrift voor Onderwijsresearch*, 16, 337-351.
- Camilli, G., & Shepard, L.A. (1994). *Methods for identifying biased test items*. Thousand Oaks: Sage Publications.
- Coenen, M., & Vallen, T. (1991). Itembias in de eindtoets basisonderwijs. *Pedagogische Studiën*, 68, 15-26.
- Doddema-Winsemius, H., & Werf, G. van der (1989). *Selectie/constructie van toetsen voor sociale redzaamheid en intelligentie ten behoeve van de evaluatie van het OVB*. Groningen: Research Instituut voor onderzoek in het noorden.
- Dorans, N.J., & Holland, P.W. (1993). Dif detection and description: Mantel-Haenszel and standardization. In P.W. Holland & H.W. Wainer (Eds.), *Differential item functioning* (pp. 35-66). Hillsdale: Lawrence Erlbaum Associates.
- Glas, C.A.W., & Ouborg, M.J. (1993). Vraagonzuiverheid. In T.J.H.M. Eggen & P.F. Sanders (Eds.), *Psychometrie in de praktijk* (pp. 349-370). Arnhem: Cito, Instituut voor toetsontwikkeling.
- Hambleton, R.K., & Rogers, H.J. (1989). Detecting potentially biased test items: comparison of IRT area and Mantel-Haenszel methods. *Applied Measurement in Education*, 2, 313-334.
- Hambleton, R.K., & Jones, R.W. (1992). *Comparison of empirical and judgmental methods for detecting differential item functioning*. Princeton: Educational Testing Service.
- Hambleton, R.K., Clauser, B.E., Mazor, K.M., & Jones, R.W. (1993). Advances in the detection of differentially functioning test items. *European Journal of Psychological Assessment*, 9, 1-18.
- Jensen, A.R. (1980). *Bias in mental testing*. Londen: Methuen.
- Jong, M. de, & Vallen, T. (1989). Linguïstische en culturele bronnen van itembias in de Eindtoets Basisonderwijs voor leerlingen uit etnische minderheidsgroepen. *Pedagogische Studiën*, 66, 390-402.
- Mellenbergh, G.J. (1989). Itembias and itemresponse theory. In R.K. Hambleton (Ed.), *Applications of itemresponse theory (special issue)*. *International Journal of Educational Research*, 13, 127-143.

- Reynolds, C.R. (1982). Methods for detecting construct and predictive bias. In R.A. Berk (Ed.), *Handbook of methods for detecting test bias* (pp. 199-227). Baltimore: The Johns Hopkins University Press.
- Schmitt, A.P., Holland, P.W., & Dorans, N.J. (1993). Evaluating hypotheses about differential item functioning. In P.W. Holland & H.W. Wainer (Eds.), *Differential item functioning* (pp. 281-315). Hillsdale: Lawrence Erlbaum Associates.
- Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. In P.W. Holland & H.W. Wainer (Eds.), *Differential item functioning* (pp. 67-113). Hillsdale: Lawrence Erlbaum Associates.
- Uiterwijk, J.H. (1994). *De bruikbaarheid van de Eindtoets Basisonderwijs voor allochtone leerlingen*. Arnhem: Cito, Instituut voor toetsontwikkeling.
- Uiterwijk, H., & Vallen, T. (1994). Talige bronnen van itembias voor allochtone leerlingen in de Eindtoets Basisonderwijs. *Spiegel*, 12, 9-29.
- Velden, R.K.W. van der (1991). *Sociale herkomst en schoolsucces*. Monografieën onderwijsonderzoek nr. 10. Groningen: Instituut voor Onderwijsonderzoek.
- Verhelst, N.D. (1988). *De Mantel-Haenszel-toetsen*. Arnhem: Cito, Instituut voor Toetsontwikkeling.
- Verhelst, N.D. (1992). *Het eenparameter logistisch model (OPLM), een theoretische inleiding en een handleiding bij het computerprogramma*. Arnhem: Cito, Instituut voor toetsontwikkeling.
- Vijver, F. van de, Willemse, G., & Rijt, B. van de (1993). Het testen van cognitieve vaardigheden van allochtone leerlingen. *De Psycholoog*, 28, 152-159.
- Waal, M. van de (1992). *Expert-oordelen over potentiële bronnen van itembias in de Eindtoets Basisonderwijs*. Tilburg: Katholieke Universiteit Brabant.
- Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P.W. Holland & H.W. Wainer (Eds.), *Differential item functioning* (pp. 337-347). Hillsdale: Lawrence Erlbaum Associates.

Adres: Instituut voor toetsontwikkeling (Cito), Postbus 1034, 6801 MG Arnhem

T. Vallen is universitair hoofddocent bij het Werkverband Taal en Minderheden van de Letterenfaculteit van de Katholieke Universiteit Brabant.

Adres: Faculteit der Letteren KUB, Postbus 90153, 5000 LE Tilburg

Abstract

Testbias and itembias for ethnic minority pupils in the Final Primary School Tests of the Dutch National Institute for Educational Measurement

H. Uiterwijk & T. Vallen. *Pedagogische Studiën*, 1997, 74, 21-32.

So far, in the Netherlands as well as in other countries hardly any substantial research has been carried out into the suitability of tests for specific subgroups within a larger population. In an interdisciplinary (educational and sociolinguistic) research project the Dutch National Institute for Educational Measurement (Cito) and the Research Group on Language and Minorities of Tilburg University have cooperated in a study into the possible existence of testbias and itembias for ethnic minority pupils in the Cito Final Primary School Test in the Netherlands. With respect to testbias it emerges that the educational success in secondary education can be predicted less accurately for ethnic minority pupils than for pupils who belong to the majority group. This conclusion holds for both the Final Primary School Test and the intelligence test used.

The research carried out into itembias shows that the several statistical procedures used for detecting itembias for ethnic minority pupils in the Final Primary School Test produce different results. Most items never appear to be biased, others sometimes, and a limited number of items always show itembias. The procedures used for discovering the reasons for itembias in terms of content revealed especially within the field of linguistics a number of potential sources of itembias (e.g. vocabulary, referential problems, amount of context) for ethnic minority pupils. The knowledge of potential sources of itembias can be useful as a tool for diminishing the risk of itembias for these pupils in future tests.

Manuscript aanvaard 9-2-1996

Auteurs

H. Uiterwijk is projectleider Eindtoets Basisonderwijs bij het Instituut voor Toetsontwikkeling (Cito).

Steun van de ouders, sekse, en het verloop van de schoolloopbaan: een retrospectieve studie*

I.A. Bok en T.W. Taris

Samenvatting

In deze studie worden sekseverschillen in de effecten van ouderlijke steun op de ontwikkeling van de opleidingsloopbaan van jongeren onderzocht. In overeenstemming met eerdere bevindingen van Rollins en Thomas (1979) en Eccles (1987) wordt verondersteld dat de opleidingsloopbaan van jongens meer profiteert van ouderlijke steun dan die van meisjes. Gegevens van een longitudinale en retrospectieve steekproef van 986 jong-volwassenen worden gebruikt om onze hypothesen middels LISREL te toetsen. De resultaten ondersteunen de hypothesen gedeeltelijk: jongens die aangeven veel ouderlijke steun te hebben ondervonden ervoeren significant minder dropout, terwijl meisjes die veel steun ervoeren juist meer tijd nodig hadden om hun opleiding af te ronden. De implicaties en beperkingen van de studie worden bediscussieerd.

Inleiding

In deze studie worden de relaties tussen sekse, de manier van opvoeden door de ouders, en het verloop van de schoolloopbaan in het voortgezet onderwijs van 986 Nederlandse jongeren onder de loep genomen. Eerder onderzoek liet zien dat de opvoedingsstijl van de ouders een systematische invloed uitoefent op de schoolprestaties van kinderen, al dan niet via medierende variabelen als prestatie-motivatie, competentie, en beheersingsoriëntatie (bijvoorbeeld Dornbusch, Ritter, Leiderman, Roberts & Fraleigh, 1987; Gottfried, Fleming & Gottfried, 1994; Grolnick & Ryan, 1989; Grolnick, Ryan & Deci, 1991; Huston, 1983; Rollins &

Thomas, 1979). Uit het overzicht van Rollins en Thomas (1979) blijkt daarnaast dat de effecten van opvoedingsstijl op de schoolprestaties voor jongens duidelijk anders zijn dan voor meisjes. Voor jongens hangt bijvoorbeeld het ervaren van veel ouderlijke steun samen met hogere schoolprestaties, terwijl voor meisjes juist het omgekeerde het geval zou zijn.

Deze conclusie wordt echter weersproken in andere studies. Walters en Stinnett (1971) concludeerden in een eerder overzicht dat schoolprestaties voor zowel jongens als meisjes positief zouden samenhangen met de mate van ervaren steun door de ouders, terwijl Farmer (1985) alleen voor *meisjes* een positief effect van ouderlijke steun op prestatiemotivatie rapporteert; voor jongens blijkt dit effect niet significant. Het is dus nog onduidelijk of jongens en meisjes wel verschillen betreffende de effecten van ouderlijke steun op hun schoolprestaties, en – als zulke verschillen inderdaad bestaan – hoe deze zich dan manifesteren¹.

In de hier gerapporteerde studie wordt deze onderzoekslijn op drie manieren verder uitgebouwd. Ten eerste, het overgrote deel van het gepubliceerde onderzoek is afkomstig uit de Angelsaksische landen. Het is echter de vraag of en in hoeverre de aldaar gevonden verschillen tussen jongens en meisjes, hun schoolprestaties, en de ouderlijke stijl van opvoeden relevant zijn voor de Nederlandse situatie. Zoals Bügel in haar overzichtsstudie naar sekseverschillen in onderwijsprestaties opmerkt, zijn het schoolstelsel, de positie van vrouwen op de arbeidsmarkt en andere relevante aspecten van de Nederlandse samenleving zó afwijkend, dat transpositie van de resultaten van Amerikaans onderzoek naar de Nederlandse situatie niet zonder meer mogelijk is (Bügel, 1991, p. 350). In deze studie beogen we daarom na te gaan in hoeverre eerdere bevindingen geldig zijn voor de Nederlandse situatie: tevens vergroten we hiermee de empirische basis van waaruit conclusies getrokken kunnen worden

* Deze studie werd ondersteund door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek NWO, nummer PGS 56-381.

omtrent de aard en richting van het verband tussen opvoedingsstijlen en schoolprestaties.

Ten tweede, in deze studie wordt de onderwijsloopbaan in het *voortgezet* onderwijs bestudeerd. Eerder Nederlands onderzoek richtte zich vrijwel uitsluitend op de schoolloopbaan gedurende het basisonderwijs (zie het overzicht van Bügel, 1991).

Ten slotte, in deze studie richten we ons op variabelen die geacht worden het *verloop* van de schoolloopbaan weer te geven: bijvoorbeeld het aantal jaren dat men voltijds onderwijs volgde, het aantal malen dat men een opleiding die men volgde zonder diploma te behalen verliet ('drop-out'), en het uiteindelijk behaalde opleidingsniveau. Eerdere studies richtten zich vooral op variabelen als prestatie-motivatie, schoolcijfers, en schooltoetsen. Dit is te verklaren uit het feit dat veel van het onderzoek naar de relaties tussen opvoeding en schoolprestaties van cross-sectionele aard is; in ons onderzoek hanteerden we echter een combinatie van een retrospectieve en longitudinale dataverzameling, waardoor het voor de hand lag de nadruk te leggen op variabelen die het verloop van de schoolloopbaan *in de tijd* representeren, en niet zozeer op meer momentgebonden variabelen als schoolcijfers en motivaties. Uiteraard mag er een sterke relatie tussen beide typen variabelen verwacht worden; het is aannemelijk dat slechte cijfers en/of een slechte motivatie zich zullen vertalen in een lager opleidingsniveau en een langere opleidingsduur.

Hieronder wordt eerst dieper ingegaan op de relaties tussen stijl van opvoeden door de ouders, en het verloop van de schoolloopbaan. Vervolgens wordt een model gepresenteerd, dat getoetst wordt middels gegevens die afkomstig zijn uit een representatieve steekproef onder 986 jong-volwassenen.

Opvoedingsstijl van de ouders: Ouderlijke steun. Gewoonlijk worden de ouderlijke gedragingen ten opzichte van het kind ondergebracht in twee hoofddimensies (bijvoorbeeld Gerris & Janssens, 1987; Parker, Tupling & Brown, 1979; Rollins & Thomas, 1979; Ten Haaf, 1993; Taris & Semin, 1995). De eerste dimensie omvat gedragingen die tot doel hebben het gedrag van het kind te *controleren*, zoals het geven van strikte regels waaraan het kind zich dient te houden, het straffen van het kind bij ongewenst gedrag, et cetera. Deze dimensie

lijkt weer uit een aantal subdimensies te bestaan, namelijk 'controle' en 'dwang' (Rollins & Thomas, 1979; vergelijk Ten Haaf, 1993).

De tweede dimensie heeft betrekking op ouderlijke gedragingen die als 'warm' en 'liefdevol' beschouwd kunnen worden; acceptatie van het kind, affectie, tijd hebben voor en luisteren naar het kind, proberen het kind te begrijpen, enzovoort. Deze gedragingen worden wel samengevat onder de noemer van *ouderlijke steun* (Rollins & Thomas, 1979). Zoals hierboven reeds werd aangegeven richten we ons hier alleen op de effecten van ouderlijke steun op het verloop van de schoolloopbaan van jongeren.

Steun, socialisatie en schoolprestaties van jongens en meisjes. Rollins en Thomas (1979) suggereerden dat meisjes die op school goed presteren het product lijken te zijn van een slechte socialisatie door de ouders. Terwijl het hebben ervaren van ouderlijke steun voor zowel jongens als meisjes gewoonlijk samen gaat met positieve zaken als een interne beheersingsoriëntatie en meer gevoelens van competentie (bijvoorbeeld Grolnick et al., 1991), een harmonieuze ouder-kindrelatie (Rueter & Conger, 1995; Steinberg, 1990), en weinig probleemgedrag (zoals drank- en druggebruik en het al op jonge leeftijd seksueel actief zijn, Barnes & Farrell, 1992; Steinberg, 1990; Taris & Semin, in druk), blijkt steun bij jongens geassocieerd met goede, maar bij meisjes juist met relatief slechte schoolprestaties.

Een mogelijke verklaring voor deze bevindingen kan worden gevonden in de theorie van Eccles (1987; zie ook Bügel, 1991, en Vloet, 1992), die weer gebaseerd is op Feather's (1982) waarden-verwachtingstheorie. Eccles (1987) stelt dat de schoolprestaties van kinderen beschouwd moeten worden als de uitkomsten van een complex sociaal-psychologisch (beslissings-)proces. De waarschijnlijkheid dat men voor een bepaalde uitkomst kiest (zoals hard werken om een hoog cijfer te behalen) worden door haar in verband gebracht met (1) de subjectieve waarde die een bepaalde keuze heeft voor het individu, en (2) de succesverwachting ten aanzien van de betreffende keuze. Bij het bepalen van de *subjectieve waarde* worden de ter beschikking staande opties door het

individueel tegen elkaar afgewogen. Een bepaalde optie kan als meer of minder aantrekkelijk worden beschouwd op basis van de geschatte opbrengsten en kosten, in termen van bijvoorbeeld het kunnen bereiken van bepaalde lange- of korte-termijndoelen, of door het onderstrepen van bepaalde persoonlijke waarden (bijvoorbeeld het bevestigen van de eigen sekserolidentiteit). Omdat mannen en vrouwen een andere waardenhiërarchie hebben – die gebaseerd is op de ideaaltypische sekserolidentiteit, en wordt overgebracht via een proces van socialisatie en internalisatie van normen en waarden: bij mannen staan geld verdienen en carrière maken centraal, vrouwen gaan ervan uit dat zij primair bestemd zijn voor de moederrol – zullen jongens en meisjes dus verschillen met betrekking tot de subjectieve waarde die zij toekennen aan het hebben van succes op school.

Ook de *succesverwachting* speelt een rol bij de keuze van jongens en meisjes. Meisjes zouden volgens Eccles (1987) harder moeten werken om goede resultaten te behalen, en daarom geneigd zijn om af te zien van het willen bereiken van bepaalde doelen (zoals het hebben van succes op school). Ook verwachten zij veel moeite te moeten doen om de rol van moeder en carrière maker te combineren (Taris, in druk; deze verwachting wordt geschraagd door de feiten, zie bijvoorbeeld Taris, Heesink & Feij, 1995), terwijl voor mannen de rol van vader en carrière maker elkaar niet in de weg staan.

Samenvattend komt Eccles' (1987) theorie erop neer dat goed-gesocialiseerde vrouwen van huis uit een andere waardenhiërarchie meekrijgen dan mannen (Brooks-Gunn & Matthews, 1979); bovendien verwachten zij harder te moeten werken en desondanks weinig succesvol te zijn bij het behalen van bepaalde doelen. Het is daarom voor meisjes minder vanzelfsprekend en nuttig om goed te (willen) presteren op school dan voor jongens. In deze optiek kan 'schoolsucces' voor meisjes opgevat worden als een teken dat men zich niet wenst te conformeren aan de traditionele vrouwelijke sekserolidentiteit (Huston, 1983), en dientengevolge als het resultaat van een slechte sekserol socialisatie. Overigens is het goed erop te wijzen dat de 'traditionele' sekserolidentiteit van zowel jongens als meisjes de laat-

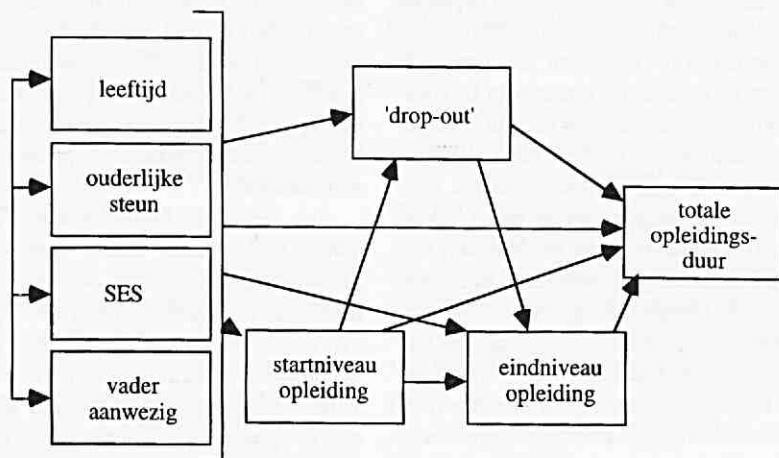
ste decennia onder druk staat (meisjes mogen tegenwoordig best assertief en zelfstandig zijn), wat betekent dat theorieën die stellen dat goede schoolprestaties van ambitieuze meisjes het gevolg zijn van een 'weinig succesvol verlopen socialisatieproces' inmiddels wat ouderwets aandoen.

Een model. In Figuur 1 wordt het in deze studie onderzochte model gepresenteerd. In principe verwachtten we een positieve relatie tussen de hoeveelheid steun die respondenten van hun ouders ervoeren, en de in deze studie gehanteerde indicatoren van wat een 'succesvolle schoolloopbaan' genoemd kan worden: een hoog aanvangsniveau na het verlaten van de lagere school, relatief weinig 'drop-out', een hoog eindniveau, en – gegeven het aanvangs- en het behaalde eindniveau van de scholing – een relatief korte studieduur. Deze relaties zullen gemodereerd worden door sekse. Op basis van de hierboven besproken literatuur lijkt het plausibel te veronderstellen dat er verschil in de sterkte van de relaties tussen opvoedingsstijl en opleidingsuitkomst zullen optreden, op een dusdanige manier dat mannen méér profijt van ouderlijke steun zullen hebben dan vrouwen.

De verwachte relaties tussen de opleidingsuitkomsten *onderling* liggen voor de hand. Een hoog aanvangsniveau zal enerzijds de kans op een hoog eindniveau vergroten, maar anderzijds zal de kans om de opleiding zonder diploma te verlaten eveneens toenemen. Personen die vaak een opleiding beëindigen zonder een diploma te behalen zullen over het algemeen een lager eindniveau bereikt hebben dan personen die altijd een diploma behaalden. Tenslotte verwachten we dat een hoog eindniveau en één of meer keren een opleiding hebben verlaten zonder diploma te behalen de totale opleidingsduur zal doen toenemen.

Naast de hierboven besproken variabelen is nog een aantal andere, tot nu toe ongenoemd gebleven, variabelen in het model opgenomen. Deze hebben betrekking op kenmerken van de respondent (leeftijd) en van de omgeving (sociaal-economische status, aanwezigheid van de vader).

Aanwezigheid van de vader. Eerder onderzoek heeft laten zien dat probleemgedrag in het algemeen (Rodgers, 1983) en schoolprestaties en 'drop-out' in het bijzonder samenhangen



Figuur 1. Weergave van de veronderstelde causale ordening van de variabelen in deze studie. Het model wordt voor mannen en vrouwen afzonderlijk onderzocht.

met de aanwezigheid van de vader in het gezin (Duncan & Duncan, 1969; Astone & McLanagan, 1994; Hauser & Phang, 1993). Dit effect kan op verschillende manieren worden geïnterpreteerd. Newcomer en Udry (1987) suggereren dat de emotionele en praktische problemen die een scheiding met zich meebrengt zich kunnen manifesteren in probleemgedrag van de zijde van de adolescent. Stern, Northman en Van Slyck (1984) stellen dat de vader een sleutelfiguur is bij de overdracht van waarden en normen, dat hij voor de (mannelijke?) adolescent als een rolmodel fungeert, en dat juist de vader gewoonlijk de ouder is die sancties uitdeelt wanneer het kind zich onwenselijk gedraagt. De relatieve bijdrage van deze interpretaties tot de verklaring van de relatie tussen de aanwezigheid van de vader en de schoolprestaties van de adolescent is echter nog onbekend; in principe verwachten we echter dat de afwezigheid van de vader de kans op problemen op school vergroot.

Sociaal-economische status. Er bestaan sterke relaties tussen enerzijds schoolsucces en prestatie-motivatie, en anderzijds sociaal-economische status (Meijnen & Riemersma, 1992; White, 1982). Kinderen afkomstig uit een lager sociaal-economisch milieu blijken over het algemeen lagere cijfers te behalen en vaker te doubleren dan kinderen uit de midden- en hogere milieus (Driessen, 1994); ook staan zulke kinderen wat minder positief ten opzichte van school (Weiner, 1982) en bleken er bij

twalfjarigen al duidelijke klasseverschillen te onderscheiden in hun prestatie-motivatie (Harvey & Kerin, 1978). Tenslotte is bekend dat ouders uit de lagere klassen minder geneigd zijn het belang van een goede scholing te benadrukken (bijvoorbeeld Katz, 1967). We verwachten daarom dat sociaal-economische status positief zal samenhangen met schoolsucces.

1 Methode

1.1 Steekproef

De gegevens werden verzameld in een representatieve en longitudinale studie onder 1775 jongeren. De respondenten waren gelijkelijk verdeeld over sekse en de geboortecohorten 1961, 1965, en 1969. De eerste meting vond plaats gedurende de winter van 1987/1988. De respondenten werden bij hen thuis geïnterviewd door getrainde interviewers, met behulp van een gestructureerde vragenlijst. Gedurende het interview kwamen attitudes, meningen en gedrag ten aanzien van verschillende levensdomeinen (zoals relatievorming, scholing, en werk) aan de orde. Ook vulden alle subjecten een vragenlijst in, waarin persoonlijkheids- en achtergrondvariabelen werden gemeten.

De tweede meting vond vier jaar later plaats (winter 1991/1992). De opzet van deze meting was vrijwel identiek aan die van de eerste meting. Van de oorspronkelijke 1775 respondenten deden 1257 personen (71%) wederom

mee. Analyse van de nonresponse liet zien dat deze niet gerelateerd was aan sekse, sociaal-economische status, huwelijkse staat, of leeftijd. Wel bleek dat hoger opgeleiden iets waren oververtegenwoordigd (zoals vaak het geval is in survey-onderzoek, vergelijk Goyder, 1987; Taris, 1996).

De aldus verzamelde gegevens stelden ons in staat na te gaan hoe de opleidingsloopbaan in het verleden verlopen was. Nadat observaties met missende waarden op één of meer variabelen listwys verwijderd waren, bedroeg de uiteindelijke steekproefgrootte 986 subjecten.

1.2 Variabelen

Opvoedingsstijl. De manier van opvoeden door de ouders werd subjectief gemeten, door de respondent voor zowel de vader als de moeder drie items afkomstig uit Parker, Tupling en Brown's (1979) 'parental bonding instrument' te laten beantwoorden (in totaal beantwoordden zij dus zes items). De items werden uiteraard eerst in het Nederlands vertaald.

De respondenten werd expliciet gevraagd om zich de ouders te herinneren zoals zij zich gedurende de *eerste twaalf jaren* van het leven van de respondent gedroegen. De mate waarin men steun van de ouders ervoer werd gemeten met de volgende items: 'mijn ouders gaven er blijk van mijn problemen en zorgen te begrijpen', 'mijn ouders gedroegen zich liefdevol tegen mij', en 'mijn ouders schenen niet te begrijpen wat ik wilde en nodig had' (omgecodeerd, (1 = 'zeer mee eens', 5 = 'zeer mee oneens')). De betrouwbaarheid van deze schaal (Cronbach's α) was .84.

Het *aanvangsniveau* en *eindniveau van de opleiding* werden geoperationaliseerd in termen van de CBS-classificatie (CBS, 1987). Het betreft hier een 9-puntsschaal waarin vergelijkbare typen opleidingen zijn gegroepeerd. Het *aanvangsniveau* van de opleiding varieerde tussen 2 (voortgezet algemeen onderwijs: VGLO, LAVO) en 6 (HAVO/VWO), met tussenliggende categorieën LBO/VBO (score 3), MAVO/4 jaar middenschool (score 4), en MBO (score 5). Het *eindniveau* van de opleiding varieerde tussen 2 en 9 (WO, afgesloten met het hebben behaald van de doctorstitel), waarbij HBO de score 7 had en score 8 aangaf dat de respondent een universitaire opleiding had afgerond, echter zonder gepromoveerd te

zijn.

De lengte van de *periode dat men een voltijdsstudie volgde* werd geoperationaliseerd als het aantal jaren liggende tussen het moment waarop men het secundaire onderwijs betrad, en weer verliet. Men werd geacht het voltijds-onderwijs te hebben verlaten als men (1) een bepaalde opleiding verlaten had (al dan niet met diploma), en (2) niet binnen één jaar weer aan een nieuwe (voltijds-)opleiding begon. Op deze manier wordt vermeden dat een kortstondige onderbreking van het volgen van onderwijs (bijvoorbeeld omdat men eerst een jaar in het buitenland wil werken alvorens te gaan studeren) gezien wordt als de afsluiting van de periode dat men voltijds-onderwijs volgt.

Een bezwaar van deze operationalisering is dat een langere opleidingsduur zowel schoolsucces als falen kan aanduiden. Immers, een aantal malen doorstromen naar een hoger opleidingstype (bijvoorbeeld MAVO-HAVO-HBO-WO) is een teken van succes: een aantal malen blijven zitten indiceert problemen op school. In beide gevallen is echter een langere opleidingsduur het resultaat. Als het aanvagen en eindniveau van de opleiding echter constant gehouden worden, zal een langere opleidingsduur echter voornamelijk falen indiceren.

Het *aantal malen dat men een studie verliet zonder een diploma te behalen* bleek een tamelijk scheef verdeelde variabele (scheefheid was 1.35; het gemiddelde aantal malen dat zulks voorkwam was voor alle respondenten .80, bij een bereik van 0 tot 6). Omdat een dergelijke scheefheid tot problemen zou kunnen leiden bij het schatten van parameters (veel analysemethoden gaan ervan uit dat de gegevens multivariaat normaal zijn verdeeld), werd deze variabele getransformeerd (we namen het natuurlijke logaritme). Dit had een weinig problematische scheefheid van .26 als resultaat.

Achtergrondvariabelen. Dit cluster van variabelen bevatte de volgende variabelen. Ten eerste, we namen mee of de *vader van de respondent* aanwezig was geweest gedurende de eerste twaalf levensjaren van de respondent. Ten tweede, de *sociaal-economische status* van het gezin, gebaseerd op drie indicatoren: (a) het beroepsniveau van de vader (gebaseerd op de classificatie van Bakker, Jonker & Oud, 1986); (b) het opleidingsniveau van de vader; en (c) het opleidingsniveau van de moeder. De drie

variabelen werden gestandaardiseerd en vervolgens gemiddeld. Als er sprake was van afwezigheid van de vader gedurende de eerste twaalf jaren van het leven van de respondent, werd de score gebaseerd op de variabelen waarvoor wel valide scores beschikbaar waren. De betrouwbaarheid van deze schaal bedroeg .83 (Cronbach's α).

Tabel 1 geeft de gemiddelden en standaardafwijkingen van de in deze studie gebruikte variabelen, voor mannen en vrouwen afzonderlijk. De tabel laat tevens zien dat er voor drie van de vier afhankelijke variabelen verschillen bestaan tussen de groepen. Mannen blijken op een lager niveau te zijn begonnen dan vrouwen (vergelijk Driessen, 1994), maar desondanks een hoger opleidingsniveau te hebben bereikt. In overeenstemming hiermee hadden zij ook wat meer tijd nodig om hun studie te voltooien; een en ander is in overeenstemming met eerdere bevindingen (zie het overzicht in Bügel, 1991). Wat betreft de overige variabelen bleken er geen belangrijke verschillen tussen de mannen en vrouwen in onze studie te bestaan.

1.3 Procedure: Specificatie en aanpassing van het model

De gegevens werden geanalyseerd met behulp van de multiple-groep procedure in het LISREL-programma van Jöreskog en Sörbom (1993). Sekse werd als moderatorvariabele opgevat. Omdat we ons richtten op verschillen in de opleidingsloopbanen van mannen en vrouwen analyseerden we de covariantiematrix. Daarnaast werden ook verschillen in gemiddelden van de variabelen met behulp van de LISREL-procedure onderzocht.

In eerste instantie werd getoetst of de

assumptie, dat de variantie-covariantiematrix voor mannen als vrouwen gelijk was, houdbaar was. Dit gaf een chi-kwadraatwaarde te zien van 95.74 met 36 vrijheidsgraden; $p < .001$. Dit resultaat betekent dat de verschillen tussen de covariantiematrices van mannen en vrouwen waarschijnlijk niet aan toevalsfluctuaties te wijten waren. Vervolgens pasten we het model aan zoals weergegeven in Figuur 1, waarbij geen restricties aangebracht werden ten aanzien van de structuur van het model over beide groepen (dat wil zeggen, de parameters van het model werden voor beide groepen afzonderlijk geschat). Omdat dit model geen vrijheidsgraden had, was de passing perfect; het werd verder als uitgangsmodel gebruikt.

Om een eerste indruk van de bron van de verschillen tussen de covariantie-matrices van mannen en vrouwen te verkrijgen, onderzochten we achtereenvolgens of (a) de relaties tussen de exogene variabelen (steun van de ouders, geboortjaar, SES, en vader van de respondent aanwezig, zie Figuur 1) voor beide groepen gelijk waren (in LISREL-termen: de parameters in de phi-matrix werd over beide groepen tegelijk geschat); (b) de effecten van de exogene variabelen op de endogene variabelen (duur van de opleiding, aanvangs- en eindniveau van de opleiding, aantal malen opleiding verlaten) voor beide groepen gelijk waren (in LISREL-termen: de gamma-matrix werd over beide groepen tegelijk geschat); en (c) de effecten tussen de endogene variabelen onderling aan elkaar gelijk waren (de beta-matrix werd over beide groepen tegelijk geschat). De resultaten gaven een significante verslechtering ten opzichte van de passing van het perfect passende uitgangsmodel te zien wanneer de

Tabel 1
Gemiddelden en standaardafwijkingen van de variabelen in deze studie, voor mannen en vrouwen afzonderlijk

variabelen	mannen (N = 485)		vrouwen (N = 509)	
	M	SD	M	SD
duur opleiding*	11.41	3.79	10.43	3.62
eindniveau opleiding*	5.59	1.86	5.36	1.76
aantal malen opleiding verlaten zonder diploma	.65	.64	.61	.61
startniveau opleiding*	4.32	1.20	4.47	1.20
vader aanwezig	.94	.24	.95	.22
sociaal-economische status	.04	.84	-.02	.82
geboortjaar	64.81	3.30	65.02	3.29
ouderlijke steun	3.37	.59	3.40	.59

* Significant verschil tussen mannen en vrouwen, $p < .05$ (via T-toets).

gamma- of beta-matrix over beide groepen tegelijk werd geschat (respectievelijk een verslechtering van 52.78 chi-kwadratpunten bij 16 vrijheidsgraden, $p < .001$, en van 26.46 bij 6 vrijheidsgraden, $p < .001$). Dit geeft aan dat er op deze onderdelen van het model belangrijke verschillen tussen mannen en vrouwen bestaan. De veronderstelling dat de relaties tussen de exogene variabelen gelijk waren bleek echter goed houdbaar (χ^2 met 10 *df* was 8.82; $p = .55$).

Vervolgens werd op basis van de standaardfouten en modificatie-indices nagegaan welke relaties voor beide groepen aan elkaar gelijk gesteld konden worden. Dit bleek het geval voor het overgrote deel van de in het model aanwezige effecten. Vervolgens werden effecten die niet significant afweken van nul uit het model verwijderd. Het uiteindelijke model leverde een chi-kwadratwaarde op van 45.56 bij 40 vrijheidsgraden, $p = .25$, NNFI = .99². Deze waarden geven aan dat het model, gezien de omvang van de steekproef, zeer goed houdbaar is.

Afsluitend werd via LISREL onderzocht of de gemiddelden van de variabelen voor beide groepen gelijk waren. Het voor beide groepen aan elkaar gelijkstellen van de gemiddelden resulteerde in een significante toename van 28.60 chi-kwadratpunten bij 8 vrijheidsgraden. De modificatie-indices lieten zien dat dit vooral werd veroorzaakt door verschillen in de duur en het eindniveau van de opleiding (mannen deden significant langer over hun opleiding, maar hadden een belangrijk hoger opleidingsniveau behaald wanneer zij het voltijds onderwijs verlieten). Voor de overige variabelen werden geen belangrijke verschillen gevonden. Deze resultaten komen grotendeels overeen met wat de *T*-toetsen in Tabel 1 al suggereerden, al valt het verschil in het aanvangsniveau van de opleiding weg; mannen en vrouwen verschillen dus niet van elkaar wat betreft het aanvangsniveau van de opleiding, wanneer er gecontroleerd wordt voor de overige in het model opgenomen variabelen.

2 Resultaten

Tabel 2 geeft de volledig gestandaardiseerde maximum likelihoodschattingen voor het uiteindelijke model. Deze effectschattingen zijn

voor beide groepen in dezelfde schaal uitgedrukt, en kunnen dus zonder meer met elkaar vergeleken worden.

De resultaten laten zien dat het startniveau van de opleiding een positief effect had op het eindniveau (een effect van .45, $p < .001$). Het eindniveau van de opleiding werd daarnaast negatief beïnvloed door het aantal malen dat men een opleiding, zonder een diploma behaald te hebben, verlaten heeft (-.15, $p < .001$). Dit gold voor zowel mannen als vrouwen. Betreffende de duur van de opleiding vonden we echter belangrijke verschillen. In principe vonden we dat de duur van de opleiding positief werd beïnvloed door het eindniveau van de opleiding en door het aantal malen dat men een opleiding zonder diploma verlaten had. Voor mannen bleken deze effecten echter aanzienlijk minder sterk dan voor vrouwen (voor mannen waren deze effecten respectievelijk .19 en .08; voor vrouwen .38 en .23). De verschillen tussen deze effecten zijn significant op $p < .001$. Met andere woorden, voor vrouwen waren voor de hand liggende factoren als het te behalen opleidingsniveau en het aantal malen dat men een opleiding afbreekt méér bepalend voor de uiteindelijke duur van de opleiding dan voor mannen.

Bij de exogene variabelen vinden we allereerst dat de *aanwezigheid van de vader* de kans dat men een studie zonder diploma verlaat wat verkleinde (-.07, $p < .05$). De *sociaal-economische status* van de ouders bleek positief samen te hangen met zowel het startniveau van de opleiding (een effect van .45, $p < .001$) als het eindniveau van de opleiding. Kinderen afkomstig uit een hoog sociaal-economisch milieu blijken dus een relatief hoog opleidingsniveau te bereiken, zelfs als men controleert voor het feit dat zulke kinderen al op een aanzienlijk hoger niveau aan het vervolg op het basisonderwijs begonnen. Met andere woorden, de verschillen tussen de sociaal-economische groepen worden alleen maar groter naarmate kinderen ouder worden. Het gaat ook zeker niet om een onaanzienlijk effect (een gestandaardiseerd effect van .20, $p < .001$).

Verder bleek dat iemands *geboortejaar* een zeer sterk effect had op de duur van de opleiding: ouderen bleken langer op school te blijven. Dit is grotendeels een artefact van de data,

Tabel 2

Maximum Likelihood-effectschattingen voor mannen en vrouwen afzonderlijk (volledig gestandaardiseerd over groepen, χ^2 met 40 df = 45.56; $p = .25$; NNFI = .99)

Tabel 2a: Effectschattingen voor de mannen

variabelen	duur opl.	eindniv. opl.	opl. verlaten	startniv. opl.
eindniveau opleiding	.19***			
aantal malen opl. verlaten	.08*	-.15***a		
startniveau opleiding		.45***a		
vader aanwezig			-.07*a	
sociaal-economische status		.20***a		.45***a
geboortjaar	-.68***		.07*a	
steun ouders			-.21***	
R^2	.49	.32	.06	.20

Tabel 2b: Effectschattingen voor de vrouwen

variabelen	duur opl.	eindniv. opl.	opl. verlaten	startniv. opl.
eindniveau opleiding	.38***			
aantal malen opl. verlaten	.23***	-.15***a		
startniveau opleiding		.45***a		
vader aanwezig			-.07*a	
sociaal-economische status		.20***a		.45***a
geboortjaar	-.50***		.07*a	
steun ouders	.14**			
R^2	.44	.37	.01	.20

* $p < .05$, ** $p < .01$, *** $p < .001$.

a Effectschatting identiek voor beide groepen.

omdat een deel van de jongeren ten tijde van het afronden van de dataverzameling voor dit onderzoek nog met enige studie bezig was (zie Taris & Feij, in druk, voor een uitgebreide beschrijving van de gegevens voor de drie geboortecohorten, alsmede een discussie van de mate waarin dit effect een artefact van de data betreft, dan wel een weergave van veranderende historische omstandigheden). Op dezelfde manier kan het negatieve effect van geboortjaar op het aantal malen dat men een opleiding zonder diploma verliet worden geïnterpreteerd.

Om te onderzoeken in hoeverre deze tekortkoming van de gegevens de resultaten beïnvloedde, hebben we de analyses herhaald op basis van de gegevens van alleen de oudste leeftijdsgroep (dertigjarigen). Vrijwel iedereen in deze leeftijdsgroep had inmiddels het voltijdsonderwijs verlaten. De effecten voor de overige variabelen op de afhankelijke variabelen in ons model bleken in deze groep goed vergelijkbaar met de bevindingen voor de totale groep, al nam de grootte van de standaardfouten uiteraard toe. Al met al lijkt het er dus op dat deze tekortkoming geen ernstige gevolgen heeft gehad voor de te trekken conclusies.

Ten slotte komen we toe aan de effecten van *opvoedingsstijl* op de afhankelijke variabelen in ons onderzoek. Tabel 2 laat zien dat hier aanzienlijke verschillen tussen mannen en vrouwen bestonden. Voor mannen bleek dat ouderlijke steun de kans dat men een opleiding zonder diploma heeft verlaten aanzienlijk reduceerde ($-.21$, $p < .001$). Het corresponderende effect bleek voor vrouwen niet significant van nul af te wijken ($-.04$, $p > .40$), wat onze veronderstellingen ondersteunt. Tenslotte bleek het hebben genoten van veel steun van de ouders voor vrouwen te leiden tot een relatief lange opleidingsduur ($.14$, $p < .01$). Bij de mannen was het corresponderende effect niet significant ($-.04$, $p > .20$); op basis van onze uitgangspunten zouden we voor mannen een kleiner of zelfs negatief effect verwachten. In die zin ondersteunt de afwezigheid van effecten onze veronderstellingen.

Indirecte effecten van steun op de afhankelijke variabelen. Opgemerkt kan worden dat ouderlijke steun niet alleen *directe* effecten op de afhankelijke variabelen heeft, maar ook indirecte effecten, via de andere in het model opgenomen variabelen. Inspectie van de indirecte effecten van steun van de ouders liet ech-

ter zien, dat deze praktisch gesproken te verwaarlozen waren (het betreft effecten die variëren tussen de -.01 en .03).

Verklaarde varianties. De in het model opgenomen variabelen verklaren een aanzienlijk deel van de variantie in de totale opleidingsduur en het behaalde opleidingsniveau (tussen de 32 en 49 procent), al is dit laatste natuurlijk grotendeels veroorzaakt doordat het aanvangsniveau eveneens in het model is opgenomen. Deze laatste variabele wordt dan ook in aanzienlijk mindere mate verklaard (20 procent in beide groepen). De verklaarde variantie in de vierde afhankelijke variabele, het aantal malen dat men een opleiding voortijdig verliet, is bijzonder laag (6 procent voor de mannen, 1 procent voor de vrouwen), dit ondanks het feit dat drie van de vier verklarende variabelen effect hadden op het aantal malen dat men een opleiding tussentijds verliet. Het lijkt er aldus op dat andere, niet in het model opgenomen, variabelen (zoals variabelen op school- en klasniveau, prestatie-motivatie en dergelijke) van meer belang zijn bij de verklaring van deze variabele.

3 Afsluitende opmerkingen

De hier gerapporteerde resultaten ondersteunden de veronderstellingen ten aanzien van de differentiële effecten van ouderlijke steun op de opleidingsloopbaan van jonge mannen en vrouwen. We verwachtten, conform de eerder door Rollins en Thomas (1979) en Eccles (1987) geformuleerde ideeën, dat het ervaren van veel ouderlijke steun voor jongens een positief effect zou hebben op het verloop van hun opleidingsloopbaan (in termen van het aanvangsniveau na de voltooiing van het lager onderwijs, het aantal malen dat men een opleiding zou verlaten zonder een diploma behaald te hebben, het eindniveau van het onderwijs, en de tijd die men nodig zou hebben om de opleiding te voltooien). Voor meisjes werd verwacht dat de effecten van veel ouderlijke steun op het verloop van de opleidingsloopbaan minder positief zouden zijn.

Onze resultaten ondersteunden deze hypothesen gedeeltelijk. We vonden dat voor jongens het ervaren van veel ouderlijke steun het minder waarschijnlijk maakte dat zij een

bepaalde opleiding voortijdig en zonder diploma zouden verlaten. Voor meisjes bleek het ervaren van veel steun het daarentegen waarschijnlijker te maken dat zij relatief veel tijd nodig zouden hebben om hun opleiding te voltooien. Deze resultaten zijn in lijn met Eccles' (1987) interpretatie van de suggestie van Rollins en Thomas (1979), dat een gunstig verlopen socialisatieproces voor meisjes relatief nadelige gevolgen heeft voor het verloop van hun opleidingsloopbaan. Wel moet hierbij worden aangetekend dat de invloed van het ervaren van veel ouderlijke steun relatief beperkt is. Het effect van – bijvoorbeeld – sociaal-economische status op de meeste van de hier onderzochte afhankelijke variabelen is aanzienlijk groter (zoals ook Bügel, 1991, concludeert).

Sociaal-economische status. Eén van de meer opvallende bevindingen van onze studie was dat sociaal-economische status van de ouders niet alleen – zoals verwacht – een sterk effect uitoefende op het niveau van de opleiding zoals gekozen onmiddellijk na het basisonderwijs, maar dat er ook nog een belangrijk effect op het eindniveau van de opleiding gevonden werd. Een en ander suggereert dat de effecten van sociaal-economische status op het opleidingsniveau van kinderen zich gedurende het *hele* opleidingstraject doen voelen, en niet alleen op een discreet aangrijpingspunt als de keuze van de opleiding na de basisvorming. Een en ander is waarschijnlijk te wijten aan het feit dat ouders uit gezinnen van minder hoge sociaal-economische status minder belang hechten aan het verwerven van een goede opleiding (Katz, 1967), waardoor zij hun kinderen ook tijdens een opleiding minder stimuleren goede resultaten te behalen: de voor-sprong van kinderen uit de hogere klassen wordt aldus steeds groter.

Aanwezigheid van de vader. Eerder werd verondersteld dat de afwezigheid van de vader nadelig zou werken voor het schoolsucces, conform eerdere studies waarin sterke relaties tussen probleemgedrag en gezinssituatie werden gerapporteerd. In onze studie vonden we dat respondenten uit gebroken gezinnen meer kans hadden voortijdig een opleiding te verlaten. Het betrof hier echter een klein en weinig systematisch ($.01 < p < .05$) effect; voor de overige drie indicatoren van het verloop van de schoolloopbaan vonden we geen effect. Het

lijkt er aldus op dat de aan- of afwezigheid van de vader inderdaad een nadelige invloed heeft op het verloop van de schoolloopbaan van de kinderen, maar dat de invloed van deze variabele zeker niet overschat moet worden.

Methodologische kanttekeningen

Een belangrijke beperking van deze studie is dat ouderlijke steun *subjectief* en *retrospectief* werd gemeten. Dit impliceert dat de verklarende variabele (steun) deels nà de te verklaren variabelen (verloop van de opleidingsloopbaan) werd gemeten, wat de mogelijkheid opent dat de evaluatie van de opvoedingsstijl van de ouders door het subject beïnvloed is door het verloop van de opleidingsloopbaan. Het is aldus in principe mogelijk dat een respondent het eigen falen met betrekking tot de scholing wijt aan de manier van opvoeden van de ouders.

We achten dit om twee redenen onwaarschijnlijk. Ten eerste, de items betreffende de opvoedingsstijl gedurende de eerste twaalf jaar van het leven van de respondent werden opzettelijk in de schriftelijke vragenlijst opgenomen (deze werd voorafgaand aan het interview ingevuld), terwijl het verloop van de opleidingsloopbaan pas gedurende het interview zélf aan de orde kwam. Omdat zo de verklarende variabele vóór de te verklaren variabelen werd gemeten, is het minder waarschijnlijk dat de evaluatie van de opvoedingsstijl van hun ouders door de respondenten werd beïnvloed door een mogelijke teleurstelling ten aanzien van het verloop van de opleidingsloopbaan.

Ten tweede, omdat de gegevens tijdens twee metingen werden verzameld was het mogelijk om voor een beperkt aantal respondenten (namelijk diegenen die nog een opleiding volgden op het moment van de eerste meting) het verloop van de opleidingsloopbaan tussen de eerste en tweede meting te relateren aan de tijdens de eerste meting gemeten subjectieve ouderlijke steun. Dit gaf een vergelijkbaar patroon van effecten te zien, al waren de standaardfouten van de effecten aanzienlijk groter vanwege een kleinere steekproef en een wat beperktere variantie betreffende de afhankelijke variabelen. Het lijkt er aldus op dat, *als* er al sprake was van contaminatie van de afhankelijke en onafhankelijke variabelen, een en ander geen majeure gevolgen voor onze resultaten heeft gehad.

Ook zou men nog kunnen opmerken dat het gebruik van subjectieve oordelen betreffende de opvoedingsstijl de resultaten onwenselijk beïnvloed zou kunnen hebben. Ten eerste is het zo dat de oordelen van kinderen over de opvoedingsstijl van hun ouders gewoonlijk positief correleren met wat ouders zelf zeggen te doen (al is deze correlatie niet altijd zo hoog als men zou wensen; zie bijvoorbeeld Grolnick et al., 1991). Ten tweede, een aantal theoretici stelt dat juist het eigen fenomenologische gezichtspunt van de kinderen van belang is bij het verklaren van hun schoolprestaties (Bronfenbrenner, 1977). Het lijkt zelfs goed mogelijk dat objectief gemeten opvoedingsstijlen de schoolprestaties alleen indirect, via de percepties van deze opvoedingsstijlen door de kinderen, beïnvloeden (Grolnick et al., 1991). Het lijkt er aldus op dat wat de één als een zwakte van deze studie zou kunnen opvatten, door een ander juist als een kracht gezien wordt.

Ten slotte willen we benadrukken dat in onze studie een 'algemene' maat bij de meting van steun werd gebruikt, namelijk in termen van emotionele steun en begrip. Hoewel de hiermee behaalde resultaten grotendeels in de lijn van onze verwachtingen lagen, zou 'steun' ook in termen van 'stimuleren tot presteren' geformuleerd kunnen worden: een dergelijke operationalisering zou wellicht beter aansluiten bij de onderzoeksvraagstelling. In principe zouden we voor een dergelijke, meer gepreciseerde, operationalisering dezelfde verwachtingen formuleren, zij het dat verwacht mag worden dat er dan sterkere effecten van 'steun' op de hier gebruikte afhankelijke variabelen gevonden zullen worden.

Samenvattend: de huidige studie liet zien dat kenmerken van de opvoeding (in casu het hebben ervaren van steun door de ouders) het verloop van de opleidingsloopbaan van jongens en meisjes systematisch beïnvloedt, alsmede dat deze effecten voor jongens en meisjes aanzienlijk van elkaar verschillen. Waar de schoolloopbaan van jongens lijkt te profiteren van het ervaren van ouderlijke steun is dit voor meisjes niet het geval. Meisjes die veel steun van hun ouders ervaren blijken een opleiding vaker voortijdig te verlaten en relatief lang over hun opleiding te doen.

Dit suggereert dat ouderlijke steun zich vertaalt in ouderlijke gedragingen die verschillen

voor jongens en meisjes. Waar 'ouderlijke steun' voor jongens lijkt te betekenen dat zij worden gestimuleerd om dóór te zetten, ook als het even moeilijk gaat met de studie, lijkt 'steun' voor meisjes juist te betekenen dat ze mogen 'uithuilen en opnieuw beginnen'. In vervolgonderzoek lijkt het daarom raadzaam onderscheid te maken tussen diverse vormen van steun die men kan ervaren (bijvoorbeeld 'algemene' emotionele steun, steun bij het willen presteren, enzovoort). Dergelijk onderzoek zou kunnen uitwijzen of steun voor jongens en meisjes iets anders betekent, alsmede in hoeverre verschillende typen steun zijn gerelateerd aan schoolprestaties. Tevens lijkt het zinvol om in verder onderzoek onderscheid te maken naar de bron van de ervaren steun (vader versus moeder), in combinatie met de sekse van de respondent.

Noten

- 1 Opvallend genoeg wordt in recent Amerikaans onderzoek naar de relatie tussen bijvoorbeeld prestatie-motivatie dan wel schoolcijfers enerzijds, en opvoedingsstijlen van de ouders anderzijds, geen onderscheid meer gemaakt tussen jongens en meisjes (bijvoorbeeld Grolnick et al., 1991; Gottfried et al., 1994). De in eerder onderzoek gevonden interacties met sekse worden blijkbaar niet meer gevonden of relevant geacht, wat – gezien eerdere resultaten – de vraag oproept of de in deze recente studies gepresenteerde resultaten wellicht vertekend zijn door een verkeerde (onvolledige) specificatie van het onderzoeksmodel.
- 2 We rapporteren hier ook de resultaten van Bentler en Bonett's (1980) non-normed fit index (NNFI). Uit Monte Carlo-onderzoek van Marsh, Balla en McDonald (1988) bleek dat veel van de gebruikelijke passingsmaten – inclusief de standaard chi-kwadraattoets – zeer gevoelig waren voor variaties in de omvang van de steekproefgrootte. Van de 30 door hen onderzochte passingsmaten bleek NNFI het minst gevoelig. Bentler en Bonett (1980) adviseren een ondergrens van .90 aan te houden; lagere waarden geven aan dat modellen nog aanzienlijk verbeterd kunnen worden.

Literatuur

- Astone, N.M., & McLanagan, S.S. (1994). Family structure, residential mobility, and school drop-out: A research note. *Demography*, 31, 575-584.
- Bakker, B.F.M., Jonker, J.K., & Oud, J.A. (1988). *Een alternatieve aggregatie van beroepen op basis van de beroepenclassificatie 1984*. Voorburg: CBS.
- Barnes, G.M., & Farrell, M.P. (1992). Parental support and control as predictors of adolescent drinking, delinquency, and related problem behaviors. *Journal of Marriage and the Family*, 54, 763-776.
- Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 32, 513-531.
- Brooks-Gunn, J., & Matthews, W.S. (1979). *He & she: How children develop their sex-role identity*. Englewood Cliffs, NY: Prentice-Hall.
- Bügel, K. (1991). Sekseverschillen in onderwijsprestaties in Nederland: Een overzicht van de literatuur en enkele nieuwe gegevens. *Pedagogische Studiën*, 68, 350-370.
- CBS (1987). *Standaard Onderwijs indeling*. Voorburg: CBS.
- Dornbusch, S.M., Ritter, P.L., Leiderman, P.H., Roberts, D.F., & Fraleigh, M.J. (1987). The relation of parenting style to adolescent school performance. *Child Development*, 58, 1244-1257.
- Driessen, J.G. (1994). Doubleren en drop-out in de eerste fase van het voortgezet onderwijs. *Pedagogische Studiën*, 71, 108-119.
- Duncan, B., & Duncan, O.D. (1969). Family stability and occupational success. *Social Problems*, 16, 286-301.
- Eccles, J.S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly*, 11, 135-172.
- Farmer, H. (1985). Model of career and achievement motivation for men and women. *Journal of Counseling Psychology*, 32, 363-390.
- Feather, N.T. (1982). *Expectations and actions. Expectancy-value models in psychology*. Hillsdale: Erlbaum.
- Gerris, J.M.R., & Janssens, J.M.A.M. (1987). Opvoedingsgedrag van ouders in overtredingssituaties: Een onderzoek naar beïnvloedende factoren. *Pedagogische Studiën*, 64, 295-307.

- Gottfried, A.E., Fleming, J.S., & Gottfried, A.W. (1994). Role of parental motivational practices in children's academic intrinsic motivation and achievement. *Journal of Educational Psychology, 86*, 104-113.
- Goyder, J. (1987). *The silent minority*. Cambridge: Blackwell.
- Grolnick, W.S., & Ryan, R.M. (1989). Parent styles associated with children's self-regulation and competence in school. *Journal of Educational Psychology, 81*, 143-154.
- Grolnick, W.S., Ryan, R.M., & Deci, E.L. (1991). Inner resources for school achievement: motivational mediators of children's perceptions of their parents. *Journal of Educational Psychology, 83*, 508-517.
- Haaf, N. ten (1993). *Opvoedingsdimensies: Convergente en discriminante validiteit*. Nijmegen: Universiteitsdrukkerij Nijmegen.
- Harvey, M.G., & Kerin, R.A. (1978). The influence of social stratification and age on occupational aspirations of adolescents. *Journal of Educational Research, 71*, 262-266.
- Hauser, R.M., & Phang, S.H. (1993). *Trends in high school dropout among white, black, and hispanic youth, 1973 to 1989*. Madison: Institute for Research on Poverty.
- Huston, A.C. (1983). Sex-typing. In P.H. Mussen (red.), *Handbook of child psychology* (4th edition, pp. 338-433). New York: Wiley.
- Jöreskog, K.G., & Sörbom, D. (1993). *LISREL-8 (computer manual)*. Chicago: Scientific Software.
- Katz, I. (1967). The socialization of academic motivation in minority group children. *Nebraska Symposium on Motivation, 15*, 133-191.
- Marsh, H.W., Balla, J.R., & McDonald, R.P. (1988). Goodness-of-Fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*, 391-410.
- Meijnen, G.W., & Riemersma, F.S.J. (1992). *Schoolcarrières: Een klassenkweszie?* Lisse: Swets & Zeitlinger.
- Newcomer, S. & Udry, J.R. (1987). Parental marriage status effects on adolescent sexual behavior. *Journal of Marriage and the Family, 49*, 235-240.
- Parker, G., Tupling, H., & Brown, L.B. (1979). A parental bonding instrument. *British Journal of Medical Psychology, 52*, 1-10.
- Rodgers, J.L. (1983). Family configuration and adolescent sexual behavior. *Developmental Psychology, 24*, 73-83.
- Rollins, B.C., & Thomas, D.L. (1979). Parental support, power, and control techniques in the socialization of children. In W.R. Burr, R. Hill, F.I. Nye en I.L. Neiss (red.), *Contemporary theories about the family: Research-based theories* (vol. I). New York: The Free Press.
- Rueter, M.A., & Conger, R.D. (1995). Antecedents of parent-adolescent disagreements. *Journal of Marriage and the Family, 57*, 435-448.
- Steinberg, L. (1990). Interdependency in the family: Autonomy, conflict, and harmony in the parent-adolescent relationship. In S.S. Feldman & G.R. Elliot (red.), *At the threshold: The developing adolescent* (pp. 255-276). Cambridge, MA: Harvard University Press.
- Stern, M., Northman, J.E., & Van Slyck, M.R. (1984). Father absence and adolescent 'problem behaviors', alcohol consumption, drug and sexual activity. *Adolescence, 19*, 301-312.
- Taris, T.W. (in druk). Fertility in the Netherlands as an expected-value process, and developmental readiness. Te verschijnen in *The Journal of Psychology*.
- Taris, T.W. (1996). Modeling nonresponse in multi-wave panel studies using discrete-time Markov models. *Quality & Quantity, 30*, 189-203.
- Taris, T.W., Heesink, J.A.M., & Feij, J.A. (1995). De continuïteit van de arbeidsloopbaan van vrouwen na het eerste kind: De effecten van de kenmerken van de werkloopbaan. *Mens en Maatschappij, 70*, 54-64.
- Taris, T.W., & Feij, J.A. (in druk). *Shifts in the educational and employment careers of young adults during the eighties: Hard times, and the changing pay-off of education*. Te verschijnen in *International Journal of Adolescence and Youth*.
- Taris, T.W., & Semin, G.R. (in druk). *Parent-child interaction during adolescence, and the adolescent's sexual experience: Control, closeness, and conflict*. Te verschijnen in *Journal of Youth and Adolescence*.
- Vloet, K. (1992). Onderwijs, psychologie en sekse. In T. Top & J. Heesink (red.), *Psychologie en sekse* (pp. 101-119). Houten: Bohn Stafleu Van Loghum.
- Walters, J., & Stinnett, N. (1971). Parent-child relationships: A decade review of research. *Journal of Marriage and the Family, 33*, 70-111.
- Weiner, I.B. (1982). *Child and adolescent psychopathology*. New York: John Wiley & Sons.
- White, K.R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*, 461-481.

Auteurs

I.A. Bok is zelfstandig gevestigd psycholoog te Hilversum. Zij is gespecialiseerd in psychologisch onderzoek van etnische minderheden, jongeren en lichamelijk/psychisch gehandicapten. Daarnaast is zij als onderzoeker verbonden aan het Instituut Werk en Stress te Bilthoven.

T.W. Taris is post-doctoraal onderzoeker aan de Vrije Universiteit, vakgroep Sociale Psychologie. Hij heeft onder meer gepubliceerd op het gebied van de (werk-)socialisatie van jong-volwassenen, alsmede over longitudinale onderzoeksmethoden.

Correspondentie-adres: T. Taris, Vrije Universiteit Amsterdam, Vakgroep Sociale Psychologie, Van der Boechorststraat 1, 1081 BT Amsterdam, e-mail AW.Taris@psy.vu.nl.

Abstract

Parental support, gender, and the development of the educational career: A retrospective study

I.A. Bok & T.W. Taris. *Pedagogische Studietoën*, 1997, 74, 33-45.

The current study examines gender differences in the effects of parental support upon the development of the educational career of young adolescents. In accordance with Rollins and Thomas' (1979) and Eccles' (1987) earlier findings, we hypothesized that the educational career of boys would benefit more from parental support than the careers of girls. These hypotheses were tested using structural equation modelling and data from a retrospective and longitudinal sample of 986 young adults. The results provided moderate support for the hypotheses: boys who indicated to have experienced much parental support experienced significantly less drop-out, whereas girls who experienced much support from their parents needed significantly more time to complete their education. Implications and limitations of the study are discussed.

De Onderwijsresearchdagen 1996 in Tilburg, 4-6 juni 1996

Inleiding

(J.N. Streumer, TO, Universiteit Twente, Enschede)

Van 4 tot en met 6 juni vonden aan de Katholieke Universiteit Brabant (KUB) te Tilburg onder auspiciën van de Vereniging voor Onderwijsresearch (VOR) voor de drieëntwintigste keer de Onderwijsresearchdagen (ORD) plaats. De organisatie was in handen van de sectie Onderwijs- en Opleidingspsychologie van de vakgroep Psychologie van de KUB. De organisatiecommissie stond onder voorzitterschap van prof. Lodewijks.

Afgaande op de deelnemerslijst kan het aantal inschrijvingen worden bepaald op 461.

Mijn persoonlijke indruk, dat er dit jaar minder deelnemers waren dan voorgaande jaren, werd geloochenstraft door de aantallen inschrijvingen die ik vond in vorige versies (1994 en 1995) van de ORD-Kroniek in Pedagogische Studiën. Kennelijk is het aantal deelnemers al jaren stabiel, hetgeen een verdienste kan worden genoemd van de organisatoren. Zij slagen er steeds weer in een aantrekkelijk programma samen te stellen en trotseren de enorme concurrentiestrijd die de afgelopen jaren is ontstaan in het steeds verder uitdijende 'conferentiecircuit'.

De presentaties (paperlezingen, symposia, postersessies, plenaire lezingen, forumdiscussies) waren gegroepeerd in een negental thema's, die in deze kroniek alle zullen worden besproken, te weten: curriculum, onderwijs en samenleving, methodologie en evaluatie, lerarenopleiding en leraarsgedrag, beroeps- en bedrijfsopleidingen, hoger onderwijs en volwasseneneducatie, beleid en organisatie van onderwijs, leren en instructie en tenslotte krachtige leeromgevingen. De eerste acht genoemde thema's sluiten aan bij de divisie-structuur van de VOR; het laatste thema werd door de KUB zelf gekozen. De motivering voor

dit extra thema werd door Lodewijks als volgt gegeven: "We hebben voor dit centrale thema gekozen, omdat wij er via ons werk in toenevende mate van overtuigd raken, dat we aan een fundamentele heroverweging toe zijn ten aanzien van de wijze waarop wij gewend zijn om scholen en scholing in te richten. De "gewone school" slaagt er – ondanks alle goede bedoelingen en inzet – nauwelijks nog in om leerlingen tot het hoogwaardige leren te brengen waar het eigenlijk om zou moeten gaan" (Programmaboek ORD 1996, p. 2).

Zou prof. De Klerk een soortgelijke gedachtingang hebben gehad toen hij in zijn openingstoespraak op dinsdagavond de knuppel in het hoenderhok gooide met de mededeling dat de onderwijskunde in Nederland, wat hem betreft, zijn langste tijd wel had gehad. Over fundamentele heroverweging gesproken! Volgens De Klerk zullen de kracht en overlevingskansen van de interdisciplinaire onderwijskunde afhangen van de mate waarin individuele wetenschappers excelleren in één van de basisdisciplines van de onderwijskunde. De onderwijskunde moet het dus hebben van monodisciplinaire georiënteerde wetenschappers! Ook heeft De Klerk ernstige twijfels over de kwaliteit van onderwijskundig onderzoek. Onderwijskundigen mogen blij zijn dat hun onderzoek niet is beoordeeld met criteria die zijn toegepast bij de onderzoeksvisitatie psychologie, aldus De Klerk. Daarmee suggereert dat er dan niet veel heel was gebleven van het onderwijskundig onderzoek! Heeft de onderzoeksvisitatiecommissie zijn werk dus toch niet goed gedaan? Inmiddels is op de rede van De Klerk gereageerd in het Tijdschrift voor Onderwijsresearch door Creemers, Lagerweij en Verloop, respectievelijk directeur ICO, voorzitter Kamer Onderwijskunde VSNU en voorzitter VOR (Creemers, Lagerweij & Verloop, 1996).

Op de woensdagmorgen werd gestart met een plenaire lezing van prof. Boekaerts: "De school van de toekomst, of de toekomst van de school". In haar lezing deed Boekaerts verslag van een project dat zij uitvoert in het secundair beroepsonderwijs. Het gaat daarbij om een grootchalig innovatieproject, waarin wordt geprobeerd het 'studiehuisconcept' te implementeren in vier regionale opleidingscentra. Centraal in dit project staat de gedragsverande-

ring van studenten, docenten en de schoolleiding. Het project wordt geflankeerd door een aantal onderzoeken, gericht op onder andere de leerresultaten van de studenten, de ontwikkeling van 'algemene' cognitieve, metacognitieve en regulatievaardigheden, de beoogde gedragverandering bij de doelgroepen en ervaringen die worden opgedaan met een zogenaamd interactief leergroepensysteem.

Donderdagmorgen was het de beurt aan prof. Simons om een plenaire lezing te houden. In zijn lezing, met dezelfde titel als die van Boekaerts, ging Simons in op een groot aantal trends die op dit moment of in de naaste toekomst hun invloed (zullen) uitoefenen op leerprocessen in scholen. Daarnaast schetste hij het gewenste verloop van die leerprocessen: van zelfstandig (samen)werken (fase 1), via zelfstandig leren (fase 2) naar zelfverantwoordelijk leren (fase 3). Ligt in de eerste fase nog de nadruk op het uitvoeren van opdrachten door leerlingen/studenten en het geven van feedback daarop door de docent, in de tweede fase staat het 'leren leren' centraal. In de derde fase wordt een begin gemaakt met 'levenslang leren'. Deze leerprocessen kunnen alleen worden gerealiseerd als de school werkt vanuit het paradigma van de lerende organisatie. Interessante onderzoeksthema's die zijn verbonden met voornoemd 'procesgericht onderwijs' zijn volgens Simons onder andere: de ontwikkeling van leerstijlinstrumenten, de ontwikkeling van discipline-specifieke en cluster-specifieke vaardigheden, het ontwerpen van procesgerichte instructie, zelfregulatie, leerprocessen van docenten.

In deze kroniek zijn de indrukken van een aantal ORD-gangers gegroepeerd en beschreven rond de eerder genoemde negen thema's. Uiteraard geven de verslagen geen volledig en objectief beeld van datgene wat in de thema's aan de orde is gesteld. De persoonlijke belangstelling van de auteurs is richtinggevend geweest bij de selectie en beoordeling van de onderwerpen.

Curriculum

(N. Nieveen, TO, Universiteit Twente, Enschede)

De organisatie van het thema "Curriculum" had dit jaar de presentaties en symposia ondergebracht in vier subthema's: (a) aspecten van

leerplanontwikkeling; (b) ontwikkelingsonderzoek op het curriculumdomein en in vakdidactiek; (c) aspecten van evaluatie en aansluiting tussen onderwijs en bedrijfsleven; en (d) internationaal vergelijkend onderzoek op het gebied van wiskunde- en science-onderwijs. In deze terugblik zal ingegaan worden op (onderdelen van) de eerste twee subthema's.

Binnen het eerste subthema stonden strategieën centraal die een bijdrage zouden kunnen leveren aan het systematiseren en rationaliseren van activiteiten binnen het leerplanontwikkelproces. Boersma (SLO) presenteerde in dit kader een procedure (gebaseerd op ideeën van Imelman) die tot meer rationaliteit zou moeten leiden bij de selectie van leerstof. Opvallend bij deze procedure is de centrale rol van criteria die 'cultuurkenners', 'kindkenners' en 'leerkrachtkenners' hanteren bij het bepalen van de relevantie en geschiktheid van leerstofinhouden. De groep cultuurkenners kan daarbij bestaan uit één of meerdere journalisten, vertegenwoordigers van kerkgenootschappen, politici, filosofen en vakbondsmedewerkers. De groep kindkenners kan bestaan uit ontwikkelings- en leerpsychologen, pedagogen, schooldirecteuren, leerkrachten en deelnemers in ouderraden. Bij de groep leerkrachtkenners kan gedacht worden aan schoolbegeleiders, PABO-docenten en medewerkers van landelijke pedagogische centra. Het betrekken van deze kenners bij leerstofselectie contrasteert met de praktijk in Nederland die laat zien dat vooral deskundigen (ontwikkelaars, opleiders en onderzoekers), overheid (als opdrachtgever) en dienstverleners (leraren en hun begeleiders) betrokken zijn bij de onderhandelingen over de ontwikkeling van kerndoelen en examenprogramma's. Om voor een betere verankering te zorgen van leerstofkeuzen in de samenleving, werd de voornoemde procedure toegepast binnen het SLO-project "Aardrijkskunde in de PABO". Een interessante uitkomst was dat de redeneringen en de daaraan ontleende criteria in eerste instantie nauwelijks vakspecifiek bleken te zijn. Hoewel het hier om een onderzoek ging van beperkte omvang, wordt voorgesteld om de reikwijdte van de criteria uit te breiden naar de overige zaakvakken. Overigens dient hierbij wel opgemerkt te worden dat met deze procedure geen antwoord wordt gegeven op het

lastige probleem van de prioriteitsvraag (welke leerstof voldoet het meest aan de criteria).

Krabbe (UT) ging in haar bijdrage in op de mogelijkheden tot standaardisatie van leerplanontwikkelactiviteiten van de afdeling beroeps- onderwijs en volwasseneneducatie (BVE) van de SLO. Bij deze afdeling zijn voornamelijk professionele ontwikkelaars werkzaam, die op basis van relatief unieke en persoonsgebonden kennis en ervaringen en met een grote mate aan vrijheid hun taken uitvoeren. Hoewel het ontwikkelen van standaarden voor professionals in bepaalde opzichten een hachelijke zaak lijkt te zijn, probeerde Krabbe aannemelijk te maken dat standaarden zowel professionals als de organisatie van een gelegenheid voorzien om professionele kennis te optimaliseren en te delen met anderen. Uit eerste deelonderzoeken blijkt dat er binnen de afdeling BVE een zekere bereidwilligheid is voor standaardisatie. Het is daarbij echter essentieel dat gestandaardiseerde methoden en instrumenten aansluiten bij de werkwijzen van professionele ontwikkelaars.

Het tweede subthema "Ontwikkelingsonderzoek op het curriculumdomein en in vakdidactiek" had de vorm van twee opeenvolgende symposia. De mogelijkheid om twee keer anderhalf uur aan één onderwerp te kunnen besteden, kwam de discussie zeker ten goede. In tegenstelling tot ander onderwijskundig onderzoek, dat veelal een beschrijvend of verklarend karakter heeft, start ontwikkelingsonderzoek met een vraag die afgeleid is van een praktisch onderwijskundig probleem. Dit kunnen onder andere curriculaire problemen (bijvoorbeeld: 'Hoe kan in een nascholingscursus de implementatie van een constructivistische benadering van natuuronderwijs gestimuleerd worden?') of didactische problemen zijn (bijvoorbeeld: 'Hoe kan ervoor gezorgd worden dat leerlingen bepaalde natuurkundige principes, waar ze moeite mee hebben, beter leren?'). Beide invullingen van ontwikkelingsonderzoek kregen een plaats binnen de afzonderlijke symposia.

Het eerste symposium, getiteld "Ontwikkelingsonderzoek op het curriculumdomein", werd verzorgd door Van den Akker, Van den Berg, Nieveen en Roes (allen UT). In een inleidende presentatie vond een oriëntatie plaats op de kenmerken van ontwikkelingsonderzoek.

Met deze onderzoeksmethode wordt een meer- voudige wetenschappelijk opbrengst nage- streefd:

- een kwalitatief hoogwaardig prototype van een curriculaire product, waarvan empirisch is aangetoond dat succesvolle implementatie mogelijk is;
- ontwerpmethodologische kennis (technologie) over het ontwerpen, evalueren en implementeren van dergelijke producten in de vorm van 'handelingsondersteuning' (bijvoorbeeld heuristieken en instrumenten) voor andere ontwerpers.

Binnen ontwikkelingsonderzoek vindt een combinatie plaats van ontwikkelactiviteiten (om tot een hoogwaardig prototype te komen) en onderzoeksactiviteiten (om de kwaliteit van het product en de gevolgde ontwerpmethodologie te bepalen). Aan de hand van drie lopende onderzoeksprojecten werd geïllustreerd hoe ontwikkelingsonderzoek binnen het curriculum-domein een concrete invulling kan krijgen.

Het tweede symposium "Ontwikkelingsonderzoek in de praktijk van vakdidactisch onderzoek en leerplanning" werd verzorgd door Lijnse (UU), Boersma (SLO) en Alblas (LUW). Dit onderzoek vertoont op veel punten overeenkomsten met het type ontwikkelingsonderzoek dat in het voorafgaande symposium werd gepresenteerd. Beide vormen van ontwikkelingsonderzoek kunnen elkaar aanvullen. Interessant was de discussie die ontstond rondom de generalisatiewaarde van dergelijk onderzoek. In dat opzicht werd aangegeven dat uitspraken over generalisatie naar andere contexten in eerste instantie maar in beperkte mate mogelijk zullen zijn. Over het algemeen wordt niet gestreefd naar 'statistische', maar naar 'analytische' generalisatie. Hierbij bepalen lezers van de onderzoeksrapportage via analogie-redeneringen op welke wijze de gerapporteerde bevindingen van toepassing zijn voor hun context. In het verlengde hiervan heeft de onderzoeker als taak een zo goed mogelijke explicitering en theoretische verankering te geven van gehanteerde ontwerpprincipes en een zorgvuldige beschrijving te geven van de implementatie-context (ecologische validiteit). Naast analytische generalisatie zal een ontwikkelingsonderzoeker ook dienen te streven naar theoretische generalisatie.

Over het geheel genomen kan gesteld worden dat criteria waaraan ontwikkelingsonderzoek dient te voldoen nader geëxpliciteerd moeten worden, zodat de discussie niet verzandt in de vraag of dergelijk onderzoek relevant is voor de onderwijspraktijk, maar leidt tot een constructieve bijdrage aan de optimalisatie van methoden voor ontwikkelingsonderzoek.

Onderwijs & Samenleving

(A. Booms, SCO-Kohnstamm Instituut, UvA, Amsterdam)

“Onderwijs & Samenleving” als jaarlijks terugkerend thema voor de ORD was dit jaar vertegenwoordigd door drie papersessies en twee symposia. Uitval, zittenblijven, aansluiten op de arbeidsmarkt, allochtonisering van het speciaal onderwijs en de structurele trekken van de ongelijke kansen voor leerlingen uit verschillende sociale milieus waren de onderwerpen, die binnen de verschillende subthema's werden behandeld.

Het symposium over voortijdig schoolverlaten was wat mij betreft het meest interessant. Met name de manier waarop De Vries, die de eigenlijke discussiant verving, de verschillende presentaties ‘samenpakte’. Hij gaf precies aan waar de blikrichting naar toe moet en wat ook dit keer weer uit het gepresenteerde onderzoek is gebleken: zonder diploma's kun je ook leven. 90% van de voortijdige schoolverlaters doet daarna weer cursussen of betaalde arbeid. Cognitieve belemmeringen van leerlingen zijn zelden de oorzaak van uitval. Volgens De Vries gaat het bij uitval om een maatschappelijk geïndiceerd motivatieprobleem door de ongehoorde verlenging en veralgemenisering van het onderwijs. De nadruk op het onderwijs als oplossing voor het werkloosheidsprobleem getuigt van naïviteit, omdat al lang is gebleken uit onderzoek in andere landen dat diploma- bezit of startkwalificatie geen structurele oplossing biedt voor werkloosheid in rijke landen. Hij vroeg zich dan ook af of je met dit soort beleid niet onbedoeld de problemen veroorzaakt: Wat heeft het voor zin om aan een startkwalificatie vast te houden voor iedereen, terwijl er gewoon geen werk is voor iedereen.

De Vries had, vanwege het ingenomen perspectief, dan ook geen goed woord over voor

Van Tilborg (UU) die de school nog eens verantwoordelijk wilde stellen voor de uitvalproblematiek van leerlingen. Zij wees op de verantwoordelijkheden van de school en de mogelijkheden die er binnen de school nog liggen om de uitval tegen te gaan. De Vries wilde niets weten van deze ‘blaming the school’, dat was al te lang zo gedaan en de rek is er wel zo'n beetje uit.

Eimers (KUN) presenteerde in het kader van dit symposium een evaluatie van projecten voor voortijdig schoolverlaters. Hij gaf inderdaad het relatief grote succes aan van het toeleiden van schoolverlaters naar werk. De Wit en Dekkers (KUN) gaven aan dat het creëren van alternatieve opleidingstrajecten of aanbieden van opleidingskredieten met name voor schoolverlaters met een diploma (maar geen startkwalificatie) weinig soelaas zal bieden. Het merendeel van de dropouts is zeer gemotiveerd....voor werk.

Vlug (EUR) wees tijdens dit symposium in een verslag van haar onderzoek naar schooluitval in het VSO-MLK wel op een heel belangrijke uitzondering en tevens een verontrustend proces, namelijk op de allochtonisering van het speciaal onderwijs en de toename van de uitstroom bij dit schooltype. De Vries deelde de ongerustheid van Vlug, omdat de leerlingen waar dit onderzoek zich op betrof ook op den duur niet aan het werk komen, zeker nu de plaatsen waar ze normaal gesproken aan het werk zouden kunnen komen – sociale werkplaatsen – commercieel moeten gaan draaien. Deze omkering van de basisprincipes, waarmee het idee van de sociale werkplaats destijds is geïntroduceerd, zorgt er voor dat alleen de beste kinderen van het VSO/MLK hierin nog aan het werk kunnen komen: een heel vervelend effect.

In de papersessie over onderwijs en arbeidsmarktpositie behandelde Van Batenburg (RUG) zijn onderzoek naar het succes van MBO-leerlingen op school en in hun vervolloopbaan. Leerlingen van het MBO waren gedurende hun schoolperiode en anderhalf jaar na verlaten van de school gevolgd, met de bedoeling om inzicht te krijgen in het externe rendement van het MBO. Extern rendement is geoperationaliseerd als de waarde van de opleiding voor de arbeidsmarkt van leerlingen afkomstig van deze opleiding. Het paper richt-

te zich op de vraag "Waar zijn leerlingen een jaar na het al dan niet gediplomeerd verlaten van de opleiding?" Schoolverlaters in het eerste leerjaar bleken in meer dan de helft van de gevallen een andere dagopleiding te zijn gaan volgen. Opvallend detail in de gegevens van Van Batenburg was dat jongeren die in het derde jaar uitvielen, na een jaar vaker werk hadden gevonden dan leerlingen die vertraagd hun diploma haalden. Dit zou te maken kunnen hebben met het sterke effect van leeftijd en de (in verhouding) zwakke relatie tussen het hebben van een diploma en het niveau van werken en het salaris dat men verdient. De behoefte van werkgevers aan jeugdige (goedkope) arbeidskrachten leek de discussianten de meest waarschijnlijke reden.

De donderdagochtend ging het binnen het thema "Onderwijs & Samenleving" over de omvang en de effecten van het zittenblijven in het basis en voortgezet onderwijs in Nederland en Vlaanderen.

Prestaties, sekse en land van herkomst van de leerling bleken (niet verrassend) de beste voorspellers van het zittenblijven. Van de Werf (RUG) onderscheidde naast zittenblijvers ook vertraagde leerlingen (leerlingen die één of meer jaren onder de reguliere leeftijd van hun klas zitten). Het bleek dat vertraagde leerlingen minder leerwinst over de jaren heen halen en dat vertraging een negatief effect had op schoolbeleving.

Belangrijk resultaat van het onderzoek van Reezigt (RUG) was dat het effect dat de school wil bereiken met zittenblijven niet wordt bereikt, omdat de prestaties weliswaar iets omhoog gaan, maar de schoolbeleving minder positief wordt. Grote verschillen zijn er tussen de manieren waarop de beoordeling plaatsvindt en de procedures die de schoolgidsen beschrijven omtrent de regels van het zittenblijven. Er was enig kwalitatief onderzoek gedaan naar de redenen voor het zittenblijven volgens ouders en leerlingen. Het bleek dat de meeste leerlingen de oorzaak van de problemen bij zichzelf leggen. Ouders geven aan dat problemen thuis vaak de oorzaak waren van het zittenblijven van hun kind.

Van Damme (KUL), de enige Vlaamse bijdrage in deze, had de luxe van de beschikking over een databestand met gegevens over het zittenblijven van de hele populatie leerlingen in

het schooljaar '93-'94. In absolute percentages was het aantal zittenblijvers 7,15% in dat jaar. Het bleek dat jongens eerder blijven zitten en meisjes vaker over gaan op voorwaarde dat ze bepaalde vakken niet zullen doen in het volgende jaar. Van Damme vond een positief effect van zittenblijven voor erg zwakke leerlingen en een negatief effect voor hen die zich al dicht bij de normaal presterende groep bevonden. Van Damme vond een positieve schoolbeleving bij Vlaamse zittenblijvers in tegenstelling tot de negatieve schoolbeleving van Nederlandse zittenblijvers. Dit verschil is heel waarschijnlijk te verklaren door de wijze van vergelijken (vergelijk je iemands schoolbeleving met de eigen schoolbeleving van het jaar daarvoor, zoals in Vlaanderen was gedaan of ten opzichte van andere klasgenoten die niet zijn blijven zitten?).

Inhoudelijk was het eerste gedeelte van het laatste symposium: "Het cohort onderzoek in het primair onderwijs (PRIMA)", vooral gericht op de leerlingen in het speciaal onderwijs ten opzichte van het basisonderwijs. In PRIMA zijn voor het eerst op landelijke schaal gegevens bekend over leerlingen, leerkrachten en directies van scholen voor speciaal onderwijs die vergelijkbaar zijn met soortgelijke gegevens die in het basisonderwijs zijn verzameld. Peetsma en Roeleveld (UvA) onderzochten (onder andere) de psycho-sociale factoren – zoals schoolwelbevinden, sociaal gedrag, werkhouding etc – bij leerlingen uit het speciaal onderwijs in vergelijking met leerlingen uit het basisonderwijs. Een opvallend gegeven was dat kinderen op het MLK een hoger cognitief zelfbeeld hebben dan kinderen op het LOM en in het basisonderwijs. Voor de oorzaak van dit verschil werd verwezen naar de leerkracht/leerling interactie die Peetsma omschreef als een effect van de 'debiliserend' dat optreedt bij een overdreven positieve feedback door de leraar. Ook werd er verwezen naar de groep die de leerling als referentiekader neemt: bij MLK-kinderen zijn dat de andere geringe presteerders, terwijl LOM-schoolkinderen zich eerder vergelijken met 'gewone' basisschool-kinderen.

Het paper van Van Langen en Jungbluth (KUN) ging over de vraag wat leerkrachten doet besluiten om een leerling als hulpbehoevend (risico-leerling) te betitelen? Erg duidelijke

lijk werd uiteengezet hoe contextgevoelig het begrip probleemleerling eigenlijk is door de relatering aan de sociaal etnische achtergrond van de leerling. Het blijkt dat kansrijke leerlingen die relatief tot de laagst presterenden bij taal behoren, in absolute termen veel beter scoren dan kansarme leerlingen die in hun klas tot de zwaksten behoren. Vrijwel hetzelfde effect, zij het iets zwakker treedt op bij rekenen. Op schoolniveau keren beide effecten terug, in die zin dat op scholen met veel kansarme leerlingen de relatief laagst presterende leerlingen absoluut gezien duidelijk zwakker presteren dan de relatief zwakst scorende leerlingen op scholen met overwegende kansrijke leerlingen. Behalve naar de taal en rekenscores om risicoleerlingen op te sporen, is er ook gekeken naar de leerlingen die door hun groepsleerkrachten getypeerd werden als risicoleerling op basis van werkhouding, zelfvertrouwen van de leerling en de inschatting van het thuisklimaat door de leerkracht. Dan lijkt het er op dat kansarme leerlingen minder kans maken om bij gelijke prestatie-scores herkend te worden als potentiële risicoleerlingen. Ingewikkeld is ook dat niet-cognitieve beoordelingen die duidelijk gerelateerd zijn aan kansarme groepen (arbeiderskinderen) een belangrijke rol spelen in de perceptie van leerkrachten ten aanzien van hun leerkrachten. Het is nog niet helemaal duidelijk wat de onbedoelde gevolgen hier weer van zijn, maar het lijkt erop dat de bestempeling tot risicoleerling een sociaal privilege kan zijn.

De onderzoeken op het terrein van PRIMA hebben in het tweede deel van dit symposium te maken met een structureel onderscheid tussen kansarme en kansrijke kinderen.

Een van de meest zorgwekkende conclusies, die getrokken kon worden uit het tweede paper dat Jungbluth presenteerde op dit symposium en wat ook de aandacht van de pers trok doordat de term standenonderwijs weer eens uit de kast was gehaald was dat kansrijke leerlingen op een school met veel kansarme leerlingen duidelijk slechter scoren op taal en rekentoetsen dan kansrijke kinderen op een school met veel andere kansrijke kinderen. Bovendien was zeer opvallend dat Turkse en Marokkaanse kinderen het nauwelijks beter doen als het gaat om taal en rekentoetsen op een kansrijke school. Dat zijn natuurlijk moeilijk te verteren resultaten. Zeker als er geen aandacht wordt gegeven

aan de mogelijkheden van eventuele positieve effecten van de mengeling van kansarme en kansrijke kinderen op scholen. Je zou hierbij kunnen denken aan sociale vaardigheden en bewustwordingsprocessen.

Ledoux en Overmaat (UvA) hadden onderzoek gedaan naar de verschillen in kwaliteitsvoorwaarden en kwaliteitszorg tussen scholen met meer en minder kansrijke leerlingen. In dit opzicht lijken scholen met veel kansarme kinderen zich eerder positief dan negatief te onderscheiden van scholen met veel kansrijke kinderen (hierbij is niet het directe handelen van de docent in de analyse betrokken). Scholen met meer kansarme kinderen richten zich vaker op niveaubewaking, planmatig handelen en evalueren. Leerkrachten van elitescholen hebben echter (uiteraard) weer meer vertrouwen in hun eigen kunnen en dat van de leerlingen. Dit laatste doet – ondanks het op zich positieve resultaat op het terrein van kwaliteitszorg en voorwaarden – toch ongewild denken aan de structurele trekken van kansongelijkheid en de mate waarin die trekken in de praktijk van het onderwijs zijn ingebed.

Lerarenopleiding & Leraarsgedrag

(A. Zanting, ICLON, RUL, Leiden)

Tijdens de ORD'96 concentreerden de bijdragen binnen het thema "Lerarenopleiding & Leraarsgedrag" zich, wat betreft nieuwe ontwikkelingen, op een tweetal gebieden. In de eerste plaats had een groot deel van de papers betrekking op de professionele ontwikkeling van leraren. In de tweede plaats was er ruime aandacht voor de veranderende rol van leraren onder invloed van het streven naar zelfstandig leren van leerlingen. De onderstaande impressie beperkt zich tot de presentaties die betrekking hadden op de twee genoemde domeinen.

In het eerste subthema, "Professionalisering van Leraren", kwam een breed scala van onderling gerelateerde onderwerpen aan de orde, namelijk de opleiding, nascholing en beoordeling van leraren. Huberman startte met een presentatie over netwerken van docenten. Door diverse ontwikkelingen binnen het onderwijs, zoals de grotere autonomie van scholen en leraren, de toenemende aandacht voor het beoordelen van leraren en de nadruk op collegialiteit en

samenwerking, is er bij docenten behoefte ontstaan om te overleggen en ervaringen uit te wisselen. In de presentatie werd duidelijk dat docent-netwerken hiervoor mogelijkheden bieden en tegelijkertijd een remedie vormen voor de gebruikelijke knelpunten bij het invoeren van onderwijsvernieuwingen. Binnen netwerken zijn docenten namelijk collectief betrokken. Bovendien zijn netwerken ingebed in de schoolcontext. Hoe kan nu aan deze netwerken vorm worden gegeven? Huberman baseerde zijn *craft*-model op de wijze waarop ervaren leraren te werk gaan bij het oplossen van problemen. Belangrijke elementen in dit netwerkmodel zijn discussiëren, experimenteren en het delen van ervaringen.

Het onderzoek van Galesloot (UU) naar netwerken van scholen in het kader van de Tweede Fase sloot uitstekend aan op Huberman. Galesloot vatte netwerken op als een mogelijkheid voor voortdurende nascholing van docenten. Een belangrijke voorwaarde om dit te doen slagen is volgens hem collegiaal of horizontaal leren, waarbij docenten op grond van hun expertise zelf inhoud geven aan netwerk-bijeenkomsten.

Een interessante parallel met de netwerken werd zichtbaar in de eigentijdse wijze waarop intervisie binnen de lerarenopleiding gestalte had gekregen in een onderzoek dat gepresenteerd werd door Admiraal en Lockhorst (UU). Zij onderzochten welke bijdrage een discussielijst voor docenten-in-opleiding, toegankelijk via e-mail, kon leveren aan de ontwikkeling van reflectievaardigheden van docenten-in-opleiding. De zo ontstane discussiegroep, waarbij docenten-in-opleiding hun stage-ervaringen beschreven en op elkaar reageerden, vertoonde door onderlinge discussie en uitwisseling van ervaringen verrassende overeenkomsten met de eerder beschreven netwerken. Het onderzoek dat zich weliswaar nog in een beginstadium bevindt, lijkt zowel vanuit het oogpunt van onderzoek als opleiding veelbelovend. Een belangrijk aandachtspunt bij dit onderzoek is de verdere ontwikkeling van een instrument om reflectievaardigheden te meten en hieraan een waardering toe te kennen.

Binnen het symposium "Beoordeling van leraren" werd verder ingegaan op één van de ontwikkelingen die door Huberman was genoemd. Door de instroom van leraren die

geen reguliere lerarenopleiding hebben gevolgd, bestaat er een groeiende belangstelling voor het beoordelen van leraren en de ontwikkeling van een hiervoor benodigde beoordelingssystematiek. In het symposium werd het beoordelen van leraren tevens opgevat als een middel tot professionele groei, omdat beoordeling leraren een concreet beeld oplevert van hun eigen functioneren.

Beijaard (RUL) benadrukte dat de ontwikkeling van een beoordelingsstelsel niet los kan worden gezien van het onderzoek naar praktijkkennis van docenten. Door deze praktijkkennis te betrekken bij de beoordeling van docenten wordt de professionaliteit van docenten erkend. Bovendien doet onderzoek naar praktijkkennis recht aan de complexiteit van het onderwijzen. Verder pleitte Beijaard voor authentieke beoordelingsmethoden en het gebruik van meerdere, met elkaar samenhangende instrumenten.

Dit laatste punt bleef niet beperkt tot woorden, maar was toegepast in het onderzoek van Klarus (CIBB), gericht op de ontwikkeling van een beoordelingsstelsel voor leraren binnen het Secundair Beroepsonderwijs, en in het onderzoek van Uhlenbeck (RUL) betreffende de ontwikkeling van een voorbeeld-instrumentarium om de bekwaamheden van beginnende docenten Engels vast te stellen. In het laatst genoemde onderzoek werd, naast theoretische over leren en onderwijzen en onderzoek naar de verwerking van een vreemde taal, expliciet de praktijkkennis van ervaren docenten betrokken. Van Tartwijk (UU) noemde, bij het gebruik van portfolio's binnen de lerarenopleiding, naast beoordelings- ook nadrukkelijk begeleidingsdoelen. De discussie die op de bijdragen volgde, richtte zich voornamelijk op de wezenlijke vraag of beoordeling van leraren legitiem is en of dit een taak van onderwijskundigen is. Duidelijk bleek dat een discussie hierover nog niet afgesloten was voor de deelnemers. Helaas bleef er daardoor ook natwelijks tijd over om dieper in te gaan op de in ontwikkeling zijnde beoordelingsinstrumenten.

In het symposium over activerende instructie en de papersessie "De docent en de zelfstandig lerende leerling" stond het tweede subthema, de nieuwe invulling van de docentrol bij zelfstandig leren, centraal. De in de papersessie gepresenteerde onderzoeken van Rademaker

(CIBB), binnen de context van het beroepsonderwijs, en van Stallen en Lunenberg (VU), binnen de context van de basiseducatie, bouwden voort op het leerstijlen-onderzoek van Vermunt. Een sterk punt in deze onderzoeken was de vertaling van leer- naar doceerstijlen. Het is namelijk nog niet duidelijk welke opvattingen en feitelijke activiteiten van leraren het zelfstandig leren van leerlingen kunnen bevorderen. De instrumenten die in beide onderzoeken gehanteerd werden, bevinden zich weliswaar nog in een stadium van verdere ontwikkeling, maar bieden zeker perspectieven om de doceerstijl voor procesgerichte instructie te concretiseren.

Een interessant resultaat uit het ontwikkelingsonderzoek van Stallen en Lunenberg was dat docenten hun doceerstijl afstemmen op de verwachtingen van cursisten. Hierdoor kunnen docenten, die procesgerichte instructie nastreven, terugvallen op een substituerende rol wanneer cursisten een centrale rol van hen verwachten in het leerproces. Geconcludeerd kan worden dat niet alleen de intenties van docenten, maar ook de verwachtingen van lerenden een doorslaggevende rol spelen bij het al dan niet slagen van onderwijsvernieuwingen.

In de discussie van de papersessie werd met recht gewaarschuwd voor het gebruik van het concept 'zelfstandig leren' als containerbegrip in zowel het onderzoek als het onderwijs. Zo worden de termen 'zelfstandig leren' en 'actief leren' veelal onterecht als equivalent gebruikt, wordt 'zelfstandig leren' verward met *zelfstandig werken* en mag ook de interpretatie van 'zelfstandig leren' door docenten zelf niet over het hoofd worden gezien.

De actieve rol van de lerende werd tijdens deze ORD sterk benadrukt. Eigenaardig is wel dat het aanbod van de ORD bijna geheel bestond uit frontale papersessies en symposia waarbij, met uitzondering van de discussies, weinig beroep gedaan werd op een actieve inbreng van de deelnemers. Bij de ORD '97 zouden een groter aantal interactieve presentatievormen een welkome aanvulling zijn.

Beroeps- & Bedrijfsopleidingen

(E.A.M. van der Schoot, TO, Universiteit Twente, Enschede)

Binnen het thema "Beroeps- & bedrijfsopleidingen" was door de coördinatoren Van der Sanden (KUB) en Versloot (UU) een breed scala aan onderwerpen samengebracht. Er stonden twee symposia, een forumdiscussie, twee papersessies en een postersessie op het programma, deze laatste bleek echter toch een papersessie te zijn. Het door mij bezochte symposium had als thema: "Onderwijskundige en beleidsmaatregelen rondom rendement in het BVE-veld". De forumdiscussie ging over performance improvement en (nieuwe) rollen van beroeps- en bedrijfsopleiders. De papersessies betroffen: Effectief opleiden; Leren voor een beroep, en Beroepsonderwijs: kennisinfrastructuur en rendement.

In de papersessie getiteld "Effectief opleiden" bracht De Vries (RUG) factoren naar voren die een training in het 'omgaan met lastige/agressieve klanten' succesvol maken. Centraal stond daarbij het oordeel over de effecten van de trainingen op de cursisten. Het niet serieus genomen worden van de werkvloer bleek hierbij het grootste knelpunt te zijn.

Vervolgens presenteerden Streumer en Van der Klink (UT) hun bevindingen over HRD (Human Resource Development) in de toekomst. Zij hadden een survey onder opleidingsfunctionarissen gehouden waarin zij hun 40 ontwikkelingen voorlegden op organisatorisch, technologisch en opleidingskundig gebied. De respondenten was gevraagd bij ieder item aan te geven: 1) de kans dat de ontwikkeling de komende vijf jaar zal optreden; 2) het effect ervan op HRD in het algemeen; 3) het effect ervan op de eigen organisatie. Opvallend was dat men de kans dat een bepaalde ontwikkeling invloed zou hebben op HRD aanzienlijk hoger inschatte dan de kans dat deze ontwikkeling invloed zou hebben op de eigen organisatie.

Mulder en Van Ginkel (UT) deelden een hand-out uit van hun onderzoek naar de determinanten van opleidingseffectiviteit. Zij onderzochten de mogelijke samenhang tussen de specificatie van een opleidingsproject enerzijds en de getrouwe implementatie daarvan en de effectiviteit van dat programma anderzijds.

Hiertoe was een survey gehouden onder opdrachtgevers van opleidingsprojecten. Het conceptuele kader was gebaseerd op procesmodellen voor het ontwerpen en implementeren van opleidingen en modellen voor het onderscheiden van niveaus in opleidingseffecten. Getrouwe implementatie bleek voor de effectiviteit van groot belang te zijn. Het was de bedoeling dat er tijdens deze presentatie ruimte zou zijn voor discussie, wat helaas in het water viel.

Tot slot brachten Van Ginkel, Streumer en Nijhof (UT) hun studie naar nieuwe managementrollen in de communicatie-industrie onder ieders aandacht. Door gebruik te maken van verscheidene onderzoeksinstrumenten hebben zij getracht de veranderende rol van managers in kaart te brengen om een nieuwe opleidingsstructuur te ontwerpen en toekomstgerichte invulling te kunnen geven aan het curriculum.

De forumdiscussie met als thema "Performance improvement en (nieuwe) rollen van beroeps- en bedrijfsopleiders" sloot de eerste dag van de ORD'96 af. Forumleden waren: Van der Zee (RUL), Van Noort (NS), Bruining (Center Parcs) en Streumer (UT). Zij discussieerden samen en met de zaal onder leiding van Mulder (UT). Performance improvement, oude wijn in nieuwe zakken, aldus Bruining (Center Parcs). Maatschappelijke ontwikkelingen nopen bedrijven tot het continu verbeteren en innoveren van werkprocessen. Een soort modern Taylorisme. Hij zag in arbeidsorganisaties een driedeling ontstaan: een elite van kenniswerkers, een laag van vaklieden, en een onderlaag van flexwerkers. Vervolgens ontstond een interessante discussie omtrent het nut van het optimaal laten functioneren van werknemers. Met name de ontmenselijking van veel ondernemingen riep bij velen ethische vragen op. Echter telkens keerde in de gevoerde redeneringen de vraag terug naar de inhoud van het concept performance improvement. Van der Zee (RUL) wees erop dat in de vakliteratuur steeds meer performance-terminologie tot het 'correcte' taalgebruik hoort. Een copernicaanse wending in het opleidersmétier, waarbij hij zich afvroeg of we daar blij mee moesten zijn.

De tweede dag begon met een symposium getiteld "Onderwijskundige en beleidsmaatregelen rondom rendement in het BVE-veld". Hövels (KUN) deed verslag van zijn analyse

van het overheidsbeleid ten aanzien van het secundair beroepsopleidingsonderwijs. Een schets werd gegeven van de belangrijkste probleemdefinities en beleidslijnen. Beleidsdoelen en instrumentaria werden eveneens onder de loep genomen. Afgesloten werd met dilemma's en uitdagingen voor het huidige beleid. Samenwerking en integratie bleken in dit kader van belang, evenals een grotere autonomie van instellingen en het opvoeren van het beleidsvoerende vermogen van schoolorganisaties.

Vervolgens presenteerden Den Boer en Harms (RUG) hun studie naar causale relatie(s) tussen beleidsimplementatie en rendement. In deze bijdrage stond het gekozen instrumentarium van de overheid centraal. Uit hun onderzoek bleek dat veel afhangt van wat instellingen er zelf van maken.

Lokman, Nieuwenhuis en Van Woerkom (STOAS) bespraken hun onderzoek, gericht op het meten van het interne rendement in het BVE-veld. Met name de individuele benadering van het rendementsvraagstuk werd uitgewerkt, hetgeen stond voor dat wat het onderwijs toevoegt aan reeds bestaande kwalificaties. Traditionele instrumenten kunnen hiervoor niet meer gebruikt worden.

Tot slot gingen Den Boer (KUN) en Moerkamp (UvA) in hun presentatie in op de betekenis van een onderwijsnummer voor evaluatie van het BVE-veld. Inderdaad leidt invoering van een onderwijsnummer tot het beter kunnen volgen van leerlingen, daling van het oneigenlijk gebruik van gelden, en ontlasting van scholen. Echter gebleken is dat de beleidsinformatie-waarde beperkt is. Tijdens dit symposium was er wel erg weinig ruimte voor interactie met de zaal.

In de daarop volgende papersessie met als thema "Leren voor een beroep" viel één presentatie weg, die van Vermeulen (UU). Slaats, Van der Sanden en Lodewijks (KUB) presenteerden hun studie naar leerstijlen en persoonlijkheidskenmerken in het MBO. Zij hadden leerstijlen gemeten volgens de ILS-MBO en volgens Kolb. ILS-MBO is een uitsprakenlijst waarop leerlingen hun opvattingen omtrent leren en hun motieven om te leren scoren. Iedere leerling krijgt een score op een reproductieve en op een constructieve factor, waarna de volgende vier leerstijlen onderscheiden kunnen worden: veelzijdige-, constructieve-, reproduc-

tieve- en de passieve leerstijl. Volgens Kolb kunnen ook vier leerstijlen onderscheiden worden: accommodatieve-, assimilatieve-, divergente- en de convergente leerstijl. Een deel van de respondenten scoorde op persoonlijkheidskenmerken van de Big Five. Van het onderzoek bleek dat op het MBO de ILS-MBO meer bruikbaar is dan Kolb. Ook bleken er relaties te bestaan tussen persoonlijkheidskenmerken en de verschillende leerstijlen.

Tot slot presenteerde Kools (LUW) een voorstudie naar vrouwelijke en mannelijke leerlingen over de kenmerken van hun toekomstig agrarische beroep. Het lijkt dat verschillende beroepen verschillende aantrekkingskrachten uitoefenen. Opvallend is ook dat leerlingen over hetzelfde beroep verschillende kenmerken benadrukken.

Het thema "Beroepsonderwijs & Bedrijfsopleidingen" werd afgesloten met een papersessie die als onderwerp had "Beroepsonderwijs: Kennisinfrastructuur en rendement". Twee presentaties volgden, die van Nieuwenhuis en Grooters (STOAS) over de bouwstenen voor responsiviteit en die van Bruijn (UvA) over de vormgeving en het rendement van de korte voltijd beroepsopleiding. Door de hitte was het aantal toehoorders danig afgeslankt, ondanks de interessante beschouwingen.

Hoger Onderwijs & Volwasseneneducatie (C.J. Sikkema, SCO-Kohnstamm Instituut, UvA, Amsterdam)

De divisie "Hoger Onderwijs" had vier papersessies en vier symposia op haar agenda staan. Twee sessies en twee symposia worden hier besproken, gevolgd door een korte algemene conclusie.

In de eerste papersessie stond het thema "Leerstijlen en leerstrategieën" centraal. In deze levendige sessie werden vier papers gepresenteerd. Na elke presentatie was er ruimte voor een aantal vragen en aan het eind gaf discussiant Vermunt (KUB) interessante opmerkingen waarover nog een tijd is gepraat.

In drie van de vier presentaties (Vermetten, KUB; Prins, UvA; Wijffels, Universiteit van Antwerpen) werd het studeergedrag van studenten, onder andere in relatie met leerresultaten benadrukt. Over het algemeen werd in de longitudinale cohortstudies gebruik gemaakt

van het instrument Inventaris Leerstijlen (ILS) van Vermunt. De uitkomsten waren echter verschillend. Vermetten vond dat studenten betekenisgerichter waren gaan werken terwijl Prins nauwelijks verschillen had gevonden. Daarnaast vond Prins geen cross-sectionele verschillen, in tegenstelling tot Wijffels die vond dat jongerejaars-studenten inzichtelijker studeerden, met meer structuur en minder nadruk op memoriseren.

Een discussiepunt dat hierbij aan de orde kwam betrof de aanpassing van studenten aan wat van hen verwacht wordt. De vraag was of studenten hun studietaken weten en de strategieën kunnen kiezen om aan de studie-eisen te voldoen. Doen studenten gewoon wat van hen verwacht wordt en passen ze zich aan als een kameleon? Het vergt bepaalde kennis om strategisch en flexibel in te kunnen springen op datgene wat van ze verwacht wordt. In de beroepspraktijk moeten die vaardigheden ook aanwezig zijn. Het is van belang meer inzicht te verkrijgen in de onderlinge samenhang van de verschillende leerstrategieën. Wijffels heeft daartoe een eerste stap gezet door factoranalyse te gebruiken. Een andere mogelijkheid om meer zicht op dit vraagstuk te verkrijgen werd door Oosterhuis (UT) genoemd. Zij benadrukte dat universitaire opleidingen te theoretisch zijn en te weinig op de arbeidsmarkt zijn afgestemd. In Twente is een cursus ontwikkeld waarin professionele vaardigheden aangeleerd worden. Daarbij is ook de ILS gebruikt, maar op een andere manier. Tijdens deze cursus worden de uitkomsten teruggekoppeld naar de aparte items. Een sterkte-zwakte analyse kan aangeven of een student bijvoorbeeld laag scoort op kritisch verwerken of memoriseren. Daar kan dan aan gewerkt worden. Hier wordt het begrip leerstijl ontleed in de aparte onderdelen.

De relatie tussen studiegedrag en leerresultaten is niet helder. Er is wel een tendens dat studenten die de toepassingsgerichte leerstijl gebruiken betere resultaten behalen, maar dit verband is niet significant. De niet gevonden relatie kan veroorzaakt zijn door de kwaliteit van het onderwijs: leerprestaties kunnen fout beoordeeld worden.

Tot slot werd ingegaan op het feit dat de ILS niet ontwikkeld is om longitudinale ontwikkeling van leergedrag te kunnen verklaren. Het is

daarom van belang om na te denken over de mate waarin schalen geschikt zijn en om meerdere meetprocedures toe te passen.

De papersessie "Studie- en beroepsloopbanen" kwam minder consistent over dan de eerste sessie, wat mede veroorzaakt zal zijn door het uitvallen van twee presentaties. De eerste spreker (Mulder, OU) vertelde over het allochtonenbeleid van de OU. Sinds kort is er een onderzoek gestart om de doorstroming van allochtonen in kaart te brengen, waarop vervolgens beleid kan worden afgestemd. De doorstroming van allochtonen wijkt niet af van de gehele OU-populatie, maar ze studeren wel veel meer uren om punten te halen. In tegenstelling tot de verwachting bleek uit enquêtes de factor taal niet problematisch te zijn. De discussie ging over de interpretatie van deze uitkomsten. Waarschijnlijk speelt de zelfselectie aan de poort een belangrijke rol. Nadere analyses zullen misschien meer helderheid geven over de interpretaties.

Het tweede paper, gepresenteerd door Cohen (RUG), ging over de invoering van een nieuw curriculum bij Geneeskunde in Groningen. Dit curriculum is probleemgeoriënteerd, biedt meer diversiteit in onderwijsvormen, legt de nadruk op kennisverwerving in plaats van op kennisoverdracht en legt meer nadruk op zelfstudie. Zowel het oude als het nieuwe curriculum is geëvalueerd en de uitkomsten zijn gunstig: het oordeel van studenten over de inhoud en vorm van het curriculum is van negatief omgezet in positief. Ook de docenten zijn positief: ze zijn aangenaam verrast over de kennis en kunde van de studenten. Het studeergedrag is veranderd, de studenten bezoeken de colleges meer, vertonen minder uitstelgedrag en verdelen de studiebelasting beter. Daarentegen doen studenten minder aan zelfstudie terwijl dat toch een uitgangspunt van de verandering was.

Verbeek (UvA) vertelde over zijn onderzoek naar verschillen tussen opleidingen (pedagogiek, onderwijskunde, biologie en medische biologie) en cohorten op loopbaanvariabelen. Een hele reeks loopbaanvariabelen werd gemeten. Er werden geen cohortverschillen gevonden, wel werden verschillen tussen de opleidingen geconstateerd.

Variabelen die de verschillen enigszins verklaren zijn: de duur van de studie (hoe langer,

hoe negatiever voor de loopbaan) en activiteiten met betrekking tot de loopbaan tijdens de studie (positief effect).

In het symposium "Hoger Activerend Onderwijs", georganiseerd door het IVLOS, stond het activeren van hogere cognitieve processen centraal. Deze gedachte werd erg letterlijk toegepast want van de 'toeschouwers' werd een actieve bijdrage verwacht. Na een inleiding van Rensman werd de groep in tweeën gesplitst. Ik maakte deel uit van de groep waarin het curriculum centraal stond. Twee papers werden gepresenteerd. Beide gingen over het invoeren van een opdrachtsysteem en de gevolgen van het (vrijwillig) deelnemen aan de opdrachten voor het studieverloop. Wagenaar paste dit systeem toe bij Scheikunde en vond onder de participanten een grotere onderwijs- en tentamendeelname, hogere slaagpercentages, hogere eindcijfers en geen verschil in zelfstudie. Gritter deed verslag van een experiment bij Economisch Publiekrecht, gericht op het verwerken van de studiestof. Haar resultaten (zonder controle vooraf) lieten een gelijksoortig beeld zien: hogere cijfers en slaagpercentages, een hogere subjectieve opbrengst (meer inzicht in het vak), maar meer tijd voor zelfstudie.

Het publiek kreeg als opdracht te kiezen welk opdrachtsysteem de voorkeur had van de docent en student en welke eventuele andere vorm gekozen zou kunnen worden. De discussie barstte los. Het eerste punt ging over het prikkelen van studenten: wanneer er meer externe prikkels zijn wordt de student meer geactiveerd en werkt hij/zij misschien harder. Maar met die prikkels wordt de student ook beperkt en treedt de docent schoolser op, waar die docent geen zin in heeft. Ten tweede zal de huidige calculerende student alleen participeren in het experiment wanneer duidelijk is wat de beloning voor deelname zal zijn.

Na de discussie werden de twee groepen weer samengevoegd en trad de discussiant Pilot op. Zijn voornaamste punt van kritiek was dat de gepresenteerde onderzoeken nog te veel op los zand berusten. De uitkomsten zijn verschillend en inconsistent. Daarom is het van belang meer naar de literatuur te kijken en vanuit de verschillende kleine studies te komen tot een theorie.

Het tweede symposium dat eveneens door

het IVLOS werd georganiseerd, had als thema "Onderzoek naar de aansluitingsproblematiek VWO-WO gericht op de propaedeuse". Drie onderzoeken werden gepresenteerd.

Admiraal keek naar de invloed van studentenfactoren op cognitieve en affectieve opbrengst bij rechtenstudenten na 1 jaar. Hij vond kleine effecten, wat mede veroorzaakt kan zijn door de selectie die had plaatsgevonden bij de tweede vragenlijst. Wientjes presenteerde een zeer interessant explorerend en kwalitatief onderzoek naar de aansluitingsproblematiek op de UU, dat in opdracht van het College van Bestuur was verricht. In het onderzoek werd vanuit het perspectief van de studenten gekeken naar de verwachtingen die aan studenten worden gesteld en naar hun werkelijke capaciteiten. Door observaties bij zowel werk- als hoorcolleges bij verschillende faculteiten en het afnemen van een vragenlijst bij een aantal geslaagde en niet-geslaagde studenten verkregen de onderzoekers een beeld van de interactie tussen student en studie-omgeving. Een belangrijke conclusie was de inconsistente leeromgeving: onderwijsdoelen worden gedefinieerd, maar de leerstof leidt niet tot die doelen. In de praktijk blijken studenten op tentamens de leerstof te moeten reproduceren. Studenten zijn tot veel meer in staat, maar passen zich aan de verwachtingen, waardoor hun intellectuele ontwikkeling wordt belemmerd.

De derde spreker (Prins) vertelde over een onderzoek gehouden bij de studie Farmacie, dat is opgezet om de knelpunten in de aansluiting helder te krijgen. Uit zijn onderzoek bleek dat niet alleen naar de interactie tussen student en onderwijs-leeromgeving gekeken dient te worden, maar ook naar de interactie tussen zowel studentenfactoren als onderwijsfactoren op zich. Zo zeggen studenten bijvoorbeeld gemotiveerd te zijn terwijl ze weinig aan zelfstudie doen. Ook blijkt dat de leerstof waarmee op het tentamen de meeste punten verdiend kunnen worden, in de colleges de minste aandacht krijgt. Het is daarom belangrijk onderzoek te verrichten naar inconsistenties, die de interactie tussen student en leeromgeving verstoren en er toe leiden dat de student niet (op tijd) leert wat hij moet leren.

In de discussie, geleid door Pilot, kwam weer naar voren dat er een te gefragmenteerd beeld bestaat. Ieder onderzoek is net iets anders

opgebouwd en leidt tot andere conclusies. Er moet meer aan een theoretisch kader gewerkt worden, waardoor een beter beeld verkregen kan worden van het samenhangend systeem waarin de processen verlopen. Het onderzoek dat door Wientjes gepresenteerd is, is een goede stap in de richting. Een andere mogelijkheid is een grootschalig onderzoek naar de aansluitingsproblematiek, waarin verschillende faculteiten en universiteiten participeren.

De bijeenkomsten waren allemaal druk bezet en er was tijd genoeg voor discussie. Opvallend was dat de onderzoeken en uitkomsten een wat onsamenhangend beeld lieten zien. Iedereen is met aparte opdrachtonderzoeken bezig, terwijl een grote lijn ontbreekt. Bij beide symposia werd dit probleem aan de kaak gesteld. Het is goed om tijdens de ORD zicht te krijgen op de vele studies die plaatsvinden en te ontdekken hoe de trend verder ontwikkeld moet worden. De taak die er nu ligt is om, naast de aparte opdrachten en problemen, een overkoepelende theorie te ontwikkelen van waaruit verder onderzoek verricht kan worden.

Beleid & Organisatie van Onderwijs

(R. Maslowski, TO, Universiteit Twente, Enschede)

Rond het thema "Beleid & Organisatie in het Onderwijs" had de organisatiecommissie dit jaar vier symposia en drie papersessies geprogrammeerd, waarin uiteenlopende onderwerpen aan bod kwamen. De symposia waren gegroepeerd rond de onderwerpen onderwijs-economie; -structuur, besluitvorming en kwaliteit van schoolorganisaties; de voortgang van de basisvorming; en toekomstig onderzoek naar schoolleiderschap. In de papersessies werd verslag gedaan van onderzoek op het terrein van personeelsbeleid, autonomie en sturing in het onderwijs, en de realisatie van Weer Samen Naar School.

De papersessie "Autonomie en sturing in het onderwijs" stond in het teken van het onderwijsbeleid van deregulering en enveloppefinanciering, zoals de lump-sum financiering van scholen in België enigszins plastisch wordt omschreven. Leysen en Van Petegem (Universiteit Antwerpen) verrichtten onderzoek naar het maatschappelijk draagvlak in

België voor een vergaande autonomie van scholen. Door de Belgen werd het door scholen in dienst kunnen nemen en bevorderen van leerkrachten, het zelfstandig vaststellen van de wijze waarop vakken onderwezen worden, en de besteding van het schoolbudget, als het meest belangrijk beschouwd. Het onderzoek van Majoor (UvA) sloot nauw op het onderzoek van Leysen en Van Petegem aan. Majoor onderzocht welke kenmerken van de organisatiestructuur samenhangen met de mate waarin scholen in Nederland gebruik maken van de mogelijkheden die hen op het terrein van de bekostiging worden geboden. Het bleek dat scholen waarin schooltaken door het team als geheel worden uitgevoerd het meest gebruik maken van de financiële mogelijkheden die door het huidige bekostigingsstelsel worden geboden. Voor het overige bleken er nauwelijks verbanden tussen de structuur en de benutting van bekostigingsmogelijkheden van een school te bestaan.

Autonomie en sturing van scholen nam ook in het symposium "Schoolorganisaties: structuur, besluitvorming en kwaliteit" een belangrijke plaats in. Karsten (UvA) ging in op de inrichtingsvrijheid van schoolbestuur en schoolleiding. Hij concludeerde dat de verschillende richtingen in verschillende mate van de hun geboden vrijheid gebruik maken. Openbare scholen worden gekenmerkt door inspraak van de verschillende betrokken geledingen bij de besluitvorming, algemeen bijzondere scholen profileren zich door een sterke nadruk op het belang van de onderwijskundige-pedagogische opvatting, terwijl het rooms-katholiek onderwijs zich op geen van deze dimensies noemenswaardig onderscheidt. Verwatering tussen de denominaties lijkt op te treden, zo concludeerde Karsten (UvA), doordat scholen zich steeds sterker profileren, maar feitelijk is van verwatering slechts in beperkte mate sprake. Scholen worden tot een sterkere profilering gedwongen, aldus Teelken (UvA), door de veranderende consumentenvraag. De kritisch wordende ouder kiest bewust een school, met als keuze-argumenten niet alleen afstand en denominatie, maar vooral kwaliteit.

Karstanje (UvA) ging na in hoeverre de grotere autonomie van scholen ten aanzien van het besteden van het nascholingsbudget consequenties heeft voor hun bestedingspatroon. Uit

het onderzoek bleek dat basisscholen redelijk gebruik maken van het aanbod van andere nascholingsinstellingen. Basisscholen zijn van mening dat er door de nascholingsinstellingen beter rekening wordt gehouden met hun wensen. De tevredenheid over de nascholing is echter niet toegenomen. In het voortgezet onderwijs vindt meer formeel overleg over nascholing plaats dan in het basisonderwijs, en wordt een groot aantal plannen voor nascholing gemaakt, maar van coördinatie van nascholingsactiviteiten is vooralsnog nauwelijks sprake. Er wordt, met andere woorden, meer gepraat, maar minder gedaan. Mogelijk dat de behoefte aan nascholing, na de invoering van de Wet op het Basisonderwijs, in basisscholen reeds nadrukkelijk aanwezig is, waar de invoering van de Basisvorming en de Tweede Fase in het voortgezet onderwijs nog niet vertaald zijn in concrete nascholingsactiviteiten.

In het symposium rond het "Toekomstig onderzoek naar schoolleiderschap" deed Witziers (UT) verslag van een meta-analyse waarin de directe invloed van de schoolleider op de leerresultaten werd onderzocht. Hij vond slechts in een zeer beperkt aantal gevallen een significant verband tussen beide variabelen. Witziers opperde dat het onderzoek naar schoolleiderschap zich mogelijk beter kan richten op de leiderstaken die door leden van de school worden uitgevoerd dan op de persoon van de schoolleider. Imants (RUL) kwam tot een soortgelijke conclusie. Hij plaatste vraagtekens bij de strikte koppeling in veel onderzoeken tussen de functie leidinggeven aan onderwijs en de positie van de schoolleider. Tevens stelde Imants het gangbare onderscheid tussen algemene managementtaken en onderwijskundige leiderschapstaken ter discussie. Peters (KUN) stelde zich hiertegen teweer. Zij gaf aan dat schoolleiders en anderen die leiderstaken binnen scholen vervullen zelf onderscheid maken naar taken die dichtbij en taken die ver van het onderwijsproces af staan.

Krüger (UvA) presenteerde een op onderzoek van het Far West Laboratory in San Francisco gebaseerd model van schoolleiderschap. Centraal in dit model staan de handelingen van de schoolleider, die worden bepaald door achtergrondfactoren als de omgeving van de school, de visie en ervaring van de schoolleider, en de wettelijke kaders en regelgeving. Het

handelen van de schoolleider beïnvloedt de schoolorganisatie, de leerresultaten, en het pedagogische onderwijskundig klimaat van de school. Krüger gaf aan dat in toekomstig onderzoek met name aandacht gegeven moet worden aan strategieën van schoolleiders. Evenals Imants wees zij op het belang van onderzoek naar mentale modellen van schoolleiders. Scheerens (UT), discussiant van het symposium, twijfelde aan het nut van dergelijk onderzoek. Van belang is met name, zo gaf Scheerens aan, wat schoolleiders doen, en in hoeverre dit handelen al dan niet direct de leerresultaten beïnvloedt. Hij pleitte er voor het Far West Laboratory model te vereenvoudigen, wanneer het er om gaat onderzoek te verrichten naar schoolleiderschap en effectiviteit. Wasink (LUW) stelde de vraag aan de orde of het, gezien de bevindingen van Witziers, wel zinvol is nader onderzoek te doen naar schoolleiderschap in ons land. Imants betoogde van wel, omdat schoolleiderschap ook van invloed is op de wijze waarop scholen met vernieuwingen omgaan, en in staat zijn zich aan veranderingen in de omgeving aan te passen. Scheerens onderschreef dit, maar gaf aan dat daarbij wel voortdurend de relatie met leerresultaten gelegd moet worden. Wanneer een schoolleider goed inspeelt op vernieuwingen of veel geld voor zijn school weet binnen te halen, maar dit uiteindelijk niet leidt tot betere resultaten, kunnen terecht vraagtekens bij het onderzoek naar schoolleiderschap geplaatst worden.

Krachtige Leeromgeving/Leren & Instructie
(H. Agerbos, TO, Universiteit Twente, Enschede)

Leren & Instructie

Op de woensdagmiddag vond een symposium over leeractiviteiten binnen het thema "Leren & Instructie" plaats. Dit symposium werd voorgezeten door Van Hout-Wolters (UvA), die deze taak bewonderenswaardig vervulde. Zij leidde het symposium in met een uiteenzetting over het onderzoeksprogramma van het Instituut voor de Lerarenopleiding (ILO) van de UvA, waarvan alle presentatoren van dit symposium afkomstig waren. De centrale vraag van het symposium was: Wat is vakspecifiek en vakoverstijgend betreffende leerstrategieën en leeractiviteiten binnen het voortgezet onderwijs? Elshout-Mohr en Volman

openden de serie presentaties met een vergelijking van de perceptie van leerlingen en leraren van leeractiviteiten bij alfa- en gamma-vakken. Verschil tussen leerling- en leraar-percepties was er nauwelijks, wel was er een verschil tussen de vakken: bij de gamma-vakken werden meer activiteiten ontplooid, waarbij de docenten vooral meer informatie moesten (voor-) bewerken, en de leerlingen vooral hoog scoorden op regulatie-activiteiten.

De tweede lezing (Ten Dam en Volman) ging in op de leeractiviteiten bij twee nieuwe vakken in de Basisvorming: Verzorging en Techniek. Hoewel deze vakken een beroep (moeten) doen op zowel psychomotorische als affectieve vaardigheden (attitude-verandering), komen deze twee factoren nauwelijks in de kerndoelen voor deze vakken voor. Opmerkelijk, gezien de maatschappelijke discussie die bij de invoering van deze vakken gevoerd werd.

Na de pauze ging Severiens in op verschillen in leerstijlen tussen seksen en tussen vakken (Nederlands, geschiedenis, wiskunde, en biologie). De resultaten lieten zien dat een op reproductie-gerichte leerstijl meer bij vrouwen dan bij mannen te vinden was, terwijl mannen hoger op ongerichtheid scoorden. Daarnaast kwam betekenisgerichtheid veel voor bij Nederlands en geschiedenis, en niet bij wiskunde, terwijl reproductiegerichtheid veel voorkwam bij biologie, en niet bij Nederlands en wiskunde. Er was geen interactie tussen sekse en vak wat betreft leerstijl. De laatste lezing (door Ferguson-Hessler) ging over de effectiviteit van verschillende leeractiviteiten. Ferguson-Hessler gaf eerst een aantal definities van begrippen die ze gebruikte, wat de helderheid van haar betoog zeer goed deed. Daarna gaf ze de resultaten van een literatuurstudie: de activiteiten classificeren, vergelijken, integreren en (extern) testen bleken als zeer effectief uit de verschillende onderzoeken naar voren te komen. Het bestuderen en uitvoeren van de doelgroep, het observeren van de uitvoering van de doelgroep door een model (bijv. docent), en het evalueren bleken daarentegen niet of nauwelijks effectief. Lodewijks (KUB) besloot het symposium met een discussie van de presentaties, waarin hij een leuke anecdote over zijn natuurkunde-geniale zoon vertelde, die de discussie over het aanleren van leeractiviteiten, de

veranderbaarheid van leerstijl, en de noodzaak voor verandering van leeractiviteiten inleidde. Helaas ging Lodewijks te lang door, zodat ik het einde van zijn betoog niet heb meegemaakt. De vraag of je een expert-model zomaar kunt opleggen aan een beginnende leerder, is evenwel interessant genoeg om nog een symposium over te beleggen.

De laatste papersessie bij *Leren & Instructie* was tevens de laatste van de ORD'96. In deze sessie moest ik mijn presentatie geven. Kanselaar (UU, de voorzitter) had mij diezelfde dag verzocht om mijn praatje in te korten tot maximaal 15 minuten. Hij wilde de dag om vijf uur beëindigen (in plaats van kwart over vijf). Het was niet zo erg druk meer; de meeste bezoekers hadden het vanwege de warmte voor gezien gehouden. De papersessie werd geopend door Otten (RUL). Ook hij presenteerde resultaten van een leerstijlen-onderzoek, ditmaal onder leerlingen van de onderbouw van het VO. Hij vond dat de activiteiten kritische verwerking en externe sturing (met betrekking tot het leerproces) niet herkend werden. Daarnaast stelde hij dat een stapsgewijze verwerking niet per se reproductief hoeft te zijn en dat de onderzochte leerlingen een duidelijk onderscheid maken tussen aanwijzingen van de docent en aanwijzingen van het boek. Dit laatste zou dus terug moeten komen in het concept van externe sturing. Drie leerstijlen werden onderscheiden: diepte-leerstijl, reproductieve leerstijl, en een stuurloze leerstijl. De laatste bleek een negatieve invloed te hebben op schoolprestaties. Klatzer (KUN) was als tweede aan de beurt. Zij ging in op leerconcepties bij leerlingen uit de laatste klas van het basisonderwijs. Ze vond drie conceptie-patronen: (1) een beperkte leerconceptie (leren is kennis opnemen, onthouden), (2) een functionele leerconceptie (leren is kennis gebruiken, begrijpen), en (3) een uitgebreide leerconceptie (leren is gebruiken van kennis voor de toekomst). Bij de laatste kwam een duidelijk sociaal aspect naar voren: samenwerken, elkaar helpen, waren begrippen die deze leerlingen vaak noemden. Daarna moest ik: mijn presentatie had betrekking op het effect van de variatie in presentatie en structuur van informatie op het leerproces en het leerresultaat van MBO-leerlingen Werktuigbouwkunde (vakgebied CNC-programmeren). Vier manieren van instructie waren onderzocht: een

tekstboek (domein-gericht), een traditionele handleiding (taakgericht, inclusief procedure-stappen), een minimale handleiding (als de traditionele, maar hier werden de procedure-stappen geïntegreerd aangeboden met de daarvoor benodigde declaratieve informatie) en een open instructie-vorm, waarbij de leerling vragen moest stellen om iets te weten te komen. De minimale handleiding kwam als beste uit de bus wat betreft de procesresultaten; hierbij werd de meeste informatie gelezen en verwerkt, en het minst overgeslagen. Bij de productresultaten bleek dat de macro-structuur van de CNC-programma's het best werd uitgevoerd bij de vragen-groep en de minimale handleiding-groep. Bij de micro-structuur (individuele acties die uitgevoerd moesten worden) bleek de minimale handleiding weer het meest effectief te zijn. De allerlaatste lezing van de ORD (bij 'Leren en Instructie' dan) werd gegeven door Waytens (Universiteit Leuven). Zij ging in op leeroriëntaties bij docenten van het VO, die ze via een enquête onderzocht had over hun onderwijsoriëntatie (leerstofgericht vs. leerlinggericht), hun opvattingen over intelligentie (stabiliteit vs. ontwikkeling), en hun opvattingen over hun vak (ruimte voor eigen inbreng, mate van hiërarchische ordening, en van dynamiek). Een belangrijke uitkomst van de enquête betrof het verschil in opvattingen over intelligentie: leraren die aan lagere klassen/schooltypen lesgeven vonden eerder dat intelligentie een vast, stabiel persoonskenmerk is, terwijl leraren van hogere klassen/schooltypen er vaker van uitgingen dat iemands intelligentie zich kan ontwikkelen door de tijd.

In gebouw Q van de KUB vond de tweede papersessie van het thema "Krachtige Leeromgevingen" plaats. De zaal zat redelijk vol. In de eerste lezing van Smeets en Mooij (KUN) werd ingegaan op multimedia-gebruik in een aantal lessen aardrijkskunde. Vier leeromgevingen, die varieerden in sturing (van heel gestructureerd naar heel open), werden vergeleken. De meest open leeromgeving bleek te leiden tot meer interactie (ook tussen leerlingen en docent) en meer zelfstandigheid bij de leerlingen. Er werd geen verschil in leerresultaat gevonden, ook waren leerlingen niet taakgerichter noch konden ze beter omgaan met ver-

schillende typen informatiebronnen, hoewel dit wel verwacht werd. Dat deze verwachting niet bevestigd werd, werd toegeschreven aan het feit dat leerlingen niet gewend waren aan zoveel vrijheid in de regulering van het leerproces. Enkele toehoorders hadden wat moeite met deze verklaring. De tweede lezing van Kommers (UT) werd overgenomen door Lan- zing (UT). Hij stelde de instructietechniek van 'concept-mapping' centraal, en vroeg het publiek om na te denken over een meetmethode, waarmee de ontwikkeling van concept- maps gemeten kan worden. Van de Ven (TUD) vervolgde de papersessie met een lezing over sturende opdrachten bij computersimulaties. Hij vergeleek de effecten van vrije opdrachten bij een simulatie met de effecten van sturende opdrachten bij een simulatie of bij een tutorial bij zwakke versus goede studenten. Hij concludeerde dat zwakke leeders het meest baat hadden bij de tutorial met sturende opdrachten en de goede leeders het meest bij de simulatie met de vrije opdrachten. De sessie werd afge- sloten door Mooij (KUN) met een lezing onder de titel "Ontwikkeling van krachtige leeromge- vingen in basis- en voortgezet onderwijs: Beginkenmerken en Leerstoflijnen". Helaas heb ik er niet zoveel van meegekregen. Vol- gens mij had dit twee oorzaken: (1) het klimaat in het zaaltje wilde niet erg meewerken: het was erg warm. (2) Mooij's transparanten wer- den steeds informatie-dichter, en steeds onlees- baarder. Erg jammer.

Over het algemeen vond ik de ORD van 1996 een geslaagde aangelegenheid. Er was veel interessants te zien en te horen (misschien wel veel) en er waren veel interessante onderzoe- kers aanwezig. Wel vond ik het jammer dat sommige sessies een beetje volgepropt waren. Bij vier lezingen binnen anderhalf uur krijgt de toehoorder weinig tijd om de informatie ook echt op te nemen. Voor discussie bleef er vaak al helemaal geen tijd over. Onderwijskundige experts zouden wat dat betreft beter moeten weten. Ten tweede was er nogal wat overlap tussen divisies. Dat is natuurlijk niet altijd te voorkomen, maar de overlap bij met name "Leren & Instructie" en "Hoger Onderwijs & Volwasseneneducatie" was wel erg opvallend dit keer. Ten derde vond ik het erg jammer dat er ter afsluiting van de ORD niets meer georga-

niseerd werd. Ik was niet de enige, want ook andere ORD-gangers kwamen naderhand de Aula binnen druppelen, keken rond, en ver- dwenen weer enigszins teleurgesteld.

Methodologie en Evaluatie

(S. Doolaard, TO, Universiteit Twente, Enschede)

Zoals gebruikelijk lopen de onderwerpen van de presentaties binnen het thema "Methodolo- gie & Evaluatie" sterk uiteen. Opvallend is het grote aandeel van effectiviteitsstudies. In de oproep tot deelname werd dit onderwerp met nadruk genoemd. Van de vier geplande sessies binnen het thema "Methodologie & Evaluatie" waren er twee gevuld met onderzoek naar de evaluatie van onderwijseffectiviteit. Op don- derdag werden in het symposium "Evaluatie van Onderwijseffectiviteit" zeven Groningse onderzoeken gepresenteerd uit een onderzoek- programma dat tot doel heeft elementen uit het onderwijseffectiviteitsmodel van Creemers te toetsen. Het GION onderzoek heeft zich de laatste jaren verbreed van het meten van voor- namelijk effecten op instructie-, klas- en leer- krachtkenmerken tot het meten en evalueren van effecten van kenmerken op verschillende niveaus op verschillende outputmaten die zich op hun beurt weer op verschillende niveaus bevinden. Het symposium maakt duidelijk dat de methodologie, gebruikt in het merendeel van de effectiviteitsstudies, niet alleen geschikt is voor doorsnee effectiviteitsonderzoek met cognitieve leerlingresultaten maar ook voor het meten van bijvoorbeeld de effectiviteit van nascholing of samenwerkingsverbanden WSNS.

In twee van de gepresenteerde onderzoeken is onder andere aandacht voor een van de vaak genoemde problemen met effectiviteitsonder- zoek: het ontbreken van (causale) verklarings- modellen en het correlatieve karakter van dit type onderzoek. Om aan dit probleem tegemoet te komen maken Brandsma en Edelenbos gebruik van een experimenteel design voor het meten van effecten van nascholing. In een aan- tal onderzoeken wordt gebruik gemaakt van de longitudinale data van de evaluatie van het onderwijsvoorrangsbeleid en/of VOCL met metingen op meerdere tijdstippen. Een andere vaak genoemde zwakheid van effectiviteitson- derzoek is de theorie-armoede. Ook hieraan

werd tijdens dit symposium aandacht besteed. Theorietoetsing stuit echter nog op problemen. Reezigt is wel in staat kenmerken te clusteren die consistent zijn en geacht worden samen sterkere en duidelijkere effecten te hebben dan de afzonderlijke kenmerken. Deze vooronderstelling wordt echter nog niet bevestigd in het onderzoek. Ook de toetsing van het model van Carroll door Deinum levert problemen op.

In de andere sessie "Effectiviteitsstudies", werd in twee delen een Vlaams onderzoek gepresenteerd. De presentaties (Coucke, Universiteit van Antwerpen; Rijmenans, Universiteit van Antwerpen) waren helder, de uitkomsten van het onderzoek voorspelbaar: scholen vertonen grote variatie in de mate waarin zij bij leerlingen goede prestaties kunnen bewerkstelligen. Wanneer covariabelen meegenomen worden neemt de variatie af. Om de verschillen tussen leerlingen inzichtelijk te maken zijn deze verschillen omgezet in leerjaren zoals in wel meer onderzoek gebeurt. Deze omrekening riep nogal wat reacties op omdat er bepaalde veronderstellingen over de toename van kennis aan ten grondslag liggen. Van de vele school- en klaskenmerken hadden enkele een significant effect. Het aantal significante effecten kwam echter nauwelijks boven het aantal effecten uit, dat op basis van het gehanteerde significantieniveau, te verwachten is. Evenals in ander schooleffectiviteitsonderzoek hadden sommige factoren significante effecten in een andere dan de veronderstelde richting. Luyten (UT) presenteerde als laatste binnen deze sessie een onderzoek naar de consistentie van effecten. Gedurende de laatste jaren is door Luyten en andere Nederlandse en buitenlandse onderzoekers de consistentie van schooleffecten tussen leerjaren, tussen leerkrachten in parallelklassen, tussen cohorten, meerdere malen het onderwerp van onderzoek geweest. Luyten onderzocht de consistentie tussen Nederlandse taal en rekenen/wiskunde en vergeleek daarbij het basisonderwijs met het voortgezet onderwijs. Op basis van de verschillen in organisatie, een leerkracht voor alle vakken in het basisonderwijs versus verschillende leerkrachten in het voortgezet onderwijs, veronderstelt hij dat er in het basisonderwijs sprake zou zijn van een grotere consistentie tussen beide vakgebieden. Deze veronderstelling bleek door het onderzoek niet waar gemaakt te

kunnen worden. Voor effectiviteitsonderzoek impliceert dit dat het gebruik van de ene outputmaat andere resultaten op kan leveren dan het gebruik van een andere maat. Overigens vonden ook de Vlaamse onderzoeksters niet alleen verschillen tussen lezen en schrijven, maar ook binnen deze vakgebieden vonden zij grotere variantie 'tussen-taken-binnen-leerlingen' dan tussen leerlingen of tussen scholen. In de praktijk, voor schoolkeuzeprocessen, betekent deze uitkomst dat ouders en leerlingen bij het zoeken van een goede school een breed beeld van een school zouden moeten krijgen. Bovendien moeten zij zelf weten waar ze naar op zoek zijn, wat zij belangrijk vinden om te leren. Elke school heeft sterke en minder sterke kanten.

De twee overige sessies hadden respectievelijk de titels "Toepassingen van testtheorie" en "Evaluatie-onderzoek". De drie onderzoeken binnen de eerst genoemde sessie kwamen overeen in het feit dat zij alle over toetsen gingen, zoals de titel al aangaf, en in het feit dat zij helder gepresenteerd waren. Op allerlei andere kenmerken liepen zij nogal uiteen. De eerste presentatie ging over een procedure ontwikkeld door het Cito om de gelijkwaardigheid van examens met wisselende onderwerpen te garanderen met equivalering volgens de IRT-benadering (Alberts en Gijssels, CITO). In de tweede presentatie (Van Vissen (KPC) en Voeten (KUN)) werd de validering van de nieuwe versie van de AVI-toetskaarten besproken met behulp van LISREL-modellen. Het onderwerp van de derde presentatie (Hofstee en Zegers, RUG) was een geautomatiseerde methode, ontwikkeld aan de R.U. Groningen, waarmee na toetsafname die items verwijderd kunnen worden die door de beste studenten slecht gemaakt worden. Hoewel in de titel van de presentatie de term 'IRT' genoemd werd, werkt het programma met p-waarden volgens de klassieke test-theorie. Van de eerste twee presentaties was het duidelijk waar de relevantie van het onderzoek lag. De derde presentatie kan getypeerd worden als een leuk idee, maar gezien het feit dat de tentamenresultaten, berekend volgens deze methode, zeer hoog correleren met resultaten die berekend zijn volgens de traditionele methode, waarbij ook items verwijderd werden, lijkt de relevantie voor studenten onduidelijk. Misschien dat deze nieuwe metho-

de voor docenten meer bruikbare informatie oplevert over de items, die later weer in het onderwijs gebruikt kan worden.

De sessie "Evaluatie-onderzoek" bestond uit twee presentaties. Bosker (UT) besprak de manier waarop in opdracht van een van de koepelorganisaties zoveel mogelijk instrumenten voor zelfevaluatie beschreven en beoordeeld zijn. De instrumenten liepen sterk uiteen in kwaliteit en prijs. Opvallend was het feit dat geen van de instrumenten aan alle vooraf gestelde eisen konden voldoen. In de tweede presentatie (Peeters, Cauwenberghe en Schollaert, Universiteit van Antwerpen) werd een analysemodel voor externe evaluatie van 'navormingsprojecten' uit Vlaanderen besproken. Hiermee wordt geprobeerd zicht te krijgen op de black box, het eigenlijke navormingsproces. Acht criteria voor 'effectieve navorming' zijn opgesteld. Opvallend afwezig zijn effect-criteria op het gebied van gedrag van de deelnemers en hun leerlingen, zoals die in het onderzoek van Brandsma en Edelenbos gebruikt werden. Hiermee ben ik beland bij een van de opvallende zaken van het thema "Methodologie & Evaluatie" tijdens deze ORD: een aanzienlijk deel Vlaamse presenta-

ties, helder gepresenteerd maar qua onderzoek zeer traditioneel. In het oog springend was verder, zoals gezegd, het grote aandeel effectiviteitsstudies, vorig jaar gebundeld binnen een speciaal thema. In veel van deze studies ligt de nadruk ook niet zozeer op de methodologie van het onderzoek, maar op de resultaten. Hierdoor passen deze onderzoeken misschien beter binnen een ander of een apart thema. De vraag is dan echter wat er nog resteert voor het thema "Methodologie & Evaluatie".

Literatuur

- Creemers, B., Lagerweij, N., & Verloop, N. (1996). Constructieve en destructieve discussie over de interdisciplinaire onderwijskunde. Een reactie op De Klerk. *Tijdschrift voor Onderwijsresearch*, 21, 284-286.

Aan deze kroniek werd meegewerkt door:
H. Agterbos, A. Booms, S. Doolaard, R. Maslowski, N. Nieveen, E.A.M. van der Schoot, C.G. Sikkema, J.N. Streumer, A. Zanting

De eindredactie werd verzorgd door
J.N. Streumer

Wordt het studiehuis volgens tekening gebouwd?

1 Inleiding

Op 17 mei 1993 heeft de minister van Onderwijs en Wetenschappen de 'Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs' ingesteld. De leden van de Stuurgroep werden uitgenodigd richting te geven aan de vernieuwing van het onderwijs in de tweede fase van HAVO en VWO. Deze vernieuwing werd noodzakelijk geacht om het onderwijs aan te laten sluiten op ontwikkelingen in de maatschappij, het functioneren van het onderwijs voor individuele leerlingen te verbeteren, de school(organisatie) te optimaliseren, en de afstemming van HAVO en VWO op aansluitende onderwijstypen zoals hoger beroepsonderwijs en wetenschappelijk onderwijs te vergroten. De Stuurgroep is in 1996 opgegaan in het 'Procesmanagement Voortgezet Onderwijs'.

2 Beleidspunten

In een aantal publicaties heeft de Stuurgroep het vernieuwingsbeleid geleidelijk aan uitgewerkt (zie bijvoorbeeld Stuurgroep, z.j., 1994, 1995a, 1995b, 1996). Het belangrijkste zijn de volgende beleidspunten:

1. de leerstof (vakinhoud) weerspiegelt de maatschappelijke behoeften;
2. het onderwijsaanbod is in belangrijke mate gemeenschappelijk;
3. profilering van het aanbod vindt plaats in vier profielen (mede gezien vanuit het vervolgonderwijs op HAVO en VWO);
4. een (gedeeltelijke) individualisering dient in het onderwijsaanbod te worden uitgewerkt, met ruimte voor eigen studieplanning van leerlingen;
5. ordening van de onderwijsprogramma's gebeurt via de 'studielastbenadering'. 'Studielast' is de grootte (uitgedrukt in uren) waarmee de omvang van een vak ofwel deelvak wordt aangegeven. De door een

leerling te leveren hoeveelheid werk en het hierop gebaseerde aantal uren is in eerste instantie afgestemd op 'de gemiddelde leerling' (mede om redenen van landelijke planning en daarnaast ook de financiering van het studiehuis). 'Studielastbenadering' verwijst naar de centrale rol van de leerling, die een inspanningsverplichting heeft wat betreft het te leveren werk en die geleidelijk aan het eigen leerproces in toenemende mate stuurt of plant (Stuurgroep, 1994, 1995b). De studievorderingen worden per leerling vastgelegd in een examendossier. De studie wordt afgesloten met één centraal (landelijk) examen per jaar;

6. kennis, inzicht en vaardigheden zijn geïntegreerd in het leerproces. Leerlingen moeten bijvoorbeeld zelfstandig opdrachten en taken kunnen uitvoeren, al of niet samen met medeleerlingen;
7. de school wordt een 'studiehuis', dat wil zeggen een instituut dat open is in de communicatie en het onderwijskundig beleid naar de omgeving. In de 'hoofdlijnen-nota' schrijft de Stuurgroep: 'De school kan een aantrekkelijk *studiehuis* worden waarin leerlingen en docenten worden uitgedaagd hun vaardigheden uit te breiden en gezamenlijk een gevarieerd, creatief (en daarvoor meer effectief) leerproces aan te gaan' (z.j., p. 50).

3 Onderwijskenmerken in het studiehuis

De beoogde veranderingen in het onderwijs kunnen uitgesplitst worden naar landelijk onderwijsniveau, schoolniveau, docentniveau en leerlingniveau. Op het niveau van **het landelijke onderwijssysteem** zijn, via de uitwerking in *vakkenprofielen*, in het onderwijsprogramma landelijke richtlijnen ter bepaling en beoordeling van de vorderingen van een leerling ontwikkeld. Hierop is de inhoud en regelgeving wat betreft het landelijke schriftelijke eindexamen mede te baseren. De landelijke richtlijnen bieden tevens handvatten ter invulling van (varianten van) het examendossier voor elk van de voorgeschreven *vakken*. De studielastbenadering per vak resulteert in een *norm-studielast* (aantal uren voor de 'gemiddelde' leerling HAVO of VWO) in elk van de vakken. Deze norm-studielast vormt de

basis voor het ontwerpen van de benodigde studiematerialen en schoolboeken.

De leerstofinhoudelijke, didactische, pedagogische en organisatorische vernieuwingen in de tweede fase krijgen concreet gestalte in de onderwijskundige en organisatorische vormgeving van het studiehuis, ofwel **de school**. De Stuurgroep ziet het studiehuis als een schoolconcept waarover met name de direct betrokkenen in de school zich zullen moeten uitspreken. Het 'studiehuis' is een open concept, dat in scholen op verschillende wijzen kan worden ingevuld (Stuurgroep, 1996).

De Stuurgroep benadrukt dat **de docent** van een (in het algemeen) sterk sturende en door overdracht van kennis getypeerde didactiek over dient te stappen op een 'begeleidende didactiek'. Essentieel hierin is het scheppen van gunstige voorwaarden voor het leerproces van de leerling. De nadruk komt te liggen op 'leren leren' en vaardigheid in het toepassen van kennis en inzicht. Belangrijk hierbij zijn het omgaan met verschillen tussen leerlingen en het kunnen hanteren van een variëteit aan werkvormen. Docenten kunnen hun werkzaamheden afwisselender invullen en gaan op meerdere wijzen en op verschillende terreinen met elkaar samenwerken. Verschillende inhoudelijk-didactische onderwijsinvullingen zijn door de docent te combineren met organisatorisch verschillende werkvormen (leerling, subgroep leerlingen, klas, meerdere klassen). Dit geldt ook voor 'norm-studielasten' per groep(ering). Bovendien dient de leerkracht (tegelijktijd, qua management) om te kunnen gaan met de in de groep aanwezige verschillen tussen leerlingen en met meerdere didactische werkvormen.

Op het niveau van **de leerling** vinden *leerprocessen* plaats in een instructie-situatie waarin, meer dan vroeger, sprake is van ofwel individualisering ofwel een gerichte sociale vormgeving. Bij deze leerprocessen dient rekening te worden gehouden met verschillen tussen leerlingen in psychologische leerkenmerken en -voorwaarden. Verwacht kan worden dat een dergelijke individualisering of gerichte socialisering van leerprocessen motivatie-verhogend zal werken, ook in onderwijsleersituaties en

zeker in relatie tot 'zelfstandig leren'.

Realisatie van dergelijke leerprocessen vraagt *instructie-situaties* waarin rekening gehouden wordt met de beginkenmerken van de leerling, de aard en grootte van leerstappen in relatie tot de mogelijkheden van de leerling (leerstijl, cognitieve vaardigheden, sociale vaardigheden), aansluiting op de motivatie, de structurering van de taak (moeilijkheidsgraad, individueel of sociaal leren, didactische werkvorm), de duur van de leerperiode, de mate van benodigde herhaling, de benodigde ondersteuning vanwege de leerkracht, en de mate van zelfstandig (kunnen) werken. Via *leerstoflijnen* is een dergelijke invulling van de instructie-situatie te realiseren, waarbij tevens de studie-last door en ook voor elke leerling geoperationaliseerd wordt (vgl. Van Hennik, 1994; Mooij, 1992, 1994).

De kenmerken van de instructie-situatie zijn in belangrijke mate mede-bepalend voor de aard van de leerprocessen van de leerling. De *mate van zelfstandig werken van een leerling* (tijdens les- of ook zelfstudie-uren) wordt mogelijk gemaakt via een individualiserende inrichting van de instructie-situatie afhankelijk van vak of activiteit per leerling, subgroep (Terwel & Mooij, 1995), klas, jaarlaag, of bijv. alle leerlingen in het studiehuis. De fysieke kenmerken van de leeromgeving zijn doorslaggevend bij actief en zelfstandig(er) leren van de leerling (samen met medeleerlingen). Deze op individueel of sociaal leren afgestemde fysieke kenmerken dienen op voorhand ingericht te zijn. Zelfstandig werken kan bijvoorbeeld geïndiceerd worden door: het zelf of met andere leerlingen kiezen van de volgorde in leeractiviteiten, het zelf of met andere leerlingen ophalen en wegbrengen van leermaterialen, het zelf registreren van het gebruik, het zelf registreren van de vorderingen, het (deels) zelf of gezamenlijk controleren van gemaakt werk, het (deels) zelf of gezamenlijk met andere leerlingen bepalen van volgende leeractiviteiten gegeven de eigen vorderingen, het gebruik van bepaalde multimedia, enzovoorts.

De *feitelijke vorderingen* van een leerling worden uitgedrukt op twee wijzen. Ten eerste in de resultaten van de door een leerling gemaakte *vorderingentoetsen in een vak* (waarmee bepaalde leerstofdelen afgesloten worden, bijv. met tenminste 80% van de opgaven goed).

Ten tweede in de door de leerling gemaakte *absolverende toetsen in het vak* (ten behoeve van het examendossier).

4 Verschil tussen norm-studielast en vorderingen in studielast

Er kan een verschil of discrepantie ontstaan tussen de voor een leerling geldende norm-studielast per vak en de in een bepaalde periode gerealiseerde studielast in termen van de feitelijke vorderingen van de leerling in het vak. Deze *normdiscrepantie* is pedagogisch en didactisch belangrijk voor de leerling én de vak- of ook begeleidingsdocenten. Het gaat er steeds om dat de *verschillen* constructief opgevangen en in de instructie, leerstoflijn(en) of ook de planning van de leerling opgenomen worden. Rekening houden met verschillen tussen leerlingen betekent dan: normdiscrepanties flexibel en constructief integreren in het onderwijsaanbod van het studiehuis.

Bij geconstateerde afwijkingen ten opzichte van de norm-studielast is essentieel dat eerst de betreffende leerling(en) toelicht(en) wat er aan de hand is. Oorzaken kunnen ook buiten de instructie-situatie liggen. Naarmate de oorzaken in de instructie-situatie liggen, zijn globaal twee aanpakken geïndiceerd:

- wanneer een leerling achterloopt op de norm-studielast zijn pedagogisch-didactische, remediale of organisatorische wijzigingen, of ook andere maatregelen, aan de orde;
- wanneer een leerling voorloopt op de norm kan het programma voor deze leerling – met dat voor tenminste één of meer andere leerlingen – verzaaid, aangevuld of versneld worden (vgl. Mooij, 1994). De leerlingen zelf kunnen hierbij de belangrijkste ondersteunende rol vervullen (o.c.).

De inhoudelijk-organisatorische kenmerken van het onderwijs in het studiehuis maken het in principe mogelijk dat de onderwijsleerprocessen inclusief de vorderingenbepaling en toetsing per leerling of per subgroep van leerlingen, maar in elk geval op leerlingniveau, gemodelleerd en van specifieke aanwijzingen of handelingsplannen voorzien worden. Afwijkingen van één of meer leerlingen van dit 'gemiddelde patroon' zijn constructief didactisch te vertalen in individueel of sociaal gemo-

delleerde leerstofconsequenties op individueel, (sub)groeps- en schoolniveau.

Het pedagogisch-didactisch, onderwijskundig, sociaal en onderwijsorganisatorisch constructief *kunnen omgaan met verschillen tussen leerlingen* wordt hiermee 'beheersbaar' of wel 'stuurbaar' (zoveel mogelijk vanuit de eigen sturing van de leerlingen). Het belangrijkste kenmerk van het studiehuis is dan de mate waarin er in de instructie echt rekening gehouden wordt met *verschillen in leren tussen leerlingen*. *Naarmate dit het geval is, zal er minder sprake kunnen zijn van eenduidige, uniforme leerprocessen en van één normatieve studielast voor de leerlingen*.

5 Toetssteen voor de implementatie van het studiehuis

De toetssteen voor de implementatie van het studiehuis is dan de mate waarin er rekening gehouden wordt met de verschillen tussen leerlingen, dus de mate waarin ten behoeve van het studiehuis niet één norm-studielast maar juist alternatieven in dit verband ontwikkeld of mogelijk worden. Het gaat er bijvoorbeeld concreet om of de leerlingen:

- binnen vakken of vakdelen kunnen kiezen uit alternatieve instructie-kenmerken (bijvoorbeeld: eenvoudiger, gemiddeld, moeilijker; of: alleen, met anderen; of: didactisch herhalend, uitbreidend) met verschillende leersnelheden;
- via individuele en sociale werkwijzen inderdaad zelfstandig(er) werken en zelfstandig(er) leren;
- variatie tonen in tempo waarin de studielast gerealiseerd wordt, in termen van feitelijke vorderingen per vak;
- afhankelijk van hun omstandigheden variatie tonen in de wijzen waarop omgegaan wordt met hun normdiscrepanties (verschillen tussen gerealiseerde en norm-studielast);
- bij het optreden van normdiscrepanties constructief ondersteuning krijgen van hun docenten, zowel bij 'positieve' als 'negatieve' normdiscrepanties.

6 Ontwikkeling

Inmiddels hebben 13 vakontwikkelgroepen de landelijke advies-examenprogramma's uitge-

werkt (zie Uitleg, 24 januari 1996, bijlage 2e fase berichten). De verschillende netwerken van scholen en docenten, en de educatieve uitgeverijen, zijn bezig met het vervaardigen van passende onderwijsleermaterialen. Deze materialen concretiseren de studielastbenadering in termen van leerstof en leeractiviteiten (voor de 'gemiddelde' leerling, met het oog op het behalen van het eindexamen). In 1998 dienen de op deze adviesprogramma's gebaseerde onderwijsleermaterialen en methoden gereed te zijn. Deze bieden als het ware een 'landelijke' invulling van het vernieuwde onderwijs in de tweede fase.

Geconstateerd kan worden dat, tot nu (1996) toe, met name slechts de norm-studielasten worden uitgewerkt. Praktijkervaringen in ontwikkelscholen wijzen tevens uit dat de huidige aanpak en werkwijze leiden tot strak geregisseerde volgorden van per onderwijstype identieke opdrachten, die elke leerling in een strak tijdschema dient af te ronden. Er is nog weinig variatie in deze opdrachtenschema's te ontdekken.

7 Conclusies

Het belangrijkste kenmerk van het studiehuis is de mate waarin er in de instructie-processen echt rekening gehouden wordt met *verschillen in leren tussen leerlingen*. Naarmate dit het geval is, zal er minder sprake kunnen zijn van uniforme leerprocessen.

Bezien wij de huidige ontwikkeling, dan is de eerste conclusie dat het nog niet zeker is dat dit belangrijkste kenmerk ook daadwerkelijk in elke school of elk studiehuis gerealiseerd gaat worden. In elk geval blijkt dit nog weinig of niet uit de inmiddels verschenen prototypen van (onderdelen van) vakwerk- of leerplannen in studiewijzers of studiekalenders e.d.. Ook de programmeringen van de instructie-processen zelf laten nog weinig variatie zien. Dit doet een groot beroep op de eigen inzet en kwaliteit van de school, hetgeen niet door elke school is op te brengen. Bovendien brengt dit als risico mee dat er pedagogisch, psychologisch of didactisch niet-optimale of verkeerde invullingen ontstaan.

Ten tweede lijkt het er op dat in de huidige landelijke uitwerking de docent 'vervangen wordt' door de studiewijzer of studiekalender.

De vraag kan gesteld worden of dit werkelijk beter uitpakt voor leerlingen die naar beneden of boven van het 'gemiddelde' afwijken. Met name in de situatie dat dergelijke leerlingen geconfronteerd worden met min of meer uniforme leerprocessen.

8 Toch naar het studiehuis

Met behulp van *leerstoflijnen* (zie par. 3) zijn als het ware operationaliseringsen te maken van de norm-studielast voor een leerling (gegeven de beginkenmerken van de leerling, het gekozen profiel en het gekozen onderwijstype). De docenten kunnen, rekening houdend met de leerlingen in school en hun eigen praktijkwensen en -invullingen, de inhouden en vormgevingen van leerstoflijnen mee-programmeren. Via de registratie verbonden aan het werken met leerstoflijnen is tevens direct inzichtelijk in welke mate zich *individuele afwijkingen of ook andere invullingen van onderwijsleerprocessen* in de school voordoen. De inhoudelijk-organisatorische implementatie van het studiehuis wordt dan uitgedrukt in de *flexibelere planning van onderwijsinhouden en werkvormen* in termen van norm-studielasten, de (toegestane) afwijkingen hiervan, en de wijzen van omgaan hiermee (bijv. ook in de vorm van concrete handelingsplannen). Er is sprake van 'leerlinggericht onderwijs'.

Deze flexibelere, leerlinggebaseerde inhouden en werkvormen maken het mogelijk dat de organisatorische implementatie van het studiehuis gekenmerkt kan worden door een *flexibelere programmering van onderwijstijden* (blokken, periodes of kavels) en door *meer variatie in sociale groeperingsvormen van leerlingen ten behoeve van de instructie* (klassikaal, subgroepen, bovenklassikaal, individueel) dan gebruikelijk. In de organisatie van het studiehuis kunnen in dit opzicht verschillende organisatie-modellen ontstaan, bijvoorbeeld organisatie op basis van een les- en studierooster, dagdelen, of blokken (Stuurgroep, 1996). De kern van het studiehuis is echter niet de organisatie ervan of het werken in blokken, maar de mate waarin er in de instructie-processen echt rekening gehouden wordt met *verschillen in leren tussen leerlingen*.

T. Mooij

Literatuur

- Hennik, D.J. van (1994). Het project begaafdheidsverschillen en onderwijsaanbod. In P.R.J. Simons & J.G.G. Zuylen (Eds.), *Actief en zelfstandig studeren in de tweede fase. Onderwijskundige en organisatorische vormgeving* (pp. 61-71). Tilburg: MesoConsult.
- Mooij, T. (1992) (Ed.). *Op weg naar een flexibele school: enkele voorbeelden van leerlijnen*. Nijmegen: Katholieke Universiteit, Instituut voor Toegepaste Sociale wetenschappen.
- Mooij, T. (1994). Leerlingverschillen in leerstoflijnen. In P.R.J. Simons & J.G.G. Zuylen (Eds.), *Actief en zelfstandig studeren in de tweede fase. Onderwijskundige en organisatorische vormgeving* (pp. 162-174). Tilburg: MesoConsult.
- Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs (z.j.). *Tweede fase. Scharnier tussen basisvorming en hoger onderwijs. Een uitwerking op hoofdlijnen*. Den Haag: auteur.
- Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs (1994). *De tweede fase vernieuwt. Scharnier tussen basisvorming en hoger onderwijs, deel 2*. Den Haag: auteur.
- Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs (1995a). *Vershil moet er zijn*. Den Haag: auteur.
- Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs (1995b). *Examen in het studiehuis*. Den Haag: auteur.
- Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs (1996). *Organisatie in het studiehuis*. Den Haag: auteur.
- Terwel, J., & Mooij, T. (1995). De relevantie van sociale contexten: klas en groepjes als sociale hulpbronnen bij het wiskunde-onderwijs. *Sociologische Gids*, 42, 301-317.

Boekbespreking

W. Pols, S. Miedema en B. Levering (red.)

Opvoeding zoals het is

De Tijdstroom, Utrecht, 1995, 258 pagina's, f 49,-
ISBN 90-352-1591-5. Verschenen in de reeks Kinder- en
Jeugdstudies

Dit boek is een sterk gewijzigde versie van een boek dat voor het eerst in 1989 verscheen. Het is destijds ontstaan, aldus de redactie in haar Inleiding "Over wetenschappelijke en pedagogische waarheid", uit onvrede met de wijze waarop volgens haar in de PABO met het verschijnsel opvoeding werd omgegaan. Naar haar mening werd de pedagogische werkelijkheid gereduceerd tot een onderwijskundig te beheersen geheel. De redactie wil de pretenties van de onderwijskunde relativeren door pedagogische vragen opnieuw onder de aandacht te brengen, en wel door een andere aanpak dan die van de traditionele pedagogieboeken, die zij te theoretisch vindt.

De andere aanpak die de redactie volgt, duidt zij aan met de term narrativiteit. Het uitgangspunt wordt genomen in documenten van verschillende aard die een weerslag zijn van pedagogische ervaringen, aldus de redactie. Deze pedagogische voorvallen vormen dan het uitgangspunt voor theoretische doordenking. De redactie acht verhalen in bepaalde opzichten superieur aan wetenschappelijke theorieën. Doordat een aantal belangrijke wetenschappelijke vragen over opvoeding en ontwikkeling in verhalen niet beantwoord worden, kan interpretatie van die verhalen directer toegang geven tot opvoeding zoals het is. Zo'n kwestie is bijv. de scheiding van het normale en het afwijkende. Zo kan een verhaal laten zien hoe het gewone en het uitzonderlijke naast elkaar kunnen bestaan. De belangrijkste kwestie is voor de redactie echter die van waarheid en onwaarheid. Een pedagogische waarheid is echt iets anders dan een waarheid in wetenschappelijke zin, aldus de redactie. Een pedagogische waarheid is een inzicht dat betekenis heeft in praktisch-normatieve zin, het gaat daarbij niet zozeer om feitelijkheden als wel om mogelijkheden. Dit onderscheid tussen wetenschappelijke waarheid en pedagogische waarheid vormt het uitgangspunt voor de betekenis van ieder van de zeven hoofdstukken waaruit het boek na de Inleiding is opgebouwd, nl. de onder-

steuning die het levert aan de stelling dat het bij de pedagogische waarheid gaat om mogelijkheden.

Het eerste hoofdstuk is getiteld: "Het verhaal van Staal. Of: over de betekenis van het verhaal voor de opvoeder en de opvoeding". Het is Theo Thijssen die in zijn boek "De gelukkige klas" de onderwijzer Staal via zijn dagboek laat reflecteren op zijn onderwijspraktijk, en Pols, de schrijver van dit eerste hoofdstuk, laat zien dat de pedagogische waarde vooral te vinden is in de wijze waarop Staal naar pedagogische mogelijkheden zoekt.

Het tweede hoofdstuk is van Levering en getiteld: "De overschatting van de macht van de opvoeder". Het is van geheel andere aard dan het eerste hoofdstuk. Levering bespreekt de vraag in hoeverre de opvoedersverantwoordelijkheid sinds de Tweede Wereldoorlog veranderd is, mede in relatie tot de kwestie van overschatting van de macht van de opvoeder, en wel op basis van een analyse van verhalen van kinderen en volwassenen over hun eigen opvoeding.

In hoofdstuk drie maakt Pols pedagogisch opmerkingen bij de roman van Simon Vestdijk: "De kinderjaren van Anton Wachter", waarin laatstgenoemde in verhaalvorm zijn eigen kinderjaren reconstrueert.

In hoofdstuk vier geeft Vreugdenhil een meer systematische beschrijving van wat opvoeding in de klas inhoudt n.a.v. fragmenten uit het boek van de Amerikaanse schrijver Kidder, die een jaar lang het werk van een lerares volgde.

Hoofdstuk 5 handelt over allochtone kinderen in 'witte' en 'zwarte' scholen: zijn ze er of zijn ze er niet?, en is van de hand van Teunissen. Het verhaal speelt in dit hoofdstuk een geringere rol, het hoofdstuk heeft meer het karakter van een wetenschappelijke verhandeling.

Hoofdstuk zes draagt als titel: "Moderne pedagogische gedachten in de achttiende eeuw". Pols bespreekt daarin de ideeën van Betje Wolff zoals zij die heeft neergelegd in haar boek: "Proeve over de opvoeding".

In hoofdstuk zeven tenslotte behandelt Miedema "Het verhaal van de wetenschap. Over opvoedingswetenschappelijke theorieën en opvoedkundige praktijken". Hij doet dit door twee relevante kritieken als verhaal te analyseren, nl. die van de Duitse anti-pedagoog Von Braunnuehl en die van de vroegere Nijmeegse lector Brus in "Didactiek naar menselijke maat".

De auteurs hebben zonder twijfel een interessante poging gedaan om op een nieuwe manier in de opleiding voor het basisonderwijs meer aandacht te vragen voor het aspect van de opvoeding in het onderwijs, al vind ik het persoonlijk jammer dat dit via een afzonderlijk boek moest gebeuren, waardoor de onderwijskunde nog meer apart wordt gezet als technologie. Men kan zich afvragen of dit de beste manier is om aan de technologiekritiek waarvan de redactie spreekt, tegemoet te komen. Hoe het ook zij, de auteurs hebben in ieder geval een energieke poging gedaan om er iets aan te doen. Toch hebben we bij deze poging een aantal kanttekeningen.

In de eerste plaats wordt niet geheel duidelijk wat de redactie met dit boek bij studenten van de PABO wil bereiken. Het is niet de bedoeling om eenvoudigweg pedagogische theorie aan te bieden, de verhalen moeten het uitgangspunt vormen voor pedagogische doordenking. Die doordenking wordt door de auteurs van het boek zelf ondernomen, en in het algemeen leveren zij daarmee een interessante bijdrage. Maar wat wordt nu van de studenten verwacht? En waarop zijn activiteiten van studenten gericht? Studie-

opdrachten en aanwijzingen voor de PABO-docent ontbreken. Dat is jammer, want een narratieve aanpak in de opleiding is nog vrij onbekend, dit in tegenstelling tot een moderne methode als de case-aanpak als ook narratief onderzoek. Het nieuwe van de aanpak die de auteurs volgen, is er, neem ik aan, de oorzaak van dat het begrip verhaal op zeer verschillende manieren wordt ingevuld, zoals uit bovenstaand overzicht van de inhoud van het boek mag blijken. Maar dat geldt overigens ook voor de inhoud van het zogenaamde narratieve onderzoek. Hoe dan ook, de didactische aanpak van de narratieve benadering in de opleiding van onderwijsgevendend verdient nadere aandacht en uitwerking, en daartoe levert dit boek een belangrijke stimulans. Wij hopen dat pedagogiek-/onderwijskundedocenten op de PABO dit boek zullen gebruiken om met deze aanpak te experimenteren, om daarmee tevens de gewenste aandacht te vragen voor het aspect van de opvoeding dat in het basisonderwijs een wezenlijk element vormt.

F.K. Kieviet

Mededelingen

Studiehuis in de steigers

Op 27 maart 1997 organiseert het Procesmanagement Voortgezet Onderwijs in de Jaarbeurs te Utrecht voor het derde jaar de tweede fase dag 'Studiehuis in de steigers'.

Voor nadere informatie en aanmelding kunt u zich wenden tot het congressecretariaat: Secretariaat Studiehuis in de steigers 3, p/a MesoConsult, Gounodlaan 15, 5049 AE Tilburg, tel.: 013 4560311, fax: 013 4563276.

Promoties

Universiteit van Amsterdam
Promovendus: Drs. C.L.M. Bodin-Baarends
Titel proefschrift: Contacten tussen ouders en leerkrachten bij leerproblemen
Promotor: Prof. Dr. K. Doornbos
Datum: 19 november 1996

Universiteit van Amsterdam
Promovendus: Drs. M.J.M. Vermeulen
Titel proefschrift: Human capital in the Hinterland
Promotor: Prof. Dr. J. Dronkers
Co-promotor: Prof. Dr. P. Nijkamp
Datum: 20 november 1996

Universiteit van Amsterdam
Promovendus: Drs. J. Meijer
Titel proefschrift: Learning potential and fear of failure
Promotor: Prof. Dr. J. Elshout
Datum: 11 december 1996

Universiteit van Amsterdam
Promovendus: Drs. M. de Graaf
Titel proefschrift: Opvoeding en vrouwelijkheid bij Jean Jacques Rousseau
Promotoren: Prof. Dr. G.F. Heyting en Prof. Dr. F.L.A. Rang
Datum: 13 december 1996

Rijksuniversiteit Groningen
Promovendus: Drs. J. Dane
Titel proefschrift: De vrucht van bijbelse opvoeding. Populaire leescultuur en opvoeding in protestants-christelijke gezinnen, circa 1880-1940.

Promotor: Prof. Dr. J. Dekker
Datum: 21 november 1996

Rijksuniversiteit Leiden
Promovendus: Drs. H.P.J. Witteman
Titel proefschrift: Styles of learning and teaching in an interactive learning group system
Promotor: Prof. Dr. M. Boekaerts
Datum: 27 februari 1997

Rijksuniversiteit Leiden
Promovendus: P. Verhoeven
Titel proefschrift: Tekstbegrip in het onderwijs Klassieke Talen
Promotoren: Prof. Dr. N. Verloop en Prof. Dr. C.M.J. Sicking
Datum: 5 maart 1997

Inhoud andere tijdschriften

Comenius
16e jaargang, nr. 4, 1996

Thema: Kinderen en de straat
Kinderkopjes, door J. Doornbal en C. van Nijnatten
Van wie is de straat? Over kinderen, genderidentiteiten en het publieke domein, door L. Karsten

De straatvrees van pedagogen. Over de opvoedkundige waarde van jeugdparticipatie, door M. de Winter
en verder:

'De pijn wakker houden, het verlangen wekken'. Enkele bedenkingen bij de nood aan waardenopvoeding. Deel 2. Over het denken, door J. Masschelein

Kind en ruimte. De positie van het kind bij de Stichting Goed Wonen (1946-1968), door W. van Moorsel
Pril geluk. De relatiedynamiek rond mishandeling van meisjes en jonge vrouwen door hun vriend, door S. Mastenbroek en R. Römken

Pedagogisch Tijdschrift
21e jaargang, nr. 4/5, 1996

Thema: Psychosociale problemen in de adolescentie
Inleiding. Gaat het goed met jongeren in Nederland? door W. Vollebergh en W. Meeus

Gezin en modernisering. Een problematische relatie voor jongeren, door M. Du Bois-Rey-mond

Tienjaars trend in het voorkomen van probleemgedrag van Nederlandse kinderen en jeugdigen, door F.C. Verhulst, J. van der Ende en A. Rietbergen

Jeugd, criminaliteit en morele paniek in Nederland, door C. Baerveldt

Opvoeding en probleemgedrag in de adolescentie, door W. Meeus, M. Deković en M. Noom

Gehechtheid, sociaal netwerk en thuisloosheid bij jongeren, door J.W.C. Tavecchio en M.A.E. Thomeer-Bouwens

Welbevinden en etnische identiteit bij allochtone jongeren, door W. Vollebergh en A. Huijberts

Het adaptief vermogen van opvoeders en adolescenten, door J. Hermanns en W. Slot

Tijdschrift voor Orthopedagogiek
35e jaargang, nr. 11, 1996

Jeugdreclassering en de motivatie van delinquente jongeren, door C.H.C.J. van Nijnatten en K.M. van Berchum

Naschoolse dagbehandeling van licht verstandelijk gehandicapte kinderen, door E.F.M. van Hoof en J. Pellen

Vraag en aanbod in de opvoedingswinkel. Een experimentele vorm van opvoedingsondersteuning in Den Bosch, door A.J.E. de Veer en A. Kerkstra

Ontvangen boeken

Baartman, H.E.M., *Opvoeden kan zeer doen Over oorzaken van kindermishandeling, hulpverlening en preventie*. Uitgeverij SWP, Utrecht, 1996, f 49,50.

Braams, T., *Dyslexie: een complex taalprobleem*. Uitgeverij Boom, Meppel, 1996, f 29,50.

Dam, G. ten, Hout, H. van, Terlouw, C., Willems, J. (red.), *Onderwijskunde hoger onderwijs. Handboek voor docenten*. Van Gorcum, Assen, 1996, f 69,50.

Dulk, H.O. den, & Kooten, A.H. van (red.), *Het kan gewoon anders. De orthopedagogische kwaliteit van de leraar*. Uitgeverij Intro, Baarn, 1996, f 29,90.

Geerligts, T., & Veen, T. van der, *Lesgeven en zelfstandig leren*. Van Gorcum, Assen, 1996, f 72,50.

Gerris, J.R.M. (red.), *Gezin: Welzijn, gezondheid en hulpverlening*. Van Gorcum, Assen, 1996, f 37,50.

Grimhuizen, T. van, *Krijg ze maar groot. Een leerboek over de adolescentie en het begeleiden van jongeren*. Uitgeverij SWP, Utrecht, 1996, f 29,50.

Haar, M. ter, *Professionals in de beleidsarena. Basisboek beleid*. Van Gorcum, Assen, 1996, f 29,50.

Hajer, M., *Leren in een tweede taal. Interactie in vakonderwijs aan een meertalige MAVO-klas (dissertatie)*. Wolters-Noordhoff, Groningen, f 49,-.

Huisman, P.E., *Kennis gewogen. Analyse van sociaal-wetenschappelijk denken. Kritiek en aanwijzingen*. Van Gorcum, Assen, 1996, f 55,-.

Jongh, A.J.C.M., de, & Scholten, I.W.T. (red.), *Opleidingsdidactiek voor de zorgsector*. Bohn Stafleu Van Loghum, Houten, 1996, f 49,75.

Kalverboer, A.F. (red.), *De nieuwe buitenbeentjes. Stoornissen in aandacht en motoriek bij kinderen*. Lemniscaat, Rotterdam, 1996, f 64,50.

Korzilius, H., *Mental representations of parenting. Structures of perceptions, cognitions, emotions and behavioral reactions in child-rearing situations (dissertatie)*. Katholieke Universiteit, Nijmegen, 1996, f 35,-.

Mossevelde, E. van, *De klas in de hand. Omgaan met en (bege)leiden van leerlingen*. Acco, Leuven/Amersfoort, 1996, f 46,75.

Ven, P.H. van de, *Moedertaalonderwijs. Interpretaties in retoriek en praktijk, heden en verleden, binnen- en buitenland (dissertatie)*. Wolters-Noordhoff, Groningen, 1996, f 79,-.

Yablonsky, L., *Van man tot man. Vader en zoon: een spannende relatie*. De Toorts, Haarlem, 1996, f 45,-.

Leeruitkomsten van leerlingen in combinatieklassen: een 'best-evidence'-synthese

S. Veenman*

Samenvatting

In dit artikel wordt een overzicht gegeven van de leeruitkomsten van de leerlingen in combinatieklassen van de basisschool. Een combinatieklas is een klas waarin leerlingen van twee of meer leerjaren bijeen zijn geplaatst in één klas. Een onderscheid wordt gemaakt tussen combinatieklassen die om beheersmatig/organisatorische motieven zijn gevormd en combinatieklassen die om pedagogisch/didactische motieven zijn gevormd. Bij de leeruitkomsten wordt een onderscheid gemaakt tussen de cognitieve en de niet-cognitieve prestaties. Samengevat worden de resultaten van 59 studies. Het resultaat van deze analyse van de beschikbare empirische gegevens (a best-evidence synthesis) laat zien dat de leerprestaties van de leerlingen in de combinatieklassen niet systematisch verschillen van de leerprestaties van de leerlingen in de enkelvoudige klassen. Ook de studies met betrekking tot de niet-cognitieve prestaties laten geen consistente verschillen zien. De in Nederland uitgevoerde studies naar de leeruitkomsten van de combinatieklassen laten eveneens geen systematische verschillen zien. Op grond van gevonden uitkomsten wordt geconcludeerd dat er op dit moment geen empirische bewijskracht bestaat voor de veronderstelling dat de combinatieklas als organisatievorm het leren van de leerlingen benadeelt.

1 Achtergrond

'Bewijs maar eens dat leerlingen in combinatieklassen niet minder leren dan leerlingen in

enkelvoudige klassen'? Deze vraag werd door verontruste ouders in 1981 aan mij gesteld als voorzitter van het bestuur van een basisschool. Wat was er aan de hand? Door het teruglopen de aantal leerlingen diende deze school noodgedwongen over te gaan tot het combineren van groepen. Op de gestelde vraag kon ik toentertijd geen antwoord geven. Onderzoek naar het functioneren van de combinatieklas was in Nederland niet verricht. Bekend was alleen het onderzoek van de Inspectie van het Basisonderwijs uit 1978, maar dit betrof slechts de opvattingen van leerkrachten en ouders over de combinatieklassen en niet de effecten van de combinatieklas op het leren van de leerlingen. In dit onderzoek kwam de inspectie tot de conclusie dat veel leerkrachten en ouders de combinatieklas in het algemeen als nadelig beschouwden voor het leren van de leerlingen. Sinds 1981 is de vraag naar het effect van de combinatieklas op het leren van de leerlingen herhaaldelijk gesteld, zowel door ouders als door leerkrachten en schoolbesturen.

In het rapport *Schaal en kwaliteit in het basisonderwijs* van de Projectgroep Schaalvergroting Basisonderwijs (1990) wordt beweerd dat vergroting van de scholen ten goede zal komen aan de kwaliteit van het basisonderwijs, met andere woorden: grote scholen zijn beter dan kleine scholen. Gewezen wordt op een positief verband tussen schoolgrootte en CITO-scores voor rekenen en taal. Omdat kleine scholen per definitie scholen zijn met combinatieklassen wordt gesuggereerd dat het leerrendement in combinatieklassen minder zou zijn dan in enkelvoudige klassen van grote scholen. Op grond van operatie *Toerusting en Bereikbaarheid*, die na het uitbrengen van het rapport in gang is gezet, dienen meer dan 1000 kleine basisscholen te verdwijnen. In 1994 concludeerde de Commissie Evaluatie Basisonderwijs (1994a, p. 54): "De combinatieklas is echter in veel gevallen een ingewikkelde klasseorganisatie waar de meeste leraren nau-

* Dank ben ik verschuldigd aan J. van Leeuwe en L. Bouts van de afdeling RTD van de subfaculteit der Pedagogische Wetenschappen en Onderwijskunde voor hun hulp bij het berekenen van effectgroottematen voor studies met onvolledige gegevens.

welijks op zijn voorbereid... De invloed van combinatieklassen op leerprestaties en extra onderwijsbehoeften van leerlingen, is echter niet duidelijk”.

Nu, na meer dan 10 jaar bemoeienis met de combinatieklas wil ik trachten antwoord te geven op de vraag waarmee ik dit artikel begonnen ben. Ik wil dat doen tegen de achtergrond van wat internationaal bekend is over de effecten van de combinatieklas op het leren van de leerling. Door de resultaten van vele studies met elkaar te vergelijken en cijfermatig samen te vatten in de vorm van effectgrootte-maten kan een antwoord worden geformuleerd dat berust op de op dit moment best beschikbare evidentie of bewijskracht.

2 De combinatieklas

Een combinatieklas is een klas waarin leerlingen van twee of meer leerjaren bijeen zijn geplaatst in één klas. In België worden dergelijke klassen graadklassen genoemd. Dit bij elkaar plaatsen van leerlingen uit verschillende leerjaren heeft tot gevolg dat de verschillen tussen de leerlingen groter zijn dan tussen de leerlingen in een enkelvoudige klas.

Overall op de wereld vinden we combinatieklassen. In de westerse geïndustrialiseerde landen komen ze vooral voor op het platteland of in dunbevolkte streken. Maar vandaag de dag komen ze ook voor in dicht bevolkte stedelijke gebieden. Dit komt omdat sinds de jaren zeventig het aantal geboorten is afgenomen waardoor de scholen minder leerlingen kregen en daardoor voor elk leerjaar geen aparte leerkracht meer konden aanstellen. Door combinatieklassen samen te stellen kan de school het hoofd bieden aan fluctuaties in de instroom van leerlingen (terugloop maar ook toeloop) en aan de ongelijke verdeling van het aantal leerlingen over de klassen of groepen.

In Nederland heeft 53% van de leerkrachten uit het basisonderwijs een combinatieklas (Commissie Evaluatie Basisonderwijs, 1994b). Ook in andere landen, binnen en buiten Europa, komt de combinatieklas veelvuldig voor (Thomas & Shaw, 1992). In een groot land als Canada zou men verwachten dat de meeste combinatieklassen voorkomen in kleine scholen in dunbevolkte gebieden. Dit is echter niet

zo. De meeste combinatieklassen zijn te vinden in stedelijke gebieden. In dit land zit één op de vijf leerlingen in een combinatieklas (Gayfer, 1991). Ontwikkelingslanden die streven naar 'basisonderwijs voor iedereen' moeten vaak om economische redenen scholen stichten met combinatieklassen. Per school zijn er te weinig leerlingen om een afzonderlijke klas te vormen. In Aziatische, Afrikaanse, en Zuid-Amerikaanse landen vinden we daarom duizenden scholen met combinatieklassen (UNESCO, 1989; Lungwangwa, 1990; Psacharopoulos, Rojas & Velez, 1993). Uit de beschikbare gegevens kan worden opgemaakt dat de combinatieklas in het basisonderwijs wereldwijd een prominente plaats inneemt.

3 Voor- en nadelen van combinatieklassen

Voor het huidige basisonderwijs is de enkelvoudige klas de norm. Dit is niet altijd zo geweest. Op schilderijen uit de 18de en de 19de eeuw kan men zien dat oorspronkelijk het volksonderwijs niet plaats vond in klassen met leeftijdgenoten. Oudere en jongere leerlingen zaten in één groep bij elkaar. Onder invloed van de industrialisatie en de verstedelijking van het platteland ontstaan in de loop van de 20ste eeuw geleidelijk aan scholen met een vaste klasse-indeling naar leeftijd. Momenteel zijn er twee redenen om van de enkelvoudige klas als norm af te wijken. De eerste reden is, zoals we gezien hebben, van beheersmatige of organisatorische aard. Combinatieklassen worden noodgedwongen gevormd als gevolg van het dalend aantal leerlingen of vanwege een ongelijke jaarlijkse instroom van leerlingen. De tweede reden om combinatieklassen te vormen zijn pedagogische of didactische motieven. In sommige vormen van onderwijs wordt het plaatsen van leerlingen met verschillende leeftijden in één klas of groep bewust toegepast. De stamgroep uit het Jenaplan-onderwijs is hier een voorbeeld van. Sommige traditionele basisscholen hebben ook uit ideële motieven combinatieklassen ingevoerd. In plaats van combinatieklassen spreekt men op deze scholen liever van klassen met heterogene groepen.

In de literatuur brengen voorstanders van

combinatieklassen of heterogene groepen meestal de volgende voordelen van deze onderwijsorganisatie naar voren: (1) onderwijs geven aan leerlingen die onderling sterk van elkaar verschillen dwingt de leerkracht tot het geven van gedifferentieerd onderwijs, (2) er ontstaan vastere bindingen tussen de leerkracht en de leerlingen omdat de leerlingen meer dan één jaar bij dezelfde leerkracht blijven, (3) omdat de heterogene groep gelijkenis vertoont met het gezin kunnen niet alleen de cognitieve doelen van de school beter gerealiseerd worden, maar ook de sociale, emotionele en creatieve doelen, (4) de nadelen van het zittenblijven worden gereduceerd, (5) leerlingen leren van elkaar: de jongere leerlingen leren van de oudere leerlingen, de oudere leerlingen leren van de jongere leerlingen (Draisey, 1985; Katz, Evangelou & Hartman, 1993). Deze voordelen worden soms ook naar voren gebracht door scholen die door externe omstandigheden gedwongen zijn om combinatieklassen te vormen. Op deze wijze maken zij van de nood een deugd.

Naast voorstanders bestaan er ook tegenstanders van combinatieklassen. De tegenstand van de kant van de leerkrachten en de schoolleiding richt zich vaak op de zwaarte van het werk of de taakbelasting in de combinatieklas. Verder worden in de literatuur veelvuldig de volgende problemen genoemd: gebrek aan aandacht voor de problemen van individuele leerlingen, geen tijd voor remediërende activiteiten, gebrekkige opleiding, inadequate leer-materialen, zorgen van de ouders over het leren van hun kinderen in de combinatieklas (Campbell, 1993; Gayfer, 1991; Pogli & Strittmatter, 1983; Veenman, Lem & Winkelmolen, 1985; Waraksa, 1989). Ouders hebben vaak problemen met de combinatieklas omdat zij denken dat leerlingen in combinatieklassen minder leren dan in enkelvoudige klassen (Inspectie Basisonderwijs, 1978; Walsh, 1989). Dit komt omdat zij het onderwijs in de enkelvoudige klas als de norm voor goed onderwijs beschouwen.

Gegeven de heersende opvattingen over de voor- en nadelen van combinatieklassen kan de vraag gesteld worden: 'Wat zijn nu de werkelijke effecten van de combinatieklas op het leren van de leerlingen?' Op deze vraag wil deze studie een antwoord geven.

4 Review-methode

In het volgende overzicht worden de resultaten van empirische studies met betrekking tot de cognitieve en niet-cognitieve effecten van de combinatieklas samengevat. De review-methode die hierbij gebruikt wordt is de 'best-evidence synthesis' zoals beschreven door Slavin (Slavin, 1986). Deze methode combineert de werkwijze van de meta-analyse (Glass, McGaw & Smith, 1981) met die van het beschrijvende verslag. In het kort komt deze werkwijze er op neer dat eerst alle beschikbare studies over het desbetreffende onderwerp worden opgespoord. Vervolgens worden methodologische en inhoudelijke criteria geformuleerd waaraan de relevante studies dienen te voldoen. Tot slot worden de resultaten van deze studies samengevat. Deze samenvattingen hebben de status van de beste op dit moment beschikbare bewijskracht. Hierbij wordt rekening gehouden met de inhoudelijke en methodologische bijdragen van de afzonderlijke studies aan deze bewijskracht (Gutiérrez & Slavin, 1992). Waar mogelijk worden de uitkomsten van de afzonderlijke studies uitgedrukt in effectgrootte-maten: het verschil tussen het gemiddelde van de experimentele groep en de controlegroep gedeeld door de standaarddeviatie van de controlegroep. Bijzonderheden met betrekking tot de toegepaste review-methode worden in de volgende paragrafen beschreven.

4.1 Het literatuuroverzicht

Om alle mogelijke relevante studies op te sporen werd een uitgebreide speurtocht door de internationale literatuur ondernomen. Aan plaats, taal en jaar werden geen restricties gesteld. De belangrijkste geraadpleegde literatuur-referaat-systemen waren: ADION (Nederlandstalige literatuur), British Education Index (BEI), Dissertation Abstracts International, Education Resources Information Center (ERIC), European Documentation and Information System for Education (EUDISED), ForschungsInformationsystem Sozialwissenschaften (FORIS), FRANCIS (Franstalige literatuur), PSYCH LIT en Social Sciences Literature Information System (SOLIS, voor Duitstalige literatuur). Verder werden de literatuurlijsten van de primaire studies nagetrok-

ken. Ongepubliceerde Amerikaanse en Canadese dissertaties werden verkregen via University Microfilms International (UMI) te Michigan. In de enkele gevallen dat een studie niet beschikbaar was via ERIC of UMI werd het betreffende instituut aangeschreven. De studies dienden te voldoen aan onderstaande criteria.

4.2 Inhoudelijke sluiscriteria

In het literatuuroverzicht werden alleen die studies opgenomen die gericht waren op de effecten van de combinatieklas op het leren van de leerlingen. Om het probleem van het vergelijken van 'appelen met peren' op het niveau van de onafhankelijke variabele te vermijden, werd een onderscheid gemaakt in studies waarin combinatieklassen werden gevormd om beheersmatig/organisatorische redenen en om pedagogisch/didactische redenen. Deze indeling in twee categorieën leverde geen moeilijkheden op omdat de meeste auteurs een nauwkeurige definitie geven van een combinatieklas. In studies die zich richten op combinatieklassen, die om beheersmatig/organisatorische redenen zijn gevormd, worden naast de redenen om te combineren meestal ook gegevens vermeld over de instroom van de leerlingen, de grootte en de locatie van de school (stad of platteland). In de Engelstalige literatuur heten deze combinatieklassen 'multigrade classes' (versus 'single-grade classes'). Soms worden ze ook wel genoemd 'combination classes', 'double grades' of 'split-grade classes'. In studies die zich richten op combinatieklassen, die om pedagogisch/didactische motieven zijn gevormd, wordt altijd uitgebreid ingegaan op de voordelen van deze vorm van groeperen ten opzichte van het groeperen van de leerlingen naar leeftijd. Deze klassen heten in de Engelstalige literatuur meestal 'multi-age classes' (versus 'single-age classes'). Uiteindelijk voldeden 59 studies aan de criteria zoals hieronder aangegeven. Deze 59 studies werden uitgevoerd in de volgende landen: 34 in de Verenigde Staten, 4 in Canada, 3 in Duitsland, 3 in Engeland, 5 in Nederland, 2 in Zweden, 2 in Finland, 2 in Australië, 1 in Colombia, 1 in Pakistan, 1 in Togo en 1 in Burkina Faso. Van deze studies hadden 48 studies betrekking op combinatieklassen die om reden van beheersmatig/organisatorische motieven waren gevormd en 11 studies op combinatieklassen die

om reden van pedagogisch/didactische motieven werden samengesteld.

Om het probleem van 'het vergelijken van appels met peren' op het niveau van de afhankelijke variabele te vermijden, werden twee vormen van afhankelijke variabelen onderscheiden: (1) cognitieve uitkomsten of leerprestaties en (2) niet-cognitieve uitkomsten. De cognitieve uitkomsten werden weer onderverdeeld in leerprestaties voor de vakken die in de desbetreffende studies werden onderzocht, bijvoorbeeld lezen, taal, rekenen/wiskunde, geschiedenis, Engels als vreemde taal. De niet-cognitieve uitkomsten werden onderverdeeld in: (1) persoonlijke aanpassing, (2) sociale aanpassing, (3) zelfbeeld, (4) attitudes ten opzichte van de school en (5) motivatie. Persoonlijke aanpassing had betrekking op variabelen als gevoelens van geborgenheid, afwezigheid van nerveuze symptomen, afwezigheid van antisociale neigingen, afwezigheid van angst. Het meetinstrument dat hiervoor veelvuldig gebruikt werd was de California Test of Personality. Sociale aanpassing had betrekking op de sociale omgang met medeleerlingen of op de aanpassing aan de school of de klas. Zelfbeeld had betrekking op attitudes ten opzichte van zichzelf, zelfredzaamheid en zelfconcept. Attitudes ten opzichte van de school richtten zich op variabelen als de houding tegenover de leerkrachten of medeleerlingen en schoolklimaat. Motivatie had betrekking op prestatie-motivatie en aspiratie-niveaus. De onderscheiden afhankelijke variabelen werden afzonderlijk in de analyse betrokken ter verhoging van de construct-validiteit (Bangert-Drowns, 1986).

4.3 Methodologische sluiscriteria

Alleen die studies werden in het onderzoek betrokken waarin expliciete vergelijkingen werden gemaakt tussen combinatieklassen en enkelvoudige klassen. De volgende sluiscriteria werden gehanteerd:

- 1 Alle studies dienden een proefopzet te hebben met een experimentele groep (een groep met combinatieklassen) en een controlegroep (een groep met enkelvoudige klassen).
- 2 Gebruik diende te worden gemaakt van gestandaardiseerde toetsen of meetinstrumenten. Rapportcijfers werden vanwege hun subjectieve karakter niet in het onder-

zoek betrokken. Niet-cognitieve uitkomsten die niet gebaseerd waren op gestandaardiseerde meetinstrumenten werden eveneens niet in beschouwing genomen.

- 3 Om te bereiken dat de onderzochte groepen bij de aanvang van het onderzoek ten aanzien van de relevante variabelen als aan elkaar gelijk konden worden beschouwd, diende gebruik gemaakt te zijn van matching van scholen of klassen, of van matching van individuele leerlingen binnen klassen of scholen. Studies die gebruik maakten van matching-procedures en gegevens overlegden waaruit bleek dat de groepen initieel vergelijkbaar waren, worden in deze studie aangeduid als 'matched studies met bewijs van initiële gelijkheid'. Studies waarin gebruik werd gemaakt van verschillen van covariantie-analyses om initiële verschillen tussen de experimentele groep en de controlegroep te verdisconten, worden in deze studie aangeduid als 'studies zonder bewijs van initiële gelijkheid met correcties voor verschillen op de voormeting'. De resultaten van deze studies moeten met enige voorzichtigheid worden geïnterpreteerd, omdat statistische correcties voor verschillen op de voormeting de invloed van deze verschillen op de nameting niet geheel onder controle houden (Reichardt, 1979). Studies waarin scholen of klassen gematched worden met gebruikmaking van een 'posttest-only design' (een statistische groepsvergelijking) worden in deze studie aangeduid als 'studies zonder bewijs van initiële gelijkheid zonder correcties voor verschillen op de voormeting'. Het ontbreken van een voormeting is een duidelijke tekortkoming van deze studies (Cook & Campbell, 1979). Hierdoor kan de mogelijkheid niet worden uitgesloten dat geconstateerde verschillen tussen de experimentele en controlegroep op de nameting verklaard kunnen worden door verschillen in de selectieprocedures voor beide groepen en niet door de experimentele behandeling (treatment). De resultaten van deze studies dienen met nog meer voorzichtigheid geïnterpreteerd te worden. Deze studies zijn voor de volledigheid opgenomen en worden steeds als laatste besproken.
- Representativiteit en gelijkheid van groepen

- kon ook worden bereikt door middel van een aselechte steekproeftrekking. Om tegemoet te komen aan de eis van generaliseerbaarheid van de gegevens en waarborging van de externe validiteit van het onderzoek, maakten deze studies gebruik van aselechte steekproeven uit de populatie van basisscholen met combinatieklassen en eenvoudige klassen. Ter controle op de interne validiteit met betrekking tot selectiefactoren (bijvoorbeeld schoolgrootte, geografische ligging, sociaal-economische achtergrond) werd veelal gebruik gemaakt van een of andere vorm van gestratificeerde steekproeftrekking. Deze studies met aselechte steekproeven zijn in een aparte categorie opgenomen.
- 4 De leerlingen dienden minstens één volledig schooljaar in een combinatieklas gezeten te hebben.
- 5 Alle studies dienden zich te richten op reguliere klassen met normale leerlingen.
- 6 De leerkrachten in de experimentele groep dienden op het moment van onderzoek geen speciale scholing gekregen te hebben die gericht was op de maximalisering van de leerprestaties. In deze studies waren trainingseffecten en groeperingseffecten met elkaar verstrengeld.
- 7 In alle studies dienden de experimentele groep uit ten minste twee combinatieklassen te bestaan en de controlegroep uit ten minste twee eenvoudige klassen. Dit criterium werd gehanteerd om in kleinschalige studies de effecten van de leerkracht en de klas op de leeruitkomsten te minimaliseren (vergl. Slavin, 1987, 1990).

Op grond van deze sluiscriteria werden 13 studies verwijderd (zie Veenman, 1996), waardoor er uiteindelijk 59 overbleven voor verdere analyse. Studies waarbij de leerlingen op een aselechte wijze aan de experimentele groep of controlegroep waren toegewezen (randomisatie-studies), werden niet gevonden. Gewerkt werd steeds met intacte scholen of klassen die na een bepaalde selectieprocedure werden toegewezen aan de experimentele of controlegroep.

Indien een studie tegemoet kwam aan bovenstaande criteria maar geen gegevens bevatte om een effectgrootte-maat te bereke-

nen, werd deze studie toch opgenomen. In de tekst wordt een indicatie gegeven van de richting en het statistisch verschil tussen de experimentele en de controlegroep.

4.4 Berekening van de effectgrootte-maat

In de literatuur over de leeruitkomsten van de combinatieklas wordt de enkelvoudige klas meestal opgevat als de controlegroep. Deze conventie wordt ook in deze studie aangehouden: positieve effecten zijn effecten ten gunste van de combinatieklas en negatieve effecten zijn effecten ten gunste van de enkelvoudige klas. Meestal werd de effectgrootte-maat (EG) berekend als het verschil tussen het gemiddelde van de experimentele groep en de controlegroep gedeeld door de standaarddeviatie van de controlegroep. Voor studies waarin de gegevens van de nameting via covariantie-analyses werden gecorrigeerd voor een of meer covariaten, werd het verschil in gecorrigeerde scores gedeeld door de ongecorrigeerde standaarddeviatie van de controlegroep. Op dezelfde wijze werd voor studies met verschillcores een effectgrootte-maat berekend door het verschil tussen de verschillcores van de experimentele groep en de controlegroep te delen door de ongecorrigeerde standaarddeviatie van de controlegroep op de nameting. Als noemer in de berekening voor de effectgrootte-maat werd genomen de ongecorrigeerde standaarddeviatie van de controlegroep op het niveau van de individuele leerling. Dit omdat het gebruik van geaggregeerde gegevens de effectgrootte-maat kunstmatig verhoogt (zie Slavin, 1987). Voor studies die gebruik maakten van klasgemiddelden (Adair, 1978; Brown & Martin, 1989; Carter, 1973; Her Majesty's Inspectorate, 1978) werd de standaarddeviatie voor de controlegroep op het niveau van de individuele leerling geschat door de standaarddeviatie van de controlegroep op het niveau van het klasgemiddelde te vermenigvuldigen met de wortel uit het gemiddeld aantal leerlingen per klas. In het geval de benodigde gemiddelden en standaarddeviaties ontbraken, werden de effectgroottematen geschat door gebruik te maken van de t -, F -, p -waarden of andere statistische gegevens volgens de door Glass, MacGaw en Smith (1981) en Slavin (1987) beschreven werkwijzen. Indien alleen de resultaten van de significantie-toetsen werden vermeld en de p -waar-

den werden afgerond tot .05 of .01 werden de gegeven waarden als exacte p -waarden beschouwd. Op deze wijze werd een behoudende schatting van de effectgrootte-maat verkregen.

Voor elke studie werd voor elke afhankelijke variabele een effectgrootte-maat berekend. Tevens werd voor elke studie een algemene effectgrootte-maat berekend: het gemiddelde van de afzonderlijke effectgrootte-maten. Voor bepaalde analyses werden de effectgroottematen voor verschillende studies samengenomen. In dit geval werden medianen berekend voor al die studies waarvoor een effectgrootte-maat berekend kon worden. Het aggregeren van effectgrootte-maten binnen goed omschreven categorieën van studies geeft volgens Slavin (1987) een bruikbare samenvatting van de omvang en de richting van de effecten. De gepoolde schatting dient echter wel geïnterpreteerd te worden vanuit de methodologische kwaliteit en de consistentie van de resultaten van de individuele studies die in de tekst beschreven worden.

Een effectgrootte-maat van .20 wordt beschouwd als een klein of minimaal effect, van .50 als een matig of gemiddeld effect en van .80 als een groot of betekenisvol effect (Cohen, 1969). Indien geen effectgrootte-maten berekend konden worden, worden de gevonden uitkomsten als volgt weergegeven. Een (+) geeft een gevonden verschil aan ten gunste van de combinatieklas, een (0) geeft aan dat er geen verschil is gevonden tussen de combinatie- en enkelvoudige klas en een (-) geeft een gevonden verschil aan ten gunste van de enkelvoudige klas. Deze werkwijze is ontleend aan de studie van Gutiérrez en Slavin (1992) met betrekking tot de effecten van de 'nongraded school'.

5 Het onderzoek over de combinatieklas

In de volgende paragrafen wordt een overzicht gegeven van de uitkomsten van het onderzoek met betrekking tot de leeruitkomsten in combinatieklassen. Eerst wordt aandacht besteed aan de leerprestaties van de leerlingen, vervolgens aan de niet-cognitieve leeruitkomsten. In elke paragraaf worden eerst de resultaten gepresenteerd van de matched studies met bewijs van

initiële gelijkheid, vervolgens van de studies die gebruik maakten van aselechte steekproeven. Daarna komen de resultaten van de studies zonder bewijs van initiële gelijkheid tussen de groepen met correcties voor verschillen op de voormeting en tenslotte de resultaten van de studies zonder bewijs van initiële gelijkheid tussen de groepen zonder correcties voor verschillen op de voormeting. De studies zijn per groep gerangschikt naar de grootte van de onderzoeksgroepen. In het algemeen geldt dat de studies die het eerst worden besproken een kwalitatief hoger methodologisch gehalte hebben dan de studies die later komen (cf. Gutiérrez & Slavin, 1992). Opgemerkt moet worden dat een aantal studies in de samenvattingen twee keer is opgenomen: één keer bij de cognitieve effecten als ze gegevens bevatten over de leerprestaties van de leerlingen en één keer bij de niet-cognitieve effecten als ze ook gegevens vermelden over het affectief functioneren van de leerlingen.

Vanwege plaatsgebrek wordt voor de uitgebreide tabellen, die bij een 'best-evidence'-synthese behoren, verwezen naar Veenman (1995, 1996b)¹. In dit artikel wordt per studie alleen de gemiddelde effectgrootte-maat vermeld. Tabel 1 geeft een beknopte samenvatting van de cognitieve en niet-cognitieve uitkomsten. Om redenen van internationale vergelijkbaarheid wordt – tenzij anders vermeld – de oude indeling in klassen aangehouden. Groep 8 in het Nederlandse basisonderwijs wordt in deze studie aangeduid als klas 6.

6 Combinatieklassen om redenen van beheersmatig/organisatorische motieven

6.1 Cognitieve effecten

In deze paragraaf worden hoofdzakelijk de resultaten van de studies besproken die zich richten op de leerprestaties van de leerlingen in combinatieklassen die om beheersmatig/organisatorische redenen zijn gevormd. Het betreft de resultaten van 40 studies. Studies met zowel cognitieve als niet-cognitieve uitkomsten worden eveneens gepresenteerd om overlap te voorkomen; zij komen niet meer voor in de paragraaf over de niet-cognitieve effecten.

Matched studies met initiële gelijkheid. In 9 studies waarin de leerprestaties van de leerlingen in combinatieklassen werden vergeleken met die van leerlingen in enkelvoudige klassen, werden gegevens gepresenteerd waaruit opgemaakt kon worden dat de groepen bij de aanvang van het onderzoek op een aantal relevante variabelen als aan elkaar gelijk konden worden beschouwd.

Rule (1983) onderzocht het effect van de combinatieklas op de lees- en rekenprestaties in klas 3 - 6 in Arizona. De combinatieklassen bestonden uit twee aanpalende jaargroepen. Uit de resultaten bleek dat de combinatieklas geen negatieve invloed had op de lees- en rekenprestaties van de leerlingen ($EG = +.01$).

Het onderzoek van Stone (1986) vond plaats in een groot stedelijk schooldistrict in de Verenigde Staten. Het richtte zich op de combinatieklas 2/3. Tussen de leerprestaties van leerlingen in combinatie- en enkelvoudige klassen werden geen significante verschillen gevonden ($EG = +.20$).

In de studie van Knight (1938) te Connecticut werden de leerprestaties van leerlingen in klas 4 onderzocht. Verschillen tussen combinatie- en enkelvoudige klassen werden niet gevonden ($EG = 0.00$). Over het algemeen onderwezen de leerkrachten de twee jaargroepen in de combinatieklassen bij rekenen en lezen apart.

In Nederland werden door Veenman e.a. (1986) gegevens verzameld over de lees- en rekenprestaties van leerlingen in klas 3 en 4. Tevens werden de lessen van leerkrachten in combinatie- en enkelvoudige klassen geobserveerd. Er werden geen significante verschillen in leerprestaties gevonden ($EG = -.02$). De observatiegegevens lieten zien dat de taakgerichte leertijd van leerlingen in combinatieklassen ongeveer 6% lager lag dan die van leerlingen in enkelvoudige klassen. Leerlingen in combinatieklassen spendeerden meer tijd aan individueel werken dan leerlingen in enkelvoudige klassen. In de meeste klassen werden de twee jaargroepen om beurten onderwezen (Veenman, Voeten & Lem, 1987).

Finley en Thompson (1963) onderzochten de verschillen in leerprestaties tussen combinatieklassen en enkelvoudige klassen in plattelandsgebieden in Californië. Het onderzoek richtte zich op klas 3 en 5. Er werden geen sig-

nificante verschillen gevonden in de leerprestaties met betrekking tot lezen, taal of rekenen ($EG = -.06$).

Lincoln (1981) ging de leesprestaties na van de leerlingen in klas 1 en 2 op een school in Connecticut. Hier werd al meer dan zeven jaar gecombineerd. De school kende een vorm van gedifferentieerd onderwijs. Tussen combinatie- en enkelvoudige klassen werden geen significante verschillen gevonden ($EG = +.04$).

Ook Adams (1953) vond geen verschillen in leerprestaties (lezen, taal en rekenen) tussen de vijfde-klas-leerlingen in combinatie- en enkelvoudige klassen in Californië ($EG = +.13$). De leerlingen zaten minstens twee jaar in een combinatieklas. Ook werden geen verschillen gevonden ten aanzien van de vriendschapskeuzen buiten de eigen klas en het aantal vriendjes in de klas ($EG = (0)$).

Furch-Krafft (1979) onderzocht de leerprestaties van leerlingen in combinatie- en enkelvoudige vierde klassen op het platteland rond Freiburg, Duitsland. De combinatieklassen bestonden uit twee of meer jaargroepen. Tussen beide groepen werden geen significante verschillen gevonden ($EG = -.18$).

Steinhofer (1980) volgde de leerprestaties van een groep kinderen vanaf de kleuterklas tot klas 1 in Californië. Of een kind geplaatst werd in een combinatieklas of in een enkelvoudige klas werd bepaald door de schoolleiding. Tussen de combinatie- en de enkelvoudige klassen werden geen significante verschillen gevonden ($EG = -.19$).

Studies met aselechte steekproeven. In 17 studies werd gebruikt gemaakt van aselechte nationale of lokale steekproeven. In 6 studies werden aanvankelijke verschillen ten aanzien van de leerprestaties op de voormeting in de analyses verdisconteerd, in 1 studie werden achtergrondvariabelen als covariaten gebruikt en in 3 studies werd een gehele populatie van combinatieklas-leerlingen uit een bepaalde staat/provincie of district vergeleken met een steekproef van leerlingen uit enkelvoudige klassen.

Aan de hand van een representatieve steekproef van Engelse basisscholen onderzocht de Inspectie (Her Majesty's Inspectorate, 1978) de lees- en rekenprestaties van de leerlingen. De meeste combinatieklassen bestonden uit twee of drie combinaties. De leerlingen van 11 jaar oud in de enkelvoudige klassen scoorden

hogerop de reken- en leestoetsen dan de leerlingen van 11 jaar oud in de combinatieklassen. Ook de leesprestaties van de 9-jarige leerlingen in de enkelvoudige klassen (rekenprestaties werden hier niet vastgesteld) waren beter dan die van de 9-jarige leerlingen in de combinatieklassen ($EG = -.26$).

Voor drie achtereenvolgende jaren onderzocht Jokinen (1979) de leerprestaties van leerlingen met betrekking tot een nationale steekproef van Finse basisscholen. In kleine scholen op het platteland bevatte een combinatieklas twee of meer jaargroepen. De resultaten van deze studie wezen uit dat de prestaties van de leerlingen voor lezen, rekenen, taal, Engels en Zweeds in combinatie- en enkelvoudige klassen niet van elkaar verschilden ($EG = +.03$).

Jarousse en Mingat (1991a,b, 1992) onderzochten de leerprestaties van leerlingen in de leerjaren 2 en 5 in Togo en in leerjaar 2 in Burkina Faso voor de vakken taal en rekenen. In beide landen bleek het onderwijs in combinatieklassen de leerprestaties van de leerlingen positief te beïnvloeden. Een effectgrootte van +.50 werd gevonden voor leerjaar 2 (in Togo en in Burkina Faso) en een effectgrootte van +.33 voor leerjaar 5 (in Togo). De leerlingen zaten in de combinaties 1/2 of 5/6. In Togo varieerde het aantal leerlingen in de onderzochte combinatie- of enkelvoudige klassen van 25 tot 130 en in Burkina Faso van 23 tot 150. Voor klassen met meer dan 55 leerlingen bleek de grootte van de combinatieklas de leerprestaties van de leerlingen minder negatief te beïnvloeden dan de grootte van een enkelvoudige klas.

Gajadharsingh (1987) onderzocht de leerprestaties van leerlingen in combinatie- en enkelvoudige klassen in Saskatchewan (Canada). Alleen die leerlingen werden in het onderzoek betrokken die in de onderzochte leerjaren (3-6) vanaf de eerste klas steeds in een combinatie- of enkelvoudige klas hadden gezeten. De leerlingen in de combinatieklassen scoorden over het algemeen significant hoger op de schooltoetsen voor taal, lezen en rekenen dan de leerlingen in de enkelvoudige klassen ($EG = +.08$).

In het onderzoek van Brandsma (1993) werden de leerprestaties van leerlingen onderzocht bij een aselechte steekproef van Nederlandse basisscholen. Gebruik werd gemaakt van een multi-niveau-analyse om de leerprestaties van

de leerlingen te analyseren op het niveau van de school, klas en leerling. De analyse richtte zich op klas 6 van de basisschool (groep 8). Over de samenstelling van de combinatieklassen werd geen informatie verstrekt. Ten behoeve van dit overzichtsartikel werden de gegevens opnieuw geanalyseerd via een covariantie-analyse met als covariaten de scores van de leerprestaties op de voormeting, het IQ en het sociaal-economisch milieu van de leerlingen². Significante maar kleine verschillen werden gevonden voor lezen ($EG = -.08$) en taal ($EG = -.09$) ten gunste van de leerlingen in de enkelvoudige klassen. Voor rekenen werden geen significante verschillen gevonden ($EG = -.04$). De gemiddelde effectgrootte bedroeg $-.07$.

Ten behoeve van dit artikel voerde Doolgaard (1996) een secundaire analyse uit op de gegevens van klas 4 (groep 6) die in het schooljaar 1987 en 1988 waren verzameld. Deze gegevens, aanwezig in de databestanden van Brandsma (1993) en Knuver (1993), werden nog niet eerder geanalyseerd vanuit de vraag naar de cognitieve en niet-cognitieve leeruitkomsten in combinatieklassen. Gegevens over de samenstelling van de combinatieklassen werden niet vermeld. Uitgevoerd werd een covariantie-analyse met als covariaten de scores van de voormeting, het IQ en het sociaal-economisch milieu van de leerlingen (bij de attitudes t.o.v. de school ontbrak een voormeting). Voor rekenen werden significante maar kleine verschillen gevonden ten gunste van de leerlingen in de enkelvoudige klas ($EG = -.15$). Voor lezen ($EG = -.00$) en taal ($EG = +.01$) werden geen verschillen gevonden. De gemiddelde effectgrootte voor de leerprestaties bedroeg $-.05$. Met betrekking tot de niet-cognitieve effecten werd ten gunste van de enkelvoudige klas een significant maar klein verschil gevonden voor de attitudes t.o.v. de school ($EG = -.08$). Voor de attitudes t.o.v. taal en rekenen en voor de prestatiemotivatie van de leerlingen werden geen verschillen gevonden.

Rojas en Castillo (1988) en Psacharopoulos, Rojas en Velez (1993) onderzochten de leerprestaties van leerlingen in klas 3 en 5 van de Escuela Nueva in Colombia. Ook werd het zelfbeeld van de leerlingen onderzocht. De Escuela Nueva is een plattelandsschool met

een of twee leerkrachten. Traditionele scholen bestaan uit enkelvoudige klassen. Uit de leerprestaties van de leerlingen bleek dat de leerlingen van de Escuela Nueva hoger scoorden dan de leerlingen van de traditionele scholen, met uitzondering van het vak rekenen in klas 5 ($EG = +.22$). Voor zelfbeeld werden geen significante verschillen gevonden ($EG = +.08$).

Spratt (1986) vergeleek de leesprestaties van de leerlingen in combinatieklassen (klas 1-6) in Virginia. Elke combinatieklas telde twee jaargroepen. In de combinatieklassen stonden naar het oordeel van de schoolleider de beste leerkrachten. Tevens zaten hier de betere leerlingen. Significante verschillen in leesprestaties tussen combinatie- en enkelvoudige klassen werden niet gevonden ($EG = +.28$).

Marklund (1962) onderzocht de leerprestaties van leerlingen uit klas 6 in kleine en grote scholen in zes Zweedse provincies. Alle scholen waren gesitueerd in kleine steden of plattelandsgemeenten. De effectgrootte voor zes schoolvakken bedroeg $-.18$. De resultaten lieten zien dat de leerlingen in combinatieklassen met niet meer dan 25 leerlingen evengoed presteerden als de leerlingen in enkelvoudige klassen. Met méér dan 25 leerlingen in een combinatieklas presteerden de leerlingen in de combinatieklassen significant lager dan de leerlingen in de enkelvoudige klassen ($EG = -.26$).

Kral (1995) bestudeerde de effecten van combinatie- versus enkelvoudige klassen voor rekenen, taal en lezen van leerlingen uit klas 2 (groep 4), 4 (groep 6) en 6 (groep 8) in Nederland. Centraal stonden de leerprestaties van leerlingen in kleine scholen (met minder dan 110 leerlingen) versus de leerprestaties van leerlingen in grote scholen (met meer dan 250 leerlingen). In de kleine scholen (in de steden en op het platteland) werd het onderwijs gegeven in combinatieklassen, in de grote scholen in enkelvoudige klassen. Gebruik werd gemaakt van multi-niveau-analyses. Ten behoeve van dit overzichtsartikel werden de gegevens geanalyseerd via een covariantie-analyse met als covariaten de scores van de leerprestaties op de voormeting, het IQ en het sociaal-economisch milieu van de leerlingen². Tussen de combinatie- en de enkelvoudige klassen werden geen significante verschillen gevonden ($EG = -.06$). Leerlingen die hun gehele school-

loopbaan in een combinatieklas hadden gezeten, presteerden niet minder dan leerlingen die steeds in een enkelvoudige klas hadden gezeten. Gegevens verzameld met behulp van de logboekmethode lieten zien dat er nauwelijks verschillen bestonden in de gemiddelde hoeveelheid instructietijd tussen enkelvoudige en combinatieklassen. Ook werden tussen deze klassen geen belangrijke verschillen gevonden met betrekking tot de organisatietijd tijdens de lessen. Werken in subgroepen kwam in beide klassen nauwelijks voor. De jaarklassikale situatie werd in combinatieklas voortgezet (Kral, Voeten & Nederlof, 1995).

Dreier (1949) vond in Minnesota geen verschillen in de leerprestaties van leerlingen in de eenklassige plattelandsschool en in de plattelandsschool met enkelvoudige klassen. De leerlingen in klas 6 van deze twee typen scholen presteerden even goed ($EG = (0)$).

Fippinger (1967) onderzocht de leerprestaties van leerlingen in combinatie- en enkelvoudige klassen bij een representatieve steekproef van leerlingen uit klas 4 in Duitsland. De leerlingen in de enkelvoudige klassen scoorden significant beter dan de leerlingen in de combinatieklassen ($EG = -.44$). Volgens Fippinger kwam dit doordat de enkelvoudige klassen bevolkt werden door intelligentere leerlingen. Deze enkelvoudige klassen waren gesitueerd in dichtbevolkte gebieden, terwijl de combinatieklassen gesitueerd waren in dunbevolkte gebieden.

Nieminen (1979) voerde een secundaire analyse uit op de Finse gegevens die verzameld waren in het evaluatie-project van de International Association for the Evaluation of Educational Achievement (IEA). In een landelijke representatieve steekproef waren gegevens beschikbaar van leerlingen uit klas 3 en 4. Tussen de leerlingen uit de combinatie- en de enkelvoudige klassen werden geen significante verschillen ten aanzien van de leerprestaties (lezen, taal, natuurkunde) gevonden ($EG = +.02$). Ook werden geen verschillen gevonden voor de attitudes t.o.v. de school ($EG = -.10$) en de motivatie van de leerlingen ($EG = -.09$).

Rowley (1992) onderzocht het effect van combinatieklassen in Pakistan voor rekenen en natuurkunde. In de combinatieklassen zaten twee of meer jaargroepen. Een significant verschil in leerprestaties werd gevonden ten gun-

ste van de leerlingen uit de enkelvoudige klassen ($EG = -.36$). Volgens Rowley werden de enkelvoudige klassen in meerderheid onderwezen door de beter opgeleide leerkrachten.

Harvey (1974) onderzocht het effect van de combinatieklas op de leerprestaties, het zelfbeeld en de sociaal-emotionele ontwikkeling van kleuters in Virginia. De combinatieklassen bestonden uit kleuters en leerlingen uit klas 1. Alle scholen lagen op het platteland. Tussen de leerlingen uit de combinatie- en de enkelvoudige klassen werden geen significante verschillen in leerprestaties gevonden ($EG = -.07$). De kleuters in de combinatieklassen scoorden ten aanzien van het zelfbeeld hoger dan de kleuters in de enkelvoudige klassen, maar niet ten aanzien van sociale aanpassing ($EG = +.44$).

Zabolotney (1983) vergeleek de leesprestaties en school-attitudes van leerlingen uit klas 4 in Arkansas. Het onderzoek betrof alleen plattelandsscholen van de zevendedagadventisten. Tussen de combinatie- en de enkelvoudige klassen werden geen verschillen gevonden voor de leesprestaties ($EG = -.14$) en de attitudes t.o.v. de school ($EG = -.39$).

Matched studies zonder bewijs van initiële gelijkheid. In 14 studies werden combinatieklassen (of leerlingen uit combinatieklassen) gematched met enkelvoudige klassen (of met leerlingen uit enkelvoudige klassen) zonder dat deze studies gegevens bevatten die duidelijk maakten dat de controle- en experimentele groepen bij de aanvang van het onderzoek gelijkwaardig waren. In 7 van deze studies werd op statistische wijze gecontroleerd voor initiële verschillen in leerprestaties op de voormeting. In de resterende 7 studies namen de onderzoekers zonder meer aan dat de experimentele en controlegroepen gelijkwaardig waren.

Stimson (1991) bestudeerde de effecten van het onderwijs in combinatieklassen (klas 3-6) op 9 scholen in Californië waar het gehele jaar door onderwijs werd gegeven (year-round multi-track education). In een bepaalde onderwijsperiode ging een gedeelte van de leerlingen naar school, terwijl een andere gedeelte op vakantie was. Stimson vond dat de leerlingen in de enkelvoudige klassen significant meer presteerden voor lezen, rekenen en taal dan de leerlingen de combinatieklassen ($EG = -.08$). Volgens de onderzoekster scoorden scholen

met een doorlopende onderwijkskalender consistent lager dan scholen met een normale onderwijsperiode.

Galton en Simon (1980) bestudeerden de invloed van combinatieklassen in Leicester (Engeland). De volgende combinaties werden onderzocht: klassen met leerlingen van 7-8 jaar oud, 8-9 jaar oud en 9-10 jaar oud. Met betrekking tot de leerprestaties (lezen, rekenen en taal) werden geen significante verschillen gevonden tussen de combinatie- en enkelvoudige klassen ($EG = (0)$). Observaties wezen uit dat de leerlingen in combinatieklassen iets minder taakgericht waren en dat meer tijd besteed werd aan de organisatie van het onderwijs.

Adair (1978) onderzocht de cognitieve en niet-cognitieve effecten van het onderwijs in combinatieklassen bij kleuters en leerlingen uit klas 1 in het noordoosten van de VS. Een aantal schoolleiders wees voor het onderzoek de 'beste combinatieklassen' aan. Voor de leerlingen uit klas 1 werden geen significante verschillen geconstateerd tussen de combinatie- en de enkelvoudige klassen ($EG = -.06$). In het eerste leerjaar stonden de leerlingen in de enkelvoudige klassen meer positief t.o.v. de school dan de leerlingen in de combinatieklassen. Voor de kleuters werd in dit opzicht geen verschil gevonden ($EG = -.16$).

In een onderzoek van Pawluk (1992) werden de leerprestaties van leerlingen uit combinatie- en enkelvoudige klassen (5-8) uit de staten Oregon en Washington met elkaar vergeleken. De leerlingen uit de groep met combinatieklassen bezochten scholen van de zeventiende- en achttiende-eeuwse met één of twee leerkrachten. Tussen de combinatie- en de enkelvoudige klassen werden geen significante verschillen gevonden ($EG = -.25$).

Higgins (1980) vergeleek de leesprestaties van leerlingen in combinatie- en enkelvoudige klassen (klas 3-5) op 3 scholen in Louisiana. Het onderwijs in de combinatieklas was in hoge mate geïndividualiseerd. Tussen de leerlingen uit de combinatie- en de enkelvoudige klassen werden geen significante verschillen in leesprestaties gevonden ($EG = +.05$).

Rehwoldt en Hamilton (1957) gingen het effect na van pas gevormde combinatieklassen in één school in Californië. De combinatieklassen bestonden uit 3 jaargroepen (klas 1-3 en

klas 4-6). De ouders beslisten of hun kinderen in een combinatieklas geplaatst werden. Tussen de combinatie- en de enkelvoudige klassen werden geen verschillen in leerprestaties (lezen, rekenen en taal) gevonden ($EG = +.15$). Voor de school-attitudes werd een niet-statistisch verschil gevonden ten gunste van de combinatieklassen ($EG = +.32$).

In de studie van Martens (1954) bezochten de leerlingen uit Iowa een eenklassige plattelandsschool of een school in een stad met enkelvoudige klassen. De leerlingen op de stadsscholen presteerden beter voor lezen, rekenen en taal dan de leerlingen op de plattelandsscholen ($EG = -.61$). Opgemerkt moet worden dat de IQ-scores van de leerlingen uit de stad significant hoger lagen dan die van de leerlingen op het platteland.

Råberg (1976) onderzocht de leerprestaties van leerlingen in combinatie- en enkelvoudige klassen (klas 3 en 6) in Västerbotten (Zweden). Tussen de combinatie- en de enkelvoudige klassen werden geen significante verschillen gevonden ($EG = +.19$).

Brown en Martin (1989) analyseerden de schoolprestaties van leerlingen uit klas 4 in New Brunswick, Canada. Volgens de leerkrachten waren de leerlingen voor plaatsing in de combinatieklassen geselecteerd op criteria als zelfredzaamheid, samenwerkingsbereidheid en zelfstandigheid. Tussen de combinatie- en enkelvoudige klassen werden geen significante verschillen gevonden ($EG = +.12$).

Way (1981) onderzocht de leerprestaties van leerlingen in combinatieklassen in New York (klas 1-5). In de combinatieklassen zaten leerlingen uit twee of drie jaargroepen. Voor de leerprestaties werden geen significante verschillen gevonden ($EG = +.01$). Ook werden geen significante verschillen gevonden voor het zelfbeeld van de leerlingen ($EG = +.17$).

Knörzer (1985) bestudeerde de leeruitkomsten van leerlingen in combinatieklassen (klas 3 en 4) van kleine scholen op het platteland in de Württemberg, Duitsland. Voor de leerprestaties (lezen, rekenen en taal) werden geen significante verschillen gevonden ($EG = (0)$). Ook niet voor de niet-cognitieve variabelen ($EG = +.28$).

Purl en Curtis (1970) onderzochten de leeruitkomsten van leerlingen in combinatieklassen in Californië. Voor lezen en zelfbeeld wer-

den geen verschillen tussen combinatie- en enkelvoudige klassen gevonden ($EG = (0)$). De leerlingen in de combinatieklas scoorden significant hoger met betrekking tot persoonlijke en sociale aanpassing ($EG = (+)$).

Chace (1961) bestudeerde de leeruitkomsten van leerlingen in klas 3-6 van twee plattelandscholen in Tennessee. Tussen de leerlingen uit de combinatie- en de enkelvoudige klassen werden geen significante verschillen gevonden ten aanzien van de leerprestaties voor lezen, rekenen en taal ($EG = +.20$). Op 5 van de 8 subtests van de California Test of Personality scoorden de leerlingen uit de combinatieklassen significant hoger ($EG = +.39$).

Dodendorf (1983) vergeleek de leerprestaties van leerlingen in een tweeklassige plattelandschool in Nebraska met die van leerlingen in enkelvoudige stadsscholen. De leerlingen in de enkelvoudige klassen presteerden beter voor de zaakvakken. Voor rekenen, taal en de exacte vakken werden geen verschillen tussen leerlingen in combinatie- en enkelvoudige klassen gevonden ($EG = -.24$).

Samenvatting van de cognitieve uitkomsten. Geconstateerd kan worden dat de resultaten van het onderzoek naar de leerprestaties van de leerlingen in de combinatieklassen in vergelijking met die van de leerlingen in de enkelvoudige klassen een tamelijk consistent beeld vertonen. De medianen voor de effectgroottematen voor de drie onderscheiden groepen van studies waren als volgt: (1) $EG = .00$ voor de 9 kwalitatief goede studies met bewijs van initiële gelijkheid, (2) $EG = -.05$ voor de 16 kwalitatief goede studies die gebruik maakten van aselechte steekproeven en (3) $EG = +.01$ voor de 11 studies zonder bewijs van initiële gelijkheid. Voor al deze 36 studies waarvoor effectgroottematen berekend konden worden, bedroeg de mediaan $-.01$ (gemiddelde = $-.02$). Voor 4 studies kon geen effectgrootte-maat berekend worden. In 30 van de 40 studies werden tussen de combinatie- en de enkelvoudige klassen geen significante verschillen in leerprestaties gevonden. De resterende studies gaven een verdeeld beeld te zien. De bovenste helft van Tabel 1 geeft een beknopte samenvatting van de belangrijkste resultaten. Gesteld kan worden dat de leerlingen in de combinatieklassen niet beter maar ook niet slechter leren

dan de leerlingen in de enkelvoudige klassen.

6.2 Niet-cognitieve effecten

In deze paragraaf worden de resultaten vermeld van 19 (deel)studies met niet-cognitieve effecten van combinatie- versus enkelvoudige klassen. Zoals gezegd zijn de studies met zowel cognitieve als niet-cognitieve leeruitkomsten in de vorige paragraaf beschreven en worden hier niet apart vermeld. Ze zijn wel opgenomen in de samenvatting.

Matched studies met initiële gelijkheid. In deze categorie vallen 4 studies. De studie van Adams (1953) is reeds in de vorige paragraaf besproken. In de 3 resterende studies werden combinatieklassen gematched met enkelvoudige klassen en werd duidelijk gemaakt dat deze matching leidde tot gelijkheid van de onderzochte groepen.

Carter (1973) bestudeerde de effecten van het onderwijs in combinatieklassen (klas 3 en 5) op het zelfbeeld en de school-attitudes in Michigan. De leerlingen hadden steeds in een combinatieklas gezeten. De leerlingen in de combinatieklassen behaalden significant hogere scores dan de leerlingen in de enkelvoudige klassen ($EG = +.38$).

Junell (1971) onderzocht de leerprestaties van 'junior-high-school'-leerlingen in de staat Washington die voorheen basisscholen hadden bezocht met combinatie- of enkelvoudige klassen. Met betrekking tot zelfbeeld, attitudes t.o.v. de school en persoonlijkheidsontwikkeling werden geen significante verschillen gevonden ($EG = +.10$).

Grove (1978) bestudeerde de effecten van het onderwijs in combinatie- en enkelvoudige klassen (4-6) ten aanzien van het zelfbeeld en de school-attitudes van de leerlingen. De scholen bevonden zich op het platteland van Idaho, Washington en Oregon. De leerkrachten waren door hun schoolleiders beoordeeld als 'excellent'. Tussen de klassen werden geen significante verschillen gevonden ($EG = +.03$).

Studies met aselechte steekproeven. In deze categorie vallen 6 (deel)studies; 5 van deze studies zijn reeds in de vorige paragraaf besproken. Knuver (1993) beperkte haar studie tot alleen de niet-cognitieve uitkomsten. Centraal in deze studie stond de vraag naar de relatie tussen klas- en schoolkenmerken en het affectief functioneren van leerlingen bij een aselechte

steekproef van Nederlandse basisscholen. Zij maakte gebruik van hetzelfde gegevensbestand dat Brandsma (1993) analyseerde met het oog op het cognitief functioneren van leerlingen in klas 6 (groep 8). Bij het analyseren van de leerling-, klas- en schoolgegevens werd een multi-niveau-analyse toegepast. Gegevens over combinatieklassen werden niet vermeld. Ten behoeve van dit overzichtartikel werden de gegevens opnieuw geanalyseerd via een covariantie-analyse met als covariaten de scores van de niet-cognitieve prestaties op de voormeting, het IQ en het sociaal-economisch milieu van de leerlingen². Tussen de leerlingen in de combinatie- en de enkelvoudige klassen werden geen significante verschillen gevonden voor de attitudes t.o.v. de school en de prestatie-motivatie van de leerlingen ($EG = .00$). Ook werden geen verschillen gevonden voor de attitudes t.o.v. taal ($EG = +.04$) en rekenen ($EG = +.02$).

Matched studies zonder bewijs van initiële gelijkheid. In deze categorie vallen 9 (deel)studies waarvan er reeds 6 zijn besproken in de vorige paragraaf. In de 3 resterende studies zonder bewijs van initiële gelijkheid (en zonder correcties voor verschillen op de voormeting) werden geen significante verschillen tussen combinatieklassen en enkelvoudige klassen gevonden. In een studie die werd uitgevoerd in West-Australië bestudeerden Pratt, Holub en Ainsworth (1989) de percepties van de leerlingen (klas 5 en 6) ten aanzien van de school. Tussen de leerlingen in de combinatie- en de enkelvoudige klassen werden geen significante verschillen gevonden ($EG = (0)$).

Pratt en Treacy (1986) vergeleken de effecten van het onderwijs in combinatie- en enkelvoudige klassen (klas 1/2) ten aanzien van de school-attitudes, eveneens in West-Australië. Tussen de combinatie- en de enkelvoudige klassen werden geen significante verschillen gevonden ($EG = (0)$). In de combinatieklassen werkten de leerkrachten meestal met de gehele klas of met één bepaalde jaargroep in de klas.

Hoer (1972) onderzocht de invloed van het onderwijs in 3 combinatieklassen in één school in British Columbia, Canada. Zijn studie richtte zich op het zelfbeeld, de motivatie en de persoonlijke aanpassing van de leerlingen in klas 5. De combinatieklassen bestonden uit 3 jaargroepen. Tussen de klassen werden geen significante verschillen gevonden ($EG = -.10$).

Samenvatting van de niet-cognitieve uitkomsten. In het algemeen zijn de uitkomsten voor de niet-cognitieve effecten van de combinatie-versus enkelvoudige klassen gelijk aan de uitkomsten voor de cognitieve effecten of de leerprestaties. De leerlingen in de combinatieklassen presteerden ten aanzien van de onderzochte niet-cognitieve uitkomsten niet beter maar ook niet slechter dan de leerlingen in de enkelvoudige klassen. De gevonden medianen voor de effectgrootte voor de drie groepen van onderscheiden studies waren als volgt: (1) $EG = +.10$ voor de 3 kwalitatief goede studies met bewijs van initiële gelijkheid, (2) $EG = -.02$ voor de 6 kwalitatief goede studies die gebruik maakten van aselechte steekproeven en (3) $EG = +.22$ voor de 6 studies zonder bewijs van initiële gelijkheid. Voor al deze 15 studies waarvoor een effectgrootte-maat berekend kon worden, bedroeg de mediane effectgrootte $+.08$ (gemiddelde = $+.09$). Voor 4 studies kon geen effectgrootte-maat berekend worden. In 14 van de 19 studies met betrekking tot de niet-cognitieve effecten werden geen significante verschillen gevonden; in de resterende 5 studies werd een significant verschil gevonden ten gunste van de combinatieklas. Deze uitkomst suggereert dat op de meer affectieve gebieden (zoals de attitudes t.o.v. de school, het zelfbeeld en de persoonlijke en sociale aanpassing) de leerlingen in combinatieklassen soms hogere scores behalen. Deze positieve effecten zijn echter zo klein dat ze niet resulteren in een cognitieve leerwinst. De bovenste helft van Tabel 1 geeft een geknopte samenvatting van de gevonden uitkomsten.

7 Combinatieklassen om reden van pedagogisch/didactische motieven

7.1 Cognitieve effecten

In deze paragraaf worden de resultaten besproken van de studies die zich richten op combinatieklassen die om reden van pedagogisch/didactische motieven zijn gevormd. In deze categorie vallen 11 studies. Studies met zowel cognitieve als niet-cognitieve uitkomsten worden eveneens gepresenteerd; zij worden niet meer aan de orde gesteld in de paragraaf over

de niet-cognitieve effecten. Studies met aselechte steekproeven komen voor deze groep van studies niet voor.

Matched studies met initiële gelijkheid. Slechts in 2 van de 11 studies over combinatieklassen die vanwege pedagogisch/didactische motieven werden samengesteld werden gegevens gepresenteerd over de gelijkheid van de experimentele en de controlegroepen bij de aanvang van het onderzoek. Mobley (1976) onderzocht de leerprestaties van leerlingen in combinatie- en enkelvoudige groepen (klas 2-3) in één school in Georgia. In alle klassen werden dezelfde methodes gebruikt. Tussen de klassen werden geen significante verschillen gevonden ($EG = .00$). Met betrekking tot het zelfbeeld van de leerlingen (klas 1-3) werden eveneens geen significante verschillen geconstateerd ($EG = +.18$).

Givens (1972) vergeleek de leerprestaties en de attitudes van leerlingen uit klas 5 van twee scholen in Missouri. De school met combinatieklassen maakte gebruik van vormen van team-teaching en geïndividualiseerd onderwijs en bezat een open architectuur. Bij toelating werden de leerlingen aan een toelatingsonderzoek onderworpen. Tussen de scholen werden geen significante verschillen gevonden voor lezen, taal en rekenen ($EG = +.06$). Ook niet tussen de meer en minder begaafde leerlingen. Voor de attitudes t.o.v. de school werd een significant effect gevonden ten gunste van de combinatieklassen ($EG = +.58$).

Matched studies zonder bewijs van initiële gelijkheid. In 9 studies werden combinatie- en enkelvoudige klassen gematched waarbij geen gegevens werden gepresenteerd over de initiële gelijkheid van de groepen. In 5 van deze studies werd op statistische wijze gecontroleerd voor aanvangsverschillen in leerprestaties op de voormeting.

Schrankler (1976) onderzocht op één school in Minnesota het effect van de combinatieklas op de lees- en rekenprestaties, het zelfbeeld en de school-attitudes (klas 1-6). Over het algemeen werden geen systematische verschillen in leerprestaties geconstateerd ($EG = (0)$). Met betrekking tot de niet-cognitieve effecten werden significante verschillen gevonden ten gunste van de combinatieklas ($EG = (+)$).

Yerry en Henderson (1964) bestudeerden de effecten van combinatieklassen op de leerpre-

taties en de niet-cognitieve variabelen als angst en vriendschapskeuzen op twee scholen in de staat New York (klas 1-6). De leerlingen in de combinatieklassen ontvingen geïndividualiseerd onderwijs of onderwijs in kleine groepen. Voor lezen, rekenen en taal werden geen significante verschillen gevonden ($EG = -.01$). Ook werden geen significante verschillen gevonden voor de niet-cognitieve variabelen ($EG = -.07$).

Burchyett (1972) stelde een onderzoek in naar de leerprestaties, de motivatie en het zelfbeeld van leerlingen in Michigan (klas 3-5). In de combinatieklassen werd gewerkt met team-teaching. Voor de leerprestaties (lezen, rekenen, zaakvakken) werden geen verschillen gevonden tussen de combinatie- en de enkelvoudige klassen ($EG = -.05$). Ook werden geen verschillen gevonden voor het zelfbeeld en de motivatie van de leerlingen ($EG = -.08$).

Vogel en Bowers (1972) analyseerden de effecten van combinatieklassen in Illinois voor de leerprestaties, de school-attitudes, persoonlijke en sociale aanpassing van de leerlingen (kleuters tot en met klas 6). De leerlingen in de enkelvoudige klassen bleken op een algemene schooltoets hoger te scoren dan de leerlingen in de combinatieklassen ($EG = -.24$). Voor de niet-cognitieve variabelen werden geen systematische verschillen gevonden ($EG = +.06$).

In een studie die door Marsh (1980) in Florida werd uitgevoerd werd onderzocht of leerlingen van een middenschool met combinatieklassen (klas 6-8) betere leerprestaties behaalden dan leerlingen van een middenschool met enkelvoudige klassen. Op de voormeting werden tussen beide groepen significante verschillen geconstateerd. Verschillen op de nameting werden getoetst via covariantie-analyse. Deze toetsing leverde significante verschillen op ten gunste van de leerlingen in de enkelvoudige klassen ($EG = -.43$). Gezien de verschillen tussen beide scholen dienen de resultaten van deze studie met grote voorzichtigheid te worden geïnterpreteerd.

Milburn (1981) bestudeerde gedurende 5 jaar het effect van combinatieklassen (klas 1-6) op één school in British Columbia (Canada). Het onderwijs was afgestemd op het individuele niveau van de leerlingen. De leerlingen bleven ongeveer 3 jaar bij dezelfde leerkracht. Verschillen in leerprestaties (lezen, rekenen) werden niet gevonden ($EG = (0)$). De leerlin-

gen in de combinatieklassen hadden een meer positieve houding t.o.v. de school en hogere zelfbeeld-scores dan de leerlingen in de enkelvoudige klassen ($EG = (+)$).

Schroeder en Nott (1974) evalueerden de effecten van combinatieklassen (klas 1-5) op één school in Cincinnati (Ohio). In klassen werd geïndividualiseerd onderwijs gegeven. Alleen voor de leerlingen in de combinatieklas 4/5 werden hogere leerprestaties gevonden ($EG = (+)$). De leerlingen in de combinatieklassen hadden verder een meer positieve attitude t.o.v. de school ($EG = (+)$).

Papay, Costello, Hedl en Spielberg (1975) analyseerden de relatie tussen angst en de rekenprestaties van risico-leerlingen in klas 1 en 2 in Dallas. Het onderwijs in de combinatieklassen was geïndividualiseerd. Voor de rekenprestaties werden geen significante verschillen gevonden ($EG = -.10$). De leerlingen in leerjaar 2 in de combinatieklassen waren minder angstig; voor leerjaar 1 werden geen significante verschillen gevonden ($EG = +.20$).

Mycock (1966) onderzocht het verschil tussen combinatie- en enkelvoudige kleuterklassen in Manchester. Alleen bij de 7-jarige kinderen werden schooltoetsen afgenomen. Voor lezen en rekenen werden geen significante verschillen gevonden ($EG = +.13$). Ook werden geen significante verschillen gevonden voor persoonlijke en sociale aanpassing ($EG = +.12$).

Samenvatting van de cognitieve uitkomsten. De resultaten van deze groep van studies laten nauwelijks verschillen zien tussen de combinatie- en de enkelvoudige klassen. In de meeste studies worden geen significante verschillen gevonden. De medianen voor de effectgroottematen voor de twee onderscheiden groepen van studies bedroegen: (1) $EG = +.03$ voor de 2 kwalitatief goede studies met bewijs van initiële gelijkheid en (2) $EG = -.07$ voor de 6 studies zonder bewijs van initiële gelijkheid. Voor al deze 8 studies waarvoor effectgroottematen berekend konden worden, bedroeg de mediaan $-.03$ (gemiddelde = $-.08$). Voor 3 studies konden geen effectgroottematen worden berekend. In 9 van de 11 studies werden geen significante verschillen gevonden tussen de combinatie- en enkelvoudige klassen. Over het algemeen blijken de leerlingen in de combina-

tieklassen die om pedagogisch/didactische redenen zijn gevormd evenveel te presteren als de leerlingen in de enkelvoudige klassen. De onderste helft van Tabel 1 geeft een beknopte samenvatting van de gevonden resultaten.

7.2 Niet-cognitieve effecten

In deze categorie vallen 11 (deel)studies: 2 studies met bewijs en 9 studies zonder bewijs van initiële gelijkheid. Studies met aselechte steekproeven komen in deze categorie niet voor. Zoals gezegd zijn de resultaten van studies met zowel cognitieve als niet-cognitieve leeruitkomsten in de vorige paragraaf vermeld. Eén studie is nog niet besproken: de studie van Hammack (1974) beperkte zich tot alleen de niet-cognitieve effecten van combinatieklassen die om pedagogisch/didactische motieven werden gevormd.

Matched studies zonder bewijs van initiële gelijkheid. Hammack (1974) onderzocht het zelfbeeld van leerlingen (3-5 jaar) in combinatie- en enkelvoudige klassen in Texas. In alle klassen werd gewerkt met leerhoeken en met dezelfde methoden. In de combinatieklassen lagen de scores voor het zelfbeeld van de leerlingen significant hoger dan in de enkelvoudige klassen ($EG = +.33$).

Samenvatting van de niet-cognitieve uitkomsten. De gevonden medianen voor de effectgroottematen voor de twee groepen van onderscheiden studies waren als volgt: (1) $EG = +.38$ voor de 2 kwalitatief goede studies met bewijs van initiële gelijkheid en (2) $EG = +.09$ voor de 6 studies zonder bewijs van initiële gelijkheid. De resultaten uit de eerste groep van studies worden voornamelijk bepaald door de uitkomsten van de studie van Givens (1972) die een effectgrootte vond van $+.58$ ten gunste van de leerlingen in gecombineerde groepen. Voor alle 8 studies waarvoor een effectgroottemaat berekend kon worden bedroeg de mediaan $+.15$ (gemiddelde = $+.18$). Voor 3 studies kon geen effectgroottemaat worden berekend. In 5 van de 11 studies werden geen significante verschillen gevonden tussen de combinatie- en de enkelvoudige klassen, in de resterende 6 studies werd een significant verschil gevonden ten gunste van de combinatieklas. De resultaten met betrekking tot het zelfbeeld en de attitudes van de leerlingen t.o.v. de school laten een

klein positief effect zien voor de leerlingen in de combinatieklassen die om pedagogisch/didactische motieven zijn samengesteld. De voornaamste resultaten zijn samengevat in de onderste helft van Tabel 1.

8 Uitkomsten van de Nederlandse studies

Tabel 2 geeft een overzicht van de 5 studies die in Nederland zijn uitgevoerd met betrekking tot de cognitieve en niet-cognitieve uitkomsten van de leerlingen in combinatieklassen in vergelijking met de leerlingen in enkelvoudige klassen. In dit overzicht zijn de uitkomsten weergegeven per groep. Voor de vakken lezen, rekenen en taal werden geen consistente significante verschillen gevonden. De mediaan van de effectgrootte-maten bedraagt $-.04$ (gemiddelde $-.05$). Van de 19 uitgevoerde vergelijkingen voor de leerprestaties waren er 13 niet significant en 6 significant ten gunste van de enkelvoudige klas. Voor de 8 uitgevoerde vergelijkingen voor de niet-cognitieve uitkomsten werd één significant verschil gevonden ten gunste van de enkelvoudige klas. De mediaan van de niet-cognitieve effect-maten bedraagt $+.02$ (gemiddelde $.00$). Ook voor de Nederlandse studies kan geconcludeerd worden dat er nauwelijks verschillen zijn in de leerprestaties tussen leerlingen in combinatie- en enkelvoudige klassen. Ook niet voor het affectief-functioneren van de leerlingen³.

9 Aanvullende gegevens

Ten aanzien van de leerprestaties werden geen verschillen gevonden tussen de leerlingen in combinatieklassen (gevormd om reden van beheersmatig/organisatorische motieven) en enkelvoudige klassen die gesitueerd waren in stedelijke gebieden of op het platteland ($EG = +.01$ om $EG = -.10$). Met betrekking tot de niet-cognitieve variabelen bleken de leerlingen in de combinatieklassen gesitueerd in landelijke gebieden hoger te scoren dan de leerlingen in de enkelvoudige klassen op het platteland. Een kleine maar positieve effectgrootte werd gevonden voor de combinatieklas op het platteland ($EG = +.28$). Combinatieklassen om reden

van pedagogisch/didactische motieven kwamen alleen voor in stedelijke gebieden. Met betrekking tot de cognitieve en de niet-cognitieve uitkomsten verschilden zij niet van de enkelvoudige klassen ($EG = -.03$ om $EG = +.15$).

Ook werden nauwelijks verschillen aangetroffen ten aanzien van de sociaal-economische achtergrond van de scholen of klassen. Een klein positief niet-cognitief effect werd gevonden voor de combinatieklas (gevormd om reden van beheersmatig/organisatorische motieven) met leerlingen uit de bovenste en middelste sociaal-economische milieus ($EG = +.25$). Studies met combinatieklassen, die gevormd zijn om reden van pedagogisch/didactische motieven, met leerlingen uit de lagere sociale klassen kwamen nauwelijks voor.

Voor het aantal jaren dat leerlingen in combinatieklassen hadden gezeten, werden eveneens geen systematische verschillen gevonden. Ook werden geen belangrijke verschillen aangetroffen tussen de onderbouw (groep 2-4), middenbouw (groep 5 en 6) en bovenbouw (groep 7 en 8). Voor de tabellen met de resultaten per locatie, SES, bouw en het aantal jaren dat doorgebracht werd in een combinatieklas wordt verwezen naar Veenman (1996a,b)⁴.

10 Discussie

De uitkomsten van deze studie geven aan dat de leerlingen in de combinatieklassen niet meer of minder leren dan de leerlingen in de enkelvoudige klassen. Hiervoor kunnen verschillende verklaringen gegeven worden. Als eerste verklaring kan gelden dat het vrij onwaarschijnlijk is dat het groeperen van leerlingen op zich het leren van de leerlingen beïnvloedt. Of leerlingen veel leren hangt veel meer af van de kwaliteit van het instructieproces dan van de groeperingsvormen op school. Op dit moment zijn weinig gegevens bekend over de feitelijke gang van zaken met betrekking tot het onderwijs in combinatieklassen. De meeste van de hier geanalyseerde studies geven geen concreet beeld van het onderwijs in combinatieklassen. In toekomstig onderzoek zou niet alleen gekeken moeten worden naar de mogelijke effecten van verschillende groeperingsvormen voor de leerlingen maar ook naar de processen die aan

deze effecten ten grondslag liggen. Uit de studies die wel gegevens verschaffen over het onderwijsproces in combinatieklassen kan worden opgemaakt dat de meest gangbare praktijk bestaat uit het instrueren door de leerkracht van één bepaalde jaargroep terwijl de andere jaargroep (of -groepen) zelfstandig de eerder onderwezen leerstof verwerkt. Meestal verdeelt de leerkracht zijn instructietijd over de jaargroepen afzonderlijk. In de combinatieklas wordt aldus het model van het onderwijs in de enkelvoudige klas gekopieerd. De mogelijkheden die de combinatieklas biedt voor het effectief groeperen van de leerlingen teneinde de heterogeniteit tussen de leerlingen te verkleinen om het onderwijsproces beter af te stemmen op de capaciteiten van de leerlingen worden onvoldoende benut. Groeperen van de leerlingen over de grenzen van de leerjaren heen geschiedt zelden. Toch laten de overzichtsstudies van Slavin (1987) en Gutiérrez en Slavin (1992) zien dat het groeperen van de leerlingen over de grenzen van de leerjaren heen de leerprestaties van de leerlingen positief kan beïnvloeden.

Een tweede verklaring voor het ontbreken van verschillen in de leeruitkomsten van de leerlingen in combinatieklassen in vergelijking met die van de leerlingen in enkelvoudige klassen ligt mogelijk in de gehanteerde selectiecriteria van de school. Uit verschillende studies blijkt dat de schoolleiders en de leerkrachten de leerlingen voor de combinatieklas vooraf selecteren (bijvoorbeeld Adair, 1978; Gayfer, 1991; Spratt, 1986). Veel voorkomende selectiecriteria zijn: zelfstandig kunnen werken, bereidheid tot samenwerking, geen gedrags- of emotionele problemen. Deze criteria worden meestal gehanteerd in scholen in stedelijke gebieden. Zij leiden tot initiële verschillen tussen de groepen. In welke mate de scholen in de hier geanalyseerde studies selectiecriteria hanteerden is onbekend. Het effect van selectiecriteria is speculatief maar kan het ontbreken van verschillen tussen combinatie- en enkelvoudige klassen verklaren. Van de andere kant moet gesteld worden dat in de meeste studies die zich richtten op de combinatieklas in landelijke gebieden, waar de scholen geen selectiecriteria kunnen hanteren, ook geen verschillen gevonden werden.

Een derde verklaring voor het ontbreken

van verschillen tussen combinatie- en enkelvoudige klassen is de onvoldoende voorbereiding van de leerkracht op het geven van onderwijs aan twee of meer jaargroepen tegelijkertijd en het ontbreken van geschikt onderwijsmateriaal voor de combinatieklas. In veel studies maken leerkrachten melding van een gebrekkige opleiding en nascholing en van gebrekkige onderwijsmaterialen. Een onvoldoende opleiding en gebrekkige materialen kunnen de negatieve instelling van de leerkrachten ten aanzien van de combinatieklas en de gangbare onderwijspraktijk van de enkelvoudige klas als norm versterken.

Een vierde verklaring voor het ontbreken van verschillen ligt mogelijk in de klacht van de leerkrachten dat de combinatieklas meer werk met zich brengt, meer voorbereidings- en correctietijd eist en een groter beroep doet op effectieve klassenmanagementvaardigheden. Op een of andere manier slagen leerkrachten in combinatieklassen er in hun leerlingen even veel te laten leren als de leerlingen in de enkelvoudige klassen. Gegeven het extra beslag dat de combinatieklas legt op de leerkracht, heeft deze weinig mogelijkheden en energie voor het introduceren van effectievere vormen van groeperen en blijft hij vasthouden aan een praktijk die ook gangbaar is in de enkelvoudige klas.

Op dit moment is niet bekend welke klassen of groepen zich het beste lenen voor het vormen van combinaties. In sommige studies worden de lagere klassen voor mogelijke combinaties aanbevolen (bijvoorbeeld Gayfer, 1991), in andere studies de klassen in de middenbouw (bijvoorbeeld Walsh, 1989) en in weer andere studies de hoogste klassen (bijvoorbeeld Daniel, 1988). In toekomstig onderzoek dient aan deze problematiek meer aandacht besteed te worden.

Ook is weinig bekend over de optimale grootte van de combinatieklas. Schoolleiders en leerkrachten vinden vaak een aantal van 20-25 leerlingen genoeg (Gayfer, 1991). Marklund (1962) kwam op grond van zijn onderzoek tot de bevinding dat de leerlingen in combinatieklassen met niet meer dan 25 leerlingen even goed presteerden als de leerlingen in vergelijkbare enkelvoudige klassen. In combinatieklassen met meer dan 25 leerlingen presteerden de leerlingen minder dan de leerlingen in verge-

lijkbare enkelvoudige klassen. Jarousse en Mingat (1991a,b, 1992) vonden in hun studies, die uitgevoerd werden in twee Afrikaanse landen, dat voor klassen met meer dan 55 leerlingen de negatieve gevolgen van een grote combinatieklas minder groot waren dan van een grote enkelvoudige klas. Deze gegevens laten zien dat de optimale grootte van een combinatieklas in toekomstig onderzoek nadere bestudering verdient, waarbij een onderscheid gemaakt dient te worden tussen geïndustrialiseerde landen en ontwikkelingslanden.

Noten

- 1 In deze tabellen staan per studie de gegevens vermeld over de onderzochte leerjaren, de grootte van de steekproef (het aantal combinatie- en enkelvoudige klassen), het aantal jaren dat de leerlingen in combinatieklassen hebben doorgebracht, de onderzoeksopzet, de gebruikte toetsen en de effectgrootte-maten per afhankelijke variabele (lezen, rekenen, taal, zelfbeeld e.d.). Incidenteel treden met betrekking tot de effectgrootte-maten kleine afwijkingen op met de resultaten zoals gepubliceerd in *Review of Educational Research* (Veenman, 1995). Dit komt omdat in dit artikel de resultaten van drie nagekomen studies (Pawluk, 1992; Knuver, 1993; Doolaard, 1996) konden worden opgenomen. Voor de meest volledige resultaten en tabellen wordt verwezen naar Veenman (1996a,b).
- 2 Ik ben S. Doolaard erkentelijk voor de her-analyse die zij uitgevoerd heeft op de data-bestanden van Brandsma (1993) en Knuver (1993). M. Kral ben ik erkentelijk voor de her-analyse van haar gegevens (Kral, 1995). Ook dank ik S. Hämäläinen (Universiteit van Turku, Finland) voor haar toelichting bij de Finse onderzoeksrapporten van Jokinen (1979) en Nieminen (1979).
- 3 In nog 2 andere Nederlandse studies werden gegevens verzameld over de leerprestaties van leerlingen in combinatie- versus enkelvoudige klassen. Een her-analyse van de gegevens van de studie van Roelofs (1993) leverde de volgende resultaten op: EG lezen = $-.08$; EG rekenen = $-.14$; EGtotaal = $-.11$ ($n = 1487$ leerlingen waarvan 1145 in combinatie- en 342 in enkelvoudige klassen; groep 3-8). Een her-analyse van de gegevens van de studie van Veenman en Raemaekers (1996) gaf de volgende resultaten: EG lezen =

$+.06$; EG rekenen = $+.07$; EGtotaal = $+.07$ ($n = 970$ leerlingen waarvan 632 in combinatie- en 338 in enkelvoudige klassen; groep 3-8). Omdat het hier twee trainingsstudies betrof, zijn deze gegevens op grond van de sluiscriteria (zie par. 4.3.) niet in de 'best-evidence'-synthese opgenomen.

- 4 Omdat effectgrootten gebaseerd op grote steekproeven nauwkeuriger zijn en derhalve in de analyse meer gewicht dienen te krijgen, is eveneens het 'random effects model' toegepast (zie Schwarzer, 1989). Deze analyse leverde de volgende gemiddelden op. Voor de combinatieklassen om reden van beheermatig/organisatorische motieven: (a) cognitieve effecten (40 studies met effectgrootten): d -index = $-.001$ (95% betrouwbaarheidsinterval $-.06$ en $.06$; steekproefomvang $n = 69913$); (b) niet-cognitieve effecten (15 studies met effectgrootten): d -index = $.08$ (95% betrouwbaarheidsinterval $-.03$ en $.20$; steekproefomvang $n = 15071$). Voor de combinatieklassen om reden van pedagogisch/didactische motieven: cognitieve effecten (8 studies met effectgrootten): d -index = $-.08$ (95% betrouwbaarheidsinterval $-.21$ en $.06$; steekproefomvang $n = 2390$); (d) niet-cognitieve effecten (8 studies met effectgrootten): d -index = $.13$ (95% betrouwbaarheidsinterval $-.02$ en $.27$; steekproefomvang $n = 2352$). Weging van de effectgrootten leidde niet tot wezenlijk andere resultaten (zie Veenman, 1996a).

Literatuur

- Adair, J.H. (1978). *An attitude and achievement comparison between kindergarten and first grade children in multi and single grade classes*. Unpublished doctoral dissertation, Boston College.
- Adams, J.J. (1953). Achievement and social adjustment of pupils in combination classes enrolling pupils of more than one grade level. *Journal of Educational Research*, 47, 151-155.
- Bangert-Drowns, R.L. (1986). Review of developments in meta-analytic method. *Psychological Bulletin*, 99(3), 388-399.
- Brandsma, H.P. (1993). *Basisschoolkenmerken en de kwaliteit van het onderwijs*. (Academisch proefschrift, Rijksuniversiteit Groningen). Groningen: RION, Instituut voor Onderwijsonderzoek.
- Brown, K.G., & Martin, A.B. (1989). Student achieve-

- ment in multigrade and single grade classes. *Education Canada*, 29(2), 10-13, 47.
- Burchyett, J.A. (1972). *A comparison of the effects of nongraded, multi-age, team teaching vs. the modified self-contained classroom at the elementary school level*. Unpublished doctoral dissertation, Michigan State University.
- Campbell, G. (1993). Assessing the impact of multi-grade classes: An annotated bibliography. *Canadian Modern Language Review*, 49(2), 345-364.
- Carter, J.B. (1973). *A study of the effects of multi-grade grouping on the attitudes toward self, peers, and school of selected third and fifth grade students*. Unpublished doctoral dissertation, Michigan State University.
- Chace, E.S. (1961). *An analysis of some aspects of multiple-grade grouping in an elementary school*. Unpublished doctoral dissertation, University of Tennessee.
- Cohen, J. (1969). *Statistical power analysis for behavioral sciences*. New York: Academic Press.
- Commissie Evaluatie Basisonderwijs. (1994a). *Onderwijs op maat*. Leiden: Distributiecentrum DOP.
- Commissie Evaluatie Basisonderwijs. (1994b). *Inhoud en opbrengsten van het basisonderwijs*. Leiden: Distributiecentrum DOP.
- Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design & analysis for field settings*. Chicago: Rand McNally.
- Daniel, I.H. (1988). "Doing the splits": Core French in the elementary schools. The impact of combined classes on students. *Canadian Modern Language Review*, 45(1), 146-154.
- Dodendorf, D.M. (1983). A unique rural school environment. *Psychology in the Schools*, 20(1), 99-104.
- Doolaard, S. (1996). *Leren in combinatiegroepen*. Enschede: Universiteit Twente, Faculteit der Toegepaste Onderwijskunde.
- Draisey, A.G. (1985). Vertical grouping in the primary school: A positive view. *Education for Development*, 9(1), 3-11.
- Dreier, W.H. (1949). The differential achievement in rural graded and ungraded school pupils. *Journal of Educational Research*, 43(3), 175-185.
- Finley, C.J., & Thompson, J.M. (1963). A comparison of the achievement of multi-graded and single-graded rural elementary school children. *Journal of Educational Research*, 56(9), 471-475.
- Fippinger, F. (1967). Empirische Untersuchung zur Leistung von Schülern aus voll und wenig gegliederten Schulen. *Schule und Psychologie*, 14(4), 97-104.
- Furch-Krafft, E. (1979). Schulische organisation im Primarbereich und Leistungsverhalten: Ein Vergleich der Schulleistungen von Schülern aus voll und wenig gegliederten Grundschulen. *Zeitschrift für Empirische Pädagogik*, 3, 297-308.
- Gajadharsingh, J.L. (1987). *The cognitive profiles of students in multi-grade and single-grade classrooms as a function of gender and locale*. Regina: Saskatchewan School Trustees Association Research Centre.
- Galton, M., & Simon, B. (Eds.) (1980). *Progress and performance in the primary classroom*. London: Routledge & Kegan Paul.
- Gayfer, M. (Ed.) (1991). *The multi-grade classroom: Myth and reality. A Canadian study*. Toronto: Canadian Education Association.
- Givens, H., Jr. (1972). *A comparative study of achievement and attitudinal characteristics of black and white intermediate pupils in individualized-multigrade and self-contained instructional programs*. Unpublished doctoral dissertation, Saint Louis University.
- Glass, G.V., McGaw, B., & Smith, M.L. (1981). *Meta-analysis in social research*. Beverly Hills: Sage.
- Grove, E.H. (1978). *An investigation of the effect of multi-age grouping on student attitude*. Unpublished doctoral dissertation, Washington State University.
- Gutiérrez, R., & Slavin, R.E. (1992). Achievement effects of the nongraded elementary school: A best evidence synthesis. *Review of Educational Research*, 62(4), 333-376.
- Hammack, B.A. (1974). *Self-concept: Evaluation of preschool children in single and multi-age classroom settings*. Unpublished doctoral dissertation, Texas Woman's University.
- Harvey, S.B. (1974). *A comparison of kindergarten children in multigrade and traditional settings on self-concept, social-emotional development, readiness development, and achievement*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University.
- Higgins, J.J. (1980). *A comparative study between the reading achievement levels of students in a combination/ungraded class and students in a graded class*. Unpublished doctoral dissertation, George Peabody College for Teachers of Vanderbilt University.
- Her Majesty's Inspectorate. (1978). *Primary education in England: A survey by HM Inspectors of Schools*. London: HMSO, Department of Education and Science.

- Hoen, R.R. (1972). *An evaluation of multi-age classes at Carnavon School, 1971-72*. Vancouver: Board of School Trustees, Department of Planning and Evaluation. (ERIC Document Reproduction Service No. ED 076 649)
- Inspectie Basisonderwijs. (1978). *Onderwijsverslag: Verslag van de staat van het onderwijs in Nederland*. 's-Gravenhage: Staatsuitgeverij.
- Jarousse, J.P., & Mingat, A. (1991a). *Efficacité pédagogique de l'enseignement à cours multiples dans le contexte africain*. Dijon: Institut de Recherche sur L'Économie de l'Éducation, Centre National de la Recherche Scientifique.
- Jarousse, J.P., & Mingat, A. (1991b). *Les facteurs agissant sur les acquisitions des élèves à l'école primaire au Burkina Faso*. Dijon: Institut de Recherche sur L'Économie de l'Éducation, Centre National de la Recherche Scientifique.
- Jarousse, J.P., & Mingat, A. (1992). *L'école primaire en Afrique: Fonctionnement, qualité, produits. Le cas du Togo*. Dijon: Institut de Recherche sur L'Économie de l'Éducation, Centre National de la Recherche Scientifique.
- Jokinen, H. (1979). *Yhdysluokkakoulujen oppilaiden koulusaavutuksista 2: Kouluhallituksen yhteisiin kokeisiin perustuva selvitys*. Jyväskylä: University of Jyväskylä, Institute for Educational Research.
- Junell, J.S. (1971). *An analysis of the effects of multi-grading on a number of noncognitive variables*. Unpublished doctoral dissertation, University of Washington.
- Katz, L.G., Evangelou, D., Hartman, J.A. (1993). *The case for mixed-age grouping in early education* (3rd ed.). Washington, DC: National Association for the Education of Young Children.
- Knight, E.E. (1938). A study of double grades in New Haven city schools. *Journal of Experimental Education*, 7(1), 11-18.
- Knörzer, W. (1985). *Sind Schüler in kombinierten Grundschulklassen benachteiligt? Eine empirische Untersuchung*. Baltmannsweiler: Pädagogischer Verlag.
- Knuver, J.W.M. (1993). *De relatie tussen klas- en schoolkenmerken en het affectief functioneren van leerlingen*. (Academisch proefschrift, Rijksuniversiteit Groningen). Groningen: RION, Instituut voor Onderwijsonderzoek.
- Kral, M. (1995). *Effecten van schoolgrootte: Een onderzoek naar effectiviteitsverschillen tussen combinatieklassen en enkelvoudige klassen in het basisonderwijs*. Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen / Vakgroep Onderwijskunde, K.U. Nijmegen.
- Kral, M., Voeten, M., & Nederlof, W. (1995). Tijdsbesteding in combinatieklassen en enkelvoudige klassen: Een descriptief onderzoek naar de besteding van de beschikbare tijd op basisscholen. *Tijdschrift voor Onderwijs Research*, 20(4), 300-317.
- Lincoln, R.D. (1981). *The effect of single-grade and multi-grade primary school classrooms on reading achievement of children*. Unpublished doctoral dissertation, University of Connecticut.
- Lungwangwa, G. (1990). *Meeting the educational needs of children in sparsely populated areas through multigrade teaching: An experience from Zambia*. Lusaka: University of Zambia. (ERIC Document Reproduction Service No. ED 339 562)
- Marklund, S. (1962). *Skolklassens storlek och struktur*. (Doctoral dissertation, University of Stockholm). Stockholm: Almqvist & Wiksell.
- Marsh, M.M. (1980). *Academic achievement and school-wide grouping of students in two middle schools*. Unpublished doctoral dissertation, University of Florida.
- Martens, C.C. (1954). Educational achievements of eight-grade pupils in one-room rural and graded town schools. *Elementary School Journal*, 54 (May), 523-525.
- Milburn, D. (1981). A study of multi-age or family-grouped classrooms. *Phi Delta Kappan*, 62(7), 513-514.
- Mobley, C.F. (1976). *A comparison of the effects of multiage grouping versus homogeneous age grouping in primary school classes of reading and mathematics achievement*. Unpublished doctoral dissertation, Nova University. (ERIC Document Reproduction Service No. ED 128 102)
- Mycock, M.A. (1966). *A comparison of vertical grouping and horizontal grouping in the infant school*. Unpublished master's thesis, University of Manchester.
- Nieminen, R. (1979). *Yhdysluokkakoulujen oppilaiden koulusaavutuksista 3: IEA-aineistoon perustuva selvitys kansakoulun oppilaiden menestymisestä luonnontiedon ja äidinkielen kokeissa*. Jyväskylä: University of Jyväskylä, Institute for Educational Research.
- Papay, J.P., Costello, R.J., Hedl, J.J., & Spielberger, C.D. (1975). Effects of trait and state anxiety on the performance of elementary school children in traditional and individualized multiage classrooms. *Journal of Educational Psychology*, 67(6), 840-846.
- Pawluk, S.T. (1992). *A comparison of the academic achievement of students in multigrade elemen-*

- ary classrooms and students in self-contained single-grade elementary classrooms. Unpublished doctoral dissertation, Montana State University.
- Poglia, E., & Strittmatter, A. (1983). *Die Situation der Mehrklassenschulen in der Schweiz: Ergebnisse einer Untersuchung*. Genf: Schweizerische Konferenz der kantonalen Erziehungsdirektoren.
- Pratt, C., Holub, K., & Ainsworth, J. (1989). Students' ratings of their classroom environment in vertically and horizontally grouped classes. *Education Research and Perspectives*, 16(2), 71-76.
- Pratt, C., & Treacy, K. (1986). *A study of student grouping practices in early childhood classes in Western Australian Government Primary Schools*. Nedlands: Education Department of Western Australia.
- Projectgroep Schaalvergroting Basisonderwijs. (1990). *Schaal en kwaliteit in het basisonderwijs*. Zoetermeer: Ministerie van Onderwijs en Wetenschappen.
- Psacharopoulos, G., Rojas, C., & Velez, E. (1993). Achievement evaluation of Colombia's Escuela Nueva: Is multigrade the answer? *Comparative Education Review*, 37(3), 263-276.
- Purl, M.C., & Curtis, J. (1970). *A look at combination class effects at Emerson Elementary School*. Riverside, CA.: Riverside Unified School District. (ERIC Document Reproduction Service No. ED 042 863)
- Råberg, A. (1976). *Glesbygdsbarn i årskurs 3 och 6: Standardprov och betyg*. Umeå: Umeå Universitet och Lärarhögskolan, Pedagogiska Institutionen.
- Rehwoldt, W., & Hamilton, W. (1957). *An analysis of some of the effects of interage and intergrade grouping in an elementary school*. Unpublished doctoral dissertation, University of Southern California.
- Reichardt, C.S. (1979). The statistical analysis of data from nonequivalent group designs. In T.C. Cook & D.T. Campbell (Eds.), *Quasi-experimentation: Design and analysis issues for field settings* (pp. 147-205). Chicago: Rand McNally.
- Roelofs, E. (1993). *Teamgerichte nascholing en coaching: een experimentele studie in scholen met combinatieklassen* (Academisch proefschrift, Katholieke Universiteit Nijmegen). Nijmegen; Universiteitsdrukkerij.
- Rojas, C., & Castillo, Z. (1988). *Evaluación del programa Escuela Nueva en Colombia*. Bogotá: Instituto SER de Investigación.
- Rowley, S.D. (1992). *Multigrade classrooms in Pakistan: How teacher conditions and practices affect student achievement*. Unpublished doctoral dissertation, Harvard University.
- Rule, J.G. (1983). *Effects of multigrade grouping on elementary student achievement in reading and mathematics*. Unpublished doctoral dissertation, Northern Arizona University.
- Schrankler, W.J. (1976). Family groupings and the affective domain. *Elementary School Journal*, 76(7), 432-439.
- Schroeder, R., & Nott, R.E. (1974). Multi-age grouping... It works! *Catalyst for Change*, 3, 15-18.
- Schwarzer, R. (1989). *Meta-analysis programs*. Berlin: Freie Universität, Institut für Psychologie.
- Slavin, R.E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher*, 15(9), 5-11.
- Slavin, R.E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57(3), 293-336.
- Slavin, R.E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research*, 60(3), 471-499.
- Spratt, B.R. (1986). *A comparative study of children enrolled in combination classes and non-combination classes in Fairfax County, Virginia Public Schools*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University.
- Steinhofer, T.L. (1980). *The effects of different classroom organizational structures on the achievement of kindergarten and first grade pupils*. Unpublished doctoral dissertation, University of San Francisco.
- Stimson, J.M. (1991). *The effects of multigrade classes on student achievement in year-round schools*. Unpublished doctoral dissertation, Northern Arizona University.
- Stone, W.M. (1986). *A study of the relationships between multigrading and academic progress of elementary school students*. Unpublished doctoral dissertation, Vanderbilt University.
- Thomas, C., & Shaw, C. (1992). *Issues in the development of multigrade schools* (World Bank technical paper number 172). Washington, DC: World Bank.
- UNESCO. (1989). *Multigrade teaching in single teacher primary schools*. Bangkok: UNESCO Principal Regional Office for Asia and the Pacific.
- Veenman, S. (1995). Cognitive and noncognitive

- effects of multigrade and multi-age classes: A best-evidence synthesis. *Review of Educational Research*, 65(4), 319-381.
- Veenman, S. (1996a). Effects of multigrade and multi-age classes reconsidered. *Review of Educational Research*, 66(3), 323-340.
- Veenman, S. (1996b). *Cognitieve en niet-cognitieve leeruitkomsten van leerlingen in combinatieklassen: een 'best-evidence'-synthese*. Nijmegen: Katholieke Universiteit, Vakgroep Onderwijskunde.
- Veenman, S., Lem, P., Voeten, M., Winkelmolen, G., & Lassche, H. (1986). *Onderwijs in combinatieklassen*. 's-Gravenhage: Instituut voor Onderzoek van het Onderwijs.
- Veenman, S., Lem, P., & Winkelmolen, B. (1985). Active learning time in mixed-age classes. *Educational Studies*, 11(3), 171-180.
- Veenman, S., & Raemaekers, J. (1996). Retentie-effecten van een nascholingsprogramma voor effectieve instructie en klasmanagement. *Pedagogische Studiën*, 73, 357-371.
- Veenman, S., Voeten, M., & Lem, P. (1987). Classroom time and achievement in mixed-age classes. *Educational Studies*, 13(1), 75-89.
- Vogel, F.X., & Bowers, N.D. (1972). Pupil behavior in a multi-age nongraded school. *Journal of Experimental Education*, 41(2), 78-86.
- Walsh, J.M. (1989). *Attitudes of students, parents, and teachers toward multigrading*. Unpublished doctoral dissertation, University of Connecticut.
- Waraksa, D.L. (1989). *Administrative considerations for multigrade classrooms: The case for split grades*. Unpublished master's thesis, University of Calgary, Alberta, Canada.
- Way, J.W. (1981). Achievement and self-concept in multiage classrooms. *Educational Research Quarterly*, 6(2), 69-75.
- Yerry, M.J., & Henderson, E. (1964). *Effects of inter-age grouping on achievement and behavior* (End-of-year report, July 1964). Bethpage, N.Y.: Plainedge Public Schools. (ERIC Document Reproduction Service No. ED 037 802)
- Zabolotney, A.B. (1983). *A comparison of reading achievement and school attitudes of rural Seventh-day Adventist multi-graded students and public school single-graded students in the state of Arkansas*. Unpublished doctoral dissertation, Brigham Young University.

Auteur

S. Veenman is als universitair hoofddocent werkzaam bij de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen.

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

Abstract

Learning effects of students in multi-grade classes: a best-evidence synthesis

S. Veenman. *Pedagogische Studiën*, 1997, 74, 73-96.

This article reviews the best evidence concerning the cognitive and noncognitive effects of multi-grade and multi-age classes in elementary schools. In these classes, students from two or more grades are taught by one teacher in one room at the same time. Multi-grade classes are formed for administrative reasons to cope with declining student enrollment or uneven class sizes. In multi-age classes, students of different ages are grouped together for educational and pedagogical benefits. This review summarizes the outcomes of 59 studies. Studies in which the cognitive effects of multi-grade and single-grade classes were compared revealed no consistent differences between these two types of classroom organization. Studies of the noncognitive effects of multi-grade and single-grade classes also produced no consistent differences. Studies in which the cognitive effects of multi-age and single-age classes were compared indicated no differences between these two types of grouping. And, finally, studies of the noncognitive effects of multi-age and single-age classes produced inconsistent effects. The results of the five Dutch studies also revealed no consistent differences between multi-grade and single-grade classes. It is concluded that no empirical evidence exists for the assumption that student learning may suffer in multi-grade or multi-age classrooms.

Tabel 1

Samenvatting van de cognitieve en niet-cognitieve effecten van combinatieklassen per vak/domein

Vak/domein	Significant positief	Geen verschil	Significant negatief	Mediaan effectgrootte
<i>Combinatieklassen om reden van beheersmatiglorganisatorische motieven</i>				
Totaal cognitief	4	30	6	-0.01 (36)
Lezen	2	21	4	-0.03 (24)
Rekenen	2	18	5	-0.03 (23)
Taal	2	17	3	+0.01 (20)
Mixtum	2	6	1	-0.02 (8)
Totaal niet-cognitief	5	14	0	+0.08 (15)
Attitudes t.o.v. de school	3	8	1	+0.05 (10)
Zelfbeeld	3	6	0	+0.13 (8)
Persoonlijke aanpassing	2	4	0	+0.29 (5)
Sociale aanpassing	2	4	0	+0.33 (3)
Motivatie	0	5	0	-0.03 (5)
<i>Combinatieklassen om reden van pedagogischdidactische motieven</i>				
Totaal cognitief	1	9	1	-.03 (8)
Lezen	0	8	0	-.03 (6)
Rekenen	0	7	2	-.10 (6)
Taal	0	2	1	-.02 (3)
Mixtum	1	0	1	-.24 (1)
Totaal niet-cognitief	6	5	0	+0.15 (8)
Attitudes t.o.v. de school	4	0	1	+0.17 (2)
Zelfbeeld	3	2	0	+0.18 (3)
Persoonlijke aanpassing	2	2	0	+0.15 (4)
Sociale aanpassing	1	0	1	+0.18 (3)
Motivatie	0	1	0	-.06 (1)

Opm. Tussen haakjes het aantal studies waarvoor een effectgrootte-maat kon worden berekend.

Tabel 2

Samenvatting van de cognitieve en niet-cognitieve effecten in combinatieklassen in Nederland

Publicatie	Onderzoeksgroep	Groep	Vak	Effectgrootte
<i>Cognitieve effecten</i>				
Veenman e.a., 1986	12 C-klassen, 12 E-klassen	5	Lezen	+0.05
		5	Rekenen	+0.01
		6	Lezen	-0.11
		6	Rekenen	-0.02
Brandsma, 1993	74 C-klassen, 118 E-klassen	8	Lezen	-0.08*
		8	Rekenen	-0.04
		8	Taal	-0.09**
		8	Lezen	-0.06
Kral, 1995	91 C- klassen, 40 E-klassen	4	Rekenen	+0.06
		4	Taal	-0.12*
		6	Lezen	-0.04
		6	Rekenen	+0.03
		6	Taal	-0.02
		8	Lezen	-0.05
		8	Rekenen	-0.17**
		8	Taal	-0.14*
Doolaard, 1996	113 C-klassen, 104 E-klassen	6	Lezen	-0.00
		6	Rekenen	-0.15**
		6	Taal	+0.01
		6	Attitude t.o.v. taal	+0.04
<i>Niet-cognitieve effecten</i>				
Knuver, 1993	83 C-klassen, 134 E-klassen	8	Attitude t.o.v. rekenen	+0.02
		8	(Prestatie)Motivatie	-0.03
		8	Attitude t.o.v. de school	+0.03
		8	Attitude t.o.v. taal	-0.01
Doolaard, 1996	113 C-klassen, 104 E-klassen	6	Attitude t.o.v. rekenen	+0.02
		6	(Prestatie)Motivatie	+0.03
		6	Attitude t.o.v. de school	-0.08*
		6	Attitude t.o.v. taal	-0.01

Opm. Nederlandse indeling in groepen aangehouden. C = aantal combinatieklassen, E = aantal enkelvoudige klassen. De studie van Doolaard bestaat uit 2 deelstudies. Zie ook noot 3.

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

De professionele oriëntatie van leraren en hun betrokkenheid bij het schoolbeleid

C.T. Jongmans en D. Beijaard

Samenvatting

Scholen moeten beschikken over beleidsvoerende vermogen om zelfstandig beleid te kunnen voeren. Het beleidsvoerende vermogen van scholen wordt in belangrijke mate bepaald door de betrokkenheid van leraren bij het schoolbeleid. Verondersteld wordt dat er een relatie is tussen de mate waarin leraren betrokken zijn bij het schoolbeleid en hun professionele oriëntatie. Voor het onderzoek hiernaar hebben 1079 leraren van 98 VBO- en MBO-scholen voor voortgezet agrarisch onderwijs een vragenlijst ingevuld. Op grond van de resultaten kunnen twee typen leraren worden onderscheiden: leraren met een minder ruime professionele oriëntatie (type 1) en leraren met een ruime professionele oriëntatie (type 2). Gebleken is dat de type 1 leraren minder betrokken zijn bij het schoolbeleid dan de type 2 leraren. Dit verschil is echter relatief: voor beide typen leraren geldt dat zij beperkt betrokken zijn bij het onderwijskundig beleid en zeer beperkt betrokken zijn bij het beheersmatig beleid van hun school. Deze resultaten wijzen erop dat een ruimere professionele oriëntatie van leraren zich alleen vertaalt in een hogere mate van betrokkenheid bij het schoolbeleid wanneer aan leraren ook de mogelijkheden wordt geboden om zich op dit terrein te begeven.

Inleiding

De overheid heeft beleid in gang gezet dat moet leiden tot een grotere mate van zelfbestuur door scholen. Vergroting van autonomie leidt ertoe dat scholen steeds meer een eigen beleid voor hun organisatie moeten ontwikkelen. De overheid verwacht dat deze ontwikkeling leidt tot een verhoging van de kwaliteit van het onderwijs.

In literatuur wordt erop gewezen dat scholen moeten beschikken over beleidsvoerende

vermogen om zelfstandiger beleid te kunnen voeren (vgl. Giesbers, Schuit & Silvertand, 1987; Marx, 1987, 1988; Marx e.a., 1995). Slegers (1991) heeft aangetoond dat scholen waar leraren in grote mate deelnemen aan de beleidsvoering beter in staat zijn zelfstandig beleid te voeren dan scholen waar dit niet of minder het geval is. Met andere woorden, de betrokkenheid van leraren bij het schoolbeleid bepaalt in belangrijke mate het beleidsvoerende vermogen van een school.

De resultaten van de studie van Slegers (1991) doen vermoeden dat het beleidsvoerende vermogen van scholen en de professionele oriëntatie van leraren aan elkaar gerelateerd zijn. Binnen het schoolorganisatie-onderzoek is het volgens hem dan ook gewenst de aandacht te richten op de professionele oriëntatie van leraren. Hiervoor is het door Hoyle (1975, 1980) gemaakte onderscheid tussen een 'restricted' (beperkte) en 'extended' (uitgebreide) professionaliteit van leraren geschikt (Slegers, 1991; Slegers & Bergen, 1992).

Dit onderzoek is gebaseerd op de veronderstelling dat leraren met een uitgebreide professionele oriëntatie meer betrokken zijn bij het schoolbeleid dan hun collega's met een beperkte professionele oriëntatie. Empirische gegevens over deze relatie ontbreken voorsnog. Een bevestiging van deze relatie betekent dat in het kader van beleidsvoering van scholen leraren zich zouden moeten ontwikkelen in de richting van een uitgebreide professionele oriëntatie (vgl. ook Slegers & Bergen, 1992). De centrale vraag van het onderzoek luidt: wat is de relatie tussen de professionele oriëntatie van leraren en de mate van hun betrokkenheid bij het schoolbeleid?

In dit artikel wordt allereerst ingegaan op de theorie die aan het onderzoek ten grondslag ligt. Deze betreft het beleidsvoerende vermogen van scholen, typen scholen die qua beleidsvoerende vermogen kunnen worden onderscheiden en de professionele oriëntatie van leraren. Ver-

volgens worden de opzet en de resultaten van het onderzoek beschreven. Het artikel besluit met enkele conclusies en een discussie over de betrokkenheid van leraren bij schoolbeleid.

1 Theoretische achtergronden

Het beleidsvoerend vermogen van scholen

Scholen worden steeds meer autonoom en moeten derhalve eigen beleid voor hun organisatie gaan ontwikkelen. In met name de (school)organisatiekundige literatuur wordt erop gewezen dat scholen moeten beschikken over beleidsvoerend vermogen om zelfstandiger beleid te kunnen voeren (Giesbers, Schuit & Silvertand, 1987; Marx, 1987, 1988; Marx e.a., 1995; Slegers, 1991). Onder beleidsvoerend vermogen wordt verstaan: "het vermogen van scholen om uit een aantal, door de overheid toegestane beleidsalternatieven, een optimale beleidslijn te ontwikkelen, rekening houdend met eigen voorkeuren en de eigen schoolsituatie" (Marx e.a., 1995, p. 46).

Het beleid van een school wordt veelal onderverdeeld in een beheersmatig en een onderwijskundig domein (Schuit & Slegers, 1990). Het beheersmatig beleid omvat aspecten die een randvoorwaardelijke functie hebben met betrekking tot het verzorgen van onderwijs, zoals personeelsbeleid, het verdelen van leraarslessen en de financiën. Het onderwijskundig beleid, ook wel inhoudelijk beleid genoemd, omvat onderdelen die het verzorgen van onderwijs betreffen. Voorbeelden hiervan zijn het formuleren van schooldoelstellingen, het toetsen van leerlingen en het maken van afspraken over didactische werkvormen (vgl. Marx e.a., 1995).

Het onderwijskundig beleid en het beheersmatig beleid zijn meestal niet of nauwelijks op elkaar afgestemd (Schuit & Slegers, 1990). Een belangrijke oorzaak hiervan is dat het onderwijskundig beleid doorgaans bepaald wordt door de leraren en het beheersmatig beleid door de schoolleiding (Slegers & Bergen, 1992). De scheiding tussen beide vormen van beleid draagt ertoe bij dat scholen slechts beperkt in staat zijn zelfstandig beleid te voeren (Giesbers, Schuit & Silvertand, 1987).

Typen scholen naar beleidsvoerend vermogen
Schoolorganisatorisch onderzoek heeft aangetoond dat scholen verschillen in het vermogen om beleid te ontwikkelen en uit te voeren (vgl. Giesbers, Schuit & Silvertand, 1987; Marx, 1987, 1988; Jongmans, 1996; Jongmans & Beijaard, 1995). Slegers (1991, p. 167) onderscheidt drie typen scholen:

- 1 scholen met een hiërarchisch georiënteerd beleidsvoerend vermogen (type 1 scholen); deze scholen kenmerken zich door het feit dat de schoolleiding meer dominant is;
- 2 scholen met een collegiaal georiënteerd beleidsvoerend vermogen (type 2 scholen); deze scholen kenmerken zich door het feit dat leraren meer dominant zijn;
- 3 scholen met een beperkt collegiaal georiënteerd beleidsvoerend vermogen (type 3 scholen); deze scholen bevinden zich tussen beide hierboven aangegeven typen scholen, waarbij ze vooral overeenkomsten vertonen met de type 2 scholen.

Uit de onderzoeksresultaten van Slegers (1991) blijkt onder andere dat de type 2 scholen meer onderwijskundig en beheersmatig beleid voeren dan de type 1 scholen. Op de type 2 scholen zijn leraren meer betrokken bij de beleidsvoering binnen het onderwijskundig en het beheersmatig domein dan leraren op de type 1 scholen. Beide typen scholen verschillen aanzienlijk in de betrokkenheid van leraren bij het beheersmatig beleid. De betrokkenheid van leraren bij het schoolbeleid is door Slegers vastgesteld door aan leraren met een coördinatietak te vragen in hoeverre leraren bij dit beleid betrokken zijn.

Het onderwijskundig domein is van oudsher het domein waarbinnen leraren overheersen. Slegers (1991) concludeert echter dat er onmiskenbaar sprake is van een doorbreking van deze overheersing. Hij wijst erop dat ook de schoolleiding binnen het onderwijskundig domein actief is. Op alle door hem onderzochte scholen is sprake van een gezamenlijke verantwoordelijkheid ten aanzien van onderwijskundige aspecten. Op scholen met een collegiaal georiënteerd beleidsvoerend vermogen, de scholen waar het meest onderwijskundig beleid wordt gevoerd, hebben de leraren binnen het onderwijskundig domein echter de grootste zeggenschap.

Uit het onderzoek van Slegers (1991) is

eveneens gebleken dat op alle onderzochte scholen de schoolleiding overheersend is binnen het beheersmatig domein. Op scholen met een collegiaal georiënteerd beleidsvoerend vermogen is de betrokkenheid van leraren bij het beheersmatig beleid het grootst. Binnen dit type school wordt ook het meest beheersmatig beleid gevoerd. Slegers vermoedt dan ook dat beheersmatig beleid het meest effectief wordt gevoerd op scholen waar leraren en schoolleiding gezamenlijk verantwoordelijk zijn voor dit beleid, met dien verstande dat de schoolleiding binnen het beheersmatig domein de eerstverantwoordelijke is. Samengevat, het beleidsvoerend vermogen van scholen is het grootst op scholen waar de schoolleiding en leraren gezamenlijk verantwoordelijk zijn voor beheersmatige en onderwijskundige taken. Er dient dan wel sprake te zijn van een duidelijke taakverdeling. De schoolleiding dient primair verantwoordelijk te zijn voor taken die het beheer en het management aangaan en leraren dienen de eerste verantwoordelijkheid te hebben voor taken die het onderwijs betreffen. Op basis van deze onderzoeksresultaten concludeert Slegers (1991) dat de betrokkenheid van leraren bij beleidsvoering in belangrijke mate het beleidsvoerend vermogen van een school bepaalt.

De professionele oriëntatie van leraren

Verondersteld wordt dat het verschil in betrokkenheid van leraren bij de beleidsvoering verwijst naar verschillen in de professionele oriëntatie van leraren (Knoers, 1994; Slegers, 1991; Slegers & Bergen 1992). Hoyle (1980) onderscheidt twee vormen van professionaliteit, namelijk 'restricted professionalism' en 'extended professionalism'. 'Restricted professionalism' omschrijft Hoyle (1980, p. 49) als "a professionalism which is intuitive, classroom-focused, and based on experience rather than theory. The good restricted professional is sensitive to the development of individual pupils, an inventive teacher and a skilful class manager. He is unencumbered with theory, is not given to comparing his work with that of others, tends not to perceive his classroom activities in a broader context, and values his classroom autonomy". De leraar met een beperkte professionele oriëntatie is vooral vakdidactisch en vakinhoudelijk georiënteerd en voornamelijk

gericht op lesgevende activiteiten. Het is de leraar die zich onafhankelijk opstelt en koning(in) is in eigen koninkrijk (Knoers, 1994).

De 'extended' professional is, aldus Hoyle (1980, p. 49), "concerned with locating his classroom teaching in a broader educational context, comparing his work with that of other teachers, evaluating his own work systematically, and collaborating with other teachers. He is interested in theory and in current educational developments. Hence he reads educational books and journals, becomes involved in various professional activities and is concerned to further his own professional development through in-service work. He sees teaching as a rational activity amenable to improvement on the basis of research and development". De leraar met een uitgebreide professionele oriëntatie vindt professionele samenwerking van groot belang en is dan ook betrokken bij vakoverstijgende activiteiten in de school. Deze leraar ziet zijn/haar functioneren als teamlid als een voorwaarde voor het adequaat kunnen uitvoeren van lesgevende taken. Eveneens achten leraren met een uitgebreide beroepsoriëntatie het van belang dat hun klassepraktijk afgestemd is op de doelstellingen en het beleid van de school. Zij spannen zich in om hun school, en daarmee ook hun klassepraktijk, te laten aansluiten bij nieuwe onderwijskundige inzichten en recente maatschappelijke ontwikkelingen (Giesbers & Bergen, 1992).

De resultaten van het onderzoek van Slegers (1991) suggereren dat de betrokkenheid van leraren bij het schoolbeleid en hun professionele oriëntatie aan elkaar gerelateerd zijn. Het is namelijk opmerkelijk dat op scholen met een collegiaal georiënteerd beleidsvoerend vermogen leraren meer betrokken zijn bij de beleidsvoering binnen het beheersmatig domein dan leraren op scholen met een hiërarchisch georiënteerd beleidsvoerend vermogen. Dit gegeven kan volgens Slegers en Bergen (1992) gezien worden als een aanwijzing voor het feit dat op scholen met een collegiaal georiënteerd beleidsvoerend vermogen vooral leraren met een uitgebreide professionele oriëntatie werkzaam zijn. Op scholen met een hiërarchisch georiënteerd beleidsvoerend vermogen zijn voornamelijk leraren met een beperkte professionele oriëntatie werkzaam.

Sleegers en Bergen (1992) vermoeden op basis van deze gegevens dat een ontwikkeling van een beperkte professionele oriëntatie naar een meer uitgebreide professionele oriëntatie van leraren binnen scholen tevens een ontwikkeling inhoudt van een hiërarchisch georiënteerd naar een collegiaal georiënteerd beleidsvoerende vermogen van scholen (vgl. ook Knoers, 1994). Een ontwikkeling van een beperkte professionele oriëntatie naar een meer uitgebreide professionele oriëntatie van leraren wordt door Knoers (1994) evenals door Sleegers en Bergen (1992) dan ook van groot belang gevonden.

2 Opzet van het onderzoek

Doelgroep en respons

Het onderzoek richtte zich op leraren die werkzaam zijn in alle VBO- en MBO-scholen voor voortgezet agrarisch onderwijs. In eerste instantie zijn de directeuren van alle 108 scholen benaderd met het verzoek om het onderzoek op hun school te mogen uitvoeren. Uiteindelijk is het onderzoek op 98 scholen

uitgevoerd. Alle 2192 leraren die op deze scholen werken hebben in februari 1995 een vragenlijst ontvangen; 1079 leraren hebben een bruikbare vragenlijst teruggestuurd. Sprake is van een respons van ruim 49%. Deze onderzoeksrespons is gelijk aan 44% van alle leraren in het voortgezet agrarisch onderwijs. Vanwege de redelijk hoge respons is geen systematisch onderzoek gedaan naar non-respons.

De vragenlijst

Voor het beantwoorden van de onderzoeksvraag is een vragenlijst ontwikkeld op basis van de hiervoor beschreven theorie met betrekking tot:

- 1 de betrokkenheid van leraren bij het onderwijskundig en beheersmatig beleid (het schoolbeleid);
 - 2 de professionele oriëntatie van leraren.
- Sleegers (1991, p. 207) onderscheidt binnen het schoolbeleid van scholen voor voortgezet onderwijs 20 taken die als belangrijk kunnen worden aangemerkt (vgl. ook Hanson, 1979). In het onderhavige onderzoek zijn deze 20 taken als uitgangspunt genomen voor het vaststellen van de betrokkenheid van leraren bij het

01. Afspraken die op school zijn gemaakt over het toepassen van didactische werkvormen tijdens de lessen (didactiek).
02. Het afstemmen van leerinhouden van verschillende vakken (leerinhoud).
03. Afspraken op school over het toetsen van leerresultaten (toetsen).
04. Afspraken op school over de inhoud van stages van leerlingen (stage-inhoud).
05. De wijze waarop leerlingbegeleiding plaatsvindt (begeleiding leerlingen).
06. Het formuleren van de onderwijsdoelen van de school (doelen).
07. Het aanschaffen van leer- en hulpmiddelen (leermiddelen).

Figuur 1. Overzicht van de taken die zijn onderscheiden om inzicht te krijgen in de betrokkenheid van leraren bij het onderwijskundig beleid van hun school

01. Contacten van de school met andere scholen in en buiten het agrarisch onderwijs (contacten).
02. Het aanstellen van onderwijsondersteunend personeel (aanstellen OOP).
03. Contacten met de inspectie landbouwonderwijs (inspectie).
04. Het verdelen van financiële middelen over diverse uitgavenposten (financiën).
05. Het vast aanstellen van leraren (vast aanstellen van leraren).
06. Het toelaten van leerlingen op de school (toelaten van leerlingen).
07. Het benoemen van nieuwe leraren (benoemen leraren).
08. Het verdelen van lesoverstijgende taken (taakverdeling).
09. De contacten van de school met het beroepenveld (beroepenveld).
10. Het werven van stageplaatsen voor de leerlingen (werven stages).
11. Het ontslaan van leraren (ontslaan van leraren).
12. Het besteden van de vrije ruimte in de lessentabel (lessentabel).
13. Het verdelen van leraarslessen (leraarslessen).
14. Het beoordelen van leraren (beoordelen van leraren).
15. Afspraken over invoering van de kwalificatiestructuur (invoering K.S.).
16. Het fuseren met andere scholen (fuseren).

Figuur 2. Overzicht van de taken die zijn onderscheiden om inzicht te krijgen in de betrokkenheid van leraren bij het beheersmatig beleid van hun school

01. Samenwerken met collega-leraren is noodzakelijk om lestaken goed te kunnen uitvoeren (samenwerken met collega's).
02. Het is voor leraren belangrijk dat zij zich blijven scholen (blijvend scholen).
03. Schoolbeleid is niet alleen een zaak van het management (betrokken zijn bij beleid).
04. Vakoverstijgende taken moeten deel uitmaken van het takenpakket van leraren (vakoverstijgende taken).
05. Het bijhouden van vakliteratuur is voor leraren van essentieel belang (vakliteratuur bijhouden).
06. Individuele leraren kunnen niet alleen bepalen welke didactische werkvormen worden gehanteerd in hun klassen. Op schoolniveau behoren daar afspraken over gemaakt te worden (afspraken over didactiek).
07. Het is belangrijk dat een leraar met collega's over zijn/haar manier van lesgeven spreekt (spreken over lesgeven).
08. Over het pedagogisch klimaat in klassen dienen op schoolniveau afspraken gemaakt te worden (afspraken pedagogisch klimaat).
09. Individuele leraren behoren niet alleen te bepalen welke leerstof ze tijdens hun lessen aanbieden. Op schoolniveau behoren daarover afspraken gemaakt te worden (afspraken over leerstof).
10. Leraren behoren nieuwe onderwijskundige inzichten te verwerken in hun lesgevendende activiteiten (onderwijskundige inzichten).
11. Ook voor leraren met ruime ervaring blijven nieuwe onderwijskundige theorieën van belang (theorie voor ervaren leraren).
12. Het is belangrijk dat leraren hun didactische activiteiten vergelijken met didactische methoden die in onderzoek effectief zijn gebleken (didactiek vergelijken).
13. Leraren moeten niet alleen beoordeeld worden op de kwaliteit van hun lesgevendende activiteiten (beoordelen leraren).

Figuur 3. Items die aan leraren zijn voorgelegd voor het verkrijgen van inzicht in hun professionele oriëntatie

schoolbeleid. Gebruik is gemaakt van het door Sleegers (1991) ontwikkelde onderzoeksinstrument; dit instrument is aangepast aan de situatie in het voorbereidend en middelbaar beroepsonderwijs (zie Figuren 1 en 2).

De wijze waarop Hoyle (1980) leraren met een beperkte en uitgebreide professionele oriëntatie karakteriseert, vormt de basis voor de vragenlijst die is ontwikkeld om inzicht te krijgen in de professionele oriëntatie van de leraren. Op basis van de beschrijving van Hoyle zijn 13 variabelen geselecteerd en verwerkt in uitspraken of items die aan de leraren zijn voorgelegd (zie Figuur 3).

Analyse en verwerking van resultaten

De betrokkenheid van leraren bij het onderwijskundig en beheersmatig beleid is direct gemeten op respondentniveau. Aan de leraren is gevraagd om op een vierpuntsschaal aan te geven in hoeverre zij bij de onderscheiden taken betrokken zijn (oplopend van niet betrokken tot in grote mate betrokken). Door te middelen over de geselecteerde items zijn de variabelen 'betrokkenheid bij het onderwijskundig beleid' en 'betrokkenheid bij het beheersmatig beleid' samengesteld. Daarna is de onderlinge correlatie tussen deze twee variabelen onder-

zocht. Er was sprake van een redelijk hoge positieve correlatie ($r = .72$), zodat deze twee variabelen zijn samengevoegd tot de variabele 'betrokkenheid van leraren bij het schoolbeleid'.

De professionele oriëntatie van leraren is ook direct gemeten op respondentniveau. Aan de leraren is gevraagd op een vierpuntsschaal aan te geven in hoeverre ze het met de 13 uitspraken eens zijn (oplopend van niet mee eens tot volledig mee eens). 'Niet mee eens' wordt in dit onderzoek opgevat als een kenmerk van een leraar met een beperkte professionele oriëntatie en 'volledig mee eens' als een eigenschap van een leraar met een uitgebreide professionele oriëntatie. Door te middelen over de geselecteerde items is de variabele 'professionele oriëntatie' samengesteld. Ten einde de interne consistentie van de hiervoor genoemde (samengestelde) variabelen te bepalen, zijn alpha-waarden berekend (zie Tabel 1). De samengestelde variabelen kennen een hoge betrouwbaarheid; er zijn geen items (variabelen) gevonden die de homogeniteit van de schalen verlagen.

Tabel 1
Alpha-waarden van de samengestelde variabelen

variabele	Aantal items	Alpha-waarden
01. betrokkenheid bij onderwijskundig beleid	7	.92
02. betrokkenheid bij beheersmatig beleid	16	.98
03. betrokkenheid bij schoolbeleid	23	.98
04. professionele oriëntatie van leraren	13	.82

3 Resultaten

Betrokkenheid bij het schoolbeleid

Uit de resultaten blijkt dat de leraren beperkt betrokken zijn bij het onderwijskundig beleid van hun school ($N=780$, $\bar{x}=2.17$, $sd=.64$) en zeer beperkt betrokken zijn bij het beheersmatig beleid van hun school ($N=681$, $\bar{x}=1.45$, $sd=.47$). Een vergelijking van de gemiddelden op de variabelen 'betrokkenheid bij het onderwijskundig beleid' en 'betrokkenheid bij het beheersmatig beleid' laat zien dat leraren meer betrokken zijn bij het onderwijskundig beleid dan bij het beheersmatig beleid van hun school ($df=587$, $t=34.02$, $p<.001$ tweezijdig). Eveneens is gebleken dat leraren die het meest betrokken zijn bij het onderwijskundig beleid ook het meest betrokken zijn bij het beheersmatig beleid ($N=588$, $r=.72$, $p<.001$ tweezijdig). Beide variabelen zijn samengevoegd tot de nieuwe variabele 'betrokkenheid bij het schoolbeleid'. De onderzoeksresultaten tonen aan dat leraren in het voortgezet agrarisch onderwijs beperkt betrokken zijn bij het beleid van hun school ($N=588$, $\bar{x}=1.78$, $sd=.52$). Nadere analyses wijzen uit dat op geen van de onderzochte scholen de leraren in redelijke mate of in hoge mate bij het schoolbeleid zijn betrokken.

Onderscheid in professionele oriëntatie

Uit het onderzoek blijkt dat leraren in het voortgezet agrarisch onderwijs beschouwd

kunnen worden als leraren met een redelijk uitgebreide professionele oriëntatie ($N=1030$, $\bar{x}=3.14$, $sd=.39$). Om na te kunnen gaan of er een relatie bestaat tussen de professionele oriëntatie van leraren en hun betrokkenheid bij het schoolbeleid, was het noodzakelijk leraren te onderscheiden naar verschillen in hun professionele oriëntatie. Hiertoe is gebruik gemaakt van de quick-cluster techniek; een factor-analyse leverde geen interpreteerbare onderverdeling op. Het resultaat van deze analyse is in Tabel 2 opgenomen.

De resultaten in Tabel 2 laten zien dat de gemiddelden van beide typen leraren op de variabele 'professionele oriëntatie' hoger zijn dan 2.5. Dit betekent dat beide typen leraren een vrij ruime professionele oriëntatie hebben en dus vooral kenmerken bezitten van 'extended professionals' (Hoyle, 1980). De type 1 leraren worden op basis van hun gemiddelde score gekarakteriseerd als leraren met een minder ruime professionele oriëntatie. De type 2 leraren worden in dit onderzoek omschreven als leraren met een ruime professionele oriëntatie. In Tabel 3 is te lezen hoe beide typen leraren hebben gescoord op de onderscheiden variabelen van professionele oriëntatie.

De resultaten in Tabel 3 laten zien dat leraren met een minder ruime professionele oriëntatie enkele beroepsopvattingen hebben die tendere naar een 'restricted professionalism' (Hoyle,

Tabel 2
Aantal respondenten, gemiddelden en standaardafwijkingen op de samengestelde variabele professionele oriëntatie per type leraar (toetsingsprocedure: oneway variantieanalyse)

variabele	Type 1 leraar			Type 2 leraar			F
	N	\bar{x}	sd	N	\bar{x}	sd	
professionele oriëntatie	405	2.77	.26	625	3.38	.24	1466.12*

NB: vierpuntsschaal: hogere waarde= in grotere mate

* $p < .05$

Tabel 3

Aantal respondenten, gemiddelden en standaardafwijkingen op de variabelen van professionele oriëntatie per type leraar (toetsingsprocedure: oneway variantieanalyse)

variabele	minder ruime beroepsoriëntatie			ruime beroepsoriëntatie			F
	N	\bar{x}	sd	N	\bar{x}	sd	
samenwerken met collega's	425	3.11	.81	647	3.70	.53	214.55*
blijvend scholen	427	3.33	.70	646	3.83	.41	208.24*
betrokken zijn bij beleid	425	2.53	1.00	643	2.83	1.04	22.21*
vakoverstijgende taken	422	2.95	1.04	644	3.51	.80	99.54*
vakliteratuur bijhouden	427	3.37	.77	647	3.70	.57	67.56*
afspraken over didactiek	426	2.03	.92	646	2.97	1.01	241.45*
spreken over lesgeven	427	2.74	.78	647	3.58	.58	408.97*
afspraken ped. klimaat	422	2.60	.84	642	3.31	.74	207.27*
afspraken over leerstof	425	2.46	.94	646	3.07	.93	108.41*
onderwijskundige inzichten	425	2.54	.79	644	3.16	.80	154.01*
theorie voor ervaren leraren	425	2.91	.99	645	3.64	.62	222.31*
didactiek vergelijken	419	2.44	.78	641	3.21	.76	258.20*
beoordelen leraren	423	2.96	.95	645	3.38	.85	57.08*

NB: vierpuntsschaal: hogere waarde= in grotere mate

* er is verschil (oneway variantieanalyse, $p < .001$)

1980). Deze leraren zijn namelijk van mening dat aangelegenheden die direct betrekking hebben op de klaspraktijk vooral de verantwoordelijkheid zijn van de individuele leraar. Op schoolniveau afspraken maken over didactische werkvormen die in de klas worden gebruikt en didactische activiteiten vergelijken met didactische methoden die in onderzoek effectief zijn gebleken, worden door deze leraren van beperkt belang gevonden. Eveneens zijn ze van mening dat het vooral de individuele leraar moet zijn die bepaalt welke leerstof aan leerlingen wordt aangeboden.

Leraren met een ruime professionele oriëntatie hebben beroepsopvattingen die in belangrijke mate overeenkomen met die van 'extended professionals' (Hoyle, 1980). Deze leraren scoren op alle variabelen een hoger gemiddelde dan hun collega's met een minder ruime professionele oriëntatie. De twee typen

leraren verschillen vooral van elkaar in hun opvattingen over aangelegenheden die de directe klaspraktijk betreffen. In tegenstelling tot de leraren met een minder ruime professionele oriëntatie zijn de leraren met een ruime professionele oriëntatie van mening dat afspraken op schoolniveau behoren te worden gemaakt over de didactische werkvormen die in de klas worden toegepast en over de leerstof die aan leerlingen wordt aangeboden. Ook vinden de leraren met een ruime professionele oriëntatie, tot op zekere hoogte in tegenstelling tot hun anders georiënteerde collega's, dat leraren hun didactische activiteiten behoren te vergelijken met didactische methoden die in onderzoek effectief zijn gebleken en dat nieuwe onderwijskundige inzichten verwerkt dienen te worden in lesgevende activiteiten.

Tabel 4

Aantal respondenten, gemiddelden en standaardafwijkingen per type leraar op de variabele betrokkenheid bij het schoolbeleid (toetsingsprocedure: oneway variantieanalyse)

variabele	minder ruime beroepsoriëntatie			ruime beroepsoriëntatie			F
	N	\bar{x}	sd	N	\bar{x}	sd	
betrokkenheid bij schoolbeleid	236	1.67	.44	350	1.85	.56	17.32*

NB: vierpuntsschaal: hogere waarde = in grotere mate

* $p < .001$

Tabel 5

Aantal respondenten, gemiddelden en standaardafwijkingen per type leraar op de variabelen betrokkenheid bij het onderwijskundig beleid en betrokkenheid bij het beheersmatig beleid (toetsingsprocedure: one-way variantieanalyse)

variabele	minder ruime beroepsoriëntatie			ruime beroepsoriëntatie			F
	N	\bar{x}	sd	N	\bar{x}	sd	
betrokkenheid bij onderwijskundig beleid	299	2.02	.59	479	2.26	.66	26.70*
betrokkenheid bij beheersmatig beleid	279	1.36	.34	399	1.52	.53	19.54*

NB: vierpuntsschaal: hogere waarde= in grotere mate

* $p < .001$

De relatie tussen de professionele oriëntatie van leraren en de betrokkenheid bij het schoolbeleid

Voor het beantwoorden van de vraag naar de relatie tussen de professionele oriëntatie van leraren en de mate van betrokkenheid bij het schoolbeleid is onderzocht of leraren met een minder ruime professionele oriëntatie verschillen van leraren met een ruime professionele oriëntatie in de mate waarin ze zijn betrokken bij het schoolbeleid (zie Tabel 4).

Uit de resultaten blijkt dat leraren met een ruime professionele oriëntatie meer betrokken zijn bij het beleid van hun school dan hun collega's met een minder ruime professionele oriëntatie. In Tabel 5 wordt duidelijk dat het verschil in betrokkenheid tussen de twee typen leraren zich voordoet binnen beide beleidsdomeinen. Ook blijkt uit deze tabel dat het verschil in betrokkenheid het grootst is binnen het onderwijskundig domein.

Uit Tabel 5 blijkt dat beide typen leraren

beperkt betrokken zijn bij het onderwijskundig beleid en zeer beperkt betrokken zijn bij het beheersmatig beleid van de school waar ze werkzaam zijn. De gemiddelden op de samengestelde variabele 'betrokkenheid bij het schoolbeleid' (Tabel 4) geven aan dat leraren met een minder ruime professionele oriëntatie zeer beperkt betrokken zijn bij het schoolbeleid en leraren met een ruime professionele oriëntatie beperkt betrokken zijn bij het beleid van hun school. Het antwoord op de onderzoeksvraag luidt dan ook: met grote mate van zekerheid blijkt in het voortgezet agrarisch onderwijs een relatie te bestaan tussen de professionele oriëntatie van leraren en hun betrokkenheid bij het schoolbeleid. Bij 18 van de 23 onderscheiden taken op onderwijskundig en beheersmatig gebied zijn de leraren met een ruime professionele oriëntatie in grotere mate betrokken dan hun collega's met een minder ruime professionele oriëntatie (zie Tabellen 6 en 7).

Tabel 6

Aantal respondenten, gemiddelden en standaardafwijkingen per type leraar op de variabelen van betrokkenheid bij onderwijskundig beleid (toetsingsprocedure: one-way variantieanalyse)

variabele	minder ruime beroepsoriëntatie			ruime beroepsoriëntatie			F
	N	\bar{x}	sd	N	\bar{x}	sd	
didactiek	369	1.84	.88	568	2.18	.94	29.67*
leerinhoud	377	2.03	.86	581	2.23	.95	9.69*
toetsen	375	2.31	.91	579	2.47	.94	7.15*
stage-inhoud	377	1.92	1.10	576	2.00	1.09	1.08
begeleiding leerlingen	406	2.19	.92	618	2.52	1.01	29.19*
doelen	405	1.71	.82	619	2.00	.93	27.07*
leermiddelen	412	2.39	.88	626	2.41	.85	.19

NB: vierpuntsschaal: hogere waarde= in grotere mate

* $p < .001$

Tabel 7

Aantal respondenten, gemiddelden en standaardafwijkingen per type leraar op de variabelen van betrokkenheid bij beheersmatig beleid (toetsingsprocedure: oneway variantieanalyse)

variabele	minder ruime beroepsoriëntatie			ruime beroepsoriëntatie			F
	N	\bar{x}	sd	N	\bar{x}	sd	
contacten	404	1.55	.74	619	1.76	.89	15.66*
aanstellen OOP personeel	407	1.15	.48	616	1.24	.63	6.47*
inspectie	404	1.09	.35	619	1.21	.55	16.27*
financiën	407	1.41	.73	624	1.47	.75	1.92
vast aanstellen van leraren	409	1.21	.55	620	1.29	.67	4.88*
toelaten van leerlingen	411	1.57	.88	630	1.95	1.11	34.25*
benoemen van leraren	407	1.26	.59	619	1.39	.75	8.58*
taakverdeling	406	1.52	.72	620	1.70	.88	10.92*
beroepenveld	387	1.92	1.00	591	2.00	1.02	1.47
werven stages	373	1.74	1.05	552	1.77	1.06	.16
ontslaan van leraren	362	1.07	.34	543	1.16	.56	8.22*
lessentabel	410	1.53	.78	621	1.79	.94	21.84*
leraarslessen	410	1.38	.74	622	1.61	.90	17.47*
beoordelen van leraren	366	1.16	.48	575	1.33	.69	15.88*
invoering K.S.	355	1.78	.92	536	1.95	.98	6.47*
fuseren	352	1.30	.58	512	1.43	.76	7.07*

NB: vierpuntsschaal: hogere waarde= in grotere mate

* $p < .05$

4 Conclusies en discussie

Uit het onderzoek blijkt dat leraren die werken in het voortgezet agrarisch onderwijs beperkt betrokken zijn bij het onderwijskundig beleid en zeer beperkt betrokken zijn bij het beheersmatig beleid van hun school. Zeer opvallend is de beperkte betrokkenheid van de leraren bij het onderwijskundig beleid. In de literatuur wordt er namelijk op gewezen dat leraren binnen het onderwijskundig domein van de school over het algemeen de grootste zeggenschap hebben. Dit laatste wordt ook bevestigd door de onderzoeksresultaten van Slegers (1991). De resultaten van het onderhavige onderzoek wijzen er dan ook op dat in scholen voor voortgezet agrarisch onderwijs beperkt onderwijskundig beleid wordt gevoerd (vgl. ook Jongmans, 1996). De eerste resultaten van een ander onderzoek dat momenteel in het voortgezet agrarisch onderwijs wordt uitgevoerd (zie Beijaard, Biemans & Jongmans, 1995; Biemans & Jongmans, 1995) bevestigen dit vermoeden. De vraag blijft echter of de leraren die werkzaam zijn in het voortgezet agrarisch onderwijs qua betrokkenheid bij het onderwijskundig beleid (en het beheersmatig beleid) verschillen van de leraren die werkzaam zijn op de scholen waar Slegers (1991) het onderzoek

heeft uitgevoerd. In het onderzoek van Slegers is namelijk aan de leraren met een coördinatietaak gevraagd aan te geven in welke mate leraren deelnemen aan het schoolbeleid. In het voortgezet agrarisch onderwijs hebben leraren de mate van betrokkenheid bij het schoolbeleid zelf aangegeven. Dit zou kunnen betekenen dat leraren met een coördinatietaak de betrokkenheid van leraren bij het schoolbeleid hoger waarderen dan de leraren zelf en dat, bijgevolg, de relatie tussen de mate van betrokkenheid van leraren bij het schoolbeleid en het beleidsvoerende vermogen van scholen minder sterk is dan nu veelal wordt aangenomen. Nader onderzoek naar deze problematiek is dan ook gewenst.

In dit onderzoek zijn twee typen leraren onderscheiden: leraren met een ruime professionele oriëntatie en leraren met een minder ruime professionele oriëntatie. Leraren met een ruime professionele oriëntatie bezitten vooral kenmerken van 'extended professionals' (Hoyle, 1980). Leraren met een minder ruime professionele oriëntatie hebben zowel beroepsopvattingen die tenderen naar een 'extended professionalism' als naar een 'restricted professionalism' (zie Jongmans & Beijaard, 1995). Met name de beroepsopvattingen van de leraren met een minder ruime professionele oriën-

tatie die overeenkomen met die van de 'restricted professionals' (Hoyle, 1980) geven reden tot enige bezorgdheid. Deze leraren zijn namelijk van mening dat het klaslokaal het domein is van de individuele leraar. Ze neigen ernaar koning(in) te willen zijn in eigen koninkrijk. De veronderstelling is dat deze leraren niet zo gemotiveerd zijn om onderwijsvernieuwingen door te voeren waarbij de rol van de leraar verandert en het klassegebeuren, meer dan voorheen, een gedeelde verantwoordelijkheid wordt (Biemans & Jongmans, 1995). Een ontwikkeling van een minder ruime professionele oriëntatie naar een ruimere professionele oriëntatie van leraren is voor het doorvoeren van onderwijskundige vernieuwingen dan ook van belang.

Uit de onderzoeksresultaten blijkt ook dat leraren met een ruime professionele oriëntatie meer betrokken zijn bij het schoolbeleid dan leraren met een minder ruime professionele oriëntatie. Met grote mate van zekerheid kan worden gesteld dat de betrokkenheid van leraren bij het schoolbeleid en de professionele oriëntatie van leraren aan elkaar gerelateerd zijn. Dit betekent dat ook in het kader van het vergroten van de betrokkenheid van leraren bij het schoolbeleid het gewenst is dat leraren zich ontwikkelen in de richting van een 'extended professionalism'.

Eerder is op de veronderstelling gewezen dat een ontwikkeling van een beperkte professionele oriëntatie naar een meer uitgebreide professionele oriëntatie van leraren binnen scholen tevens een ontwikkeling inhoudt van een hiërarchisch georiënteerd naar een collegiaal georiënteerd beleidsvoerend vermogen van scholen (Sleegers & Bergen, 1992). Dit onderzoek heeft echter aan het licht gebracht dat ook de leraren met een ruime professionele oriëntatie slechts beperkt betrokken zijn bij het schoolbeleid. De veronderstelling in deze studie is dat van de ontwikkeling die Sleegers en Bergen (1992) beschrijven alleen sprake kan zijn wanneer het schoolmanagement bereid is om op het gebied van de schoolorganisatie met de leraren samen te werken. Een ruime professionele oriëntatie van leraren vertaalt zich alleen in een hogere mate van betrokkenheid bij het schoolbeleid wanneer aan de leraren ook de mogelijkheden worden geboden om zich op dit terrein te begeven. In scholen voor voortge-

zet agrarisch onderwijs, zo lijkt het, worden op dit moment aan de leraren nog onvoldoende mogelijkheden geboden om te participeren op beleidsniveau. Uit lopend onderzoek (Biemans & Jongmans, 1995) komt naar voren dat schoolleiders hiertoe wel bereid zijn, maar nog onvoldoende mogelijkheden zien om leraren te betrekken bij het schoolbeleid.

Literatuur

- Biemans, H.J.A., & Jongmans, C.T. (1995). Secondary agricultural education from a 'learning to learn' perspective. *European Journal of Agricultural Education and Extension*, 2(3), 41-50.
- Beijaard, D., Biemans, H.J.A., & Jongmans, C.T. (1995). Leren leren in het Middelbaar Agrarisch Onderwijs. *Agrarisch Onderwijs*, 37(13), 6-9.
- Giesbers, J.H.G.I., & Bergen, Th. (1992). *Professionaliteit en professionalisering van leraren*. Culemborg: Educaboek.
- Giesbers, J., Schuit, H., & Silvertand, R. (1987). Beleidsvorming in relatief autonome schoolorganisaties. In Th. Bergen, J. Giesbers & C. Morsch (red.), *Professionalisering van onderwijsgeven* (pp. 83-97). Lisse: Swets & Zeitlinger.
- Hanson, M.E. (1979). *Educational Administration and Organizational Behavior*. Boston: Allyn and Bacon.
- Hoyle, E. (1975). Leadership and decision-making in education. In M.G. Hughes (Ed.), *Administering Education: International Challenge* (pp. 30-44). London: The Athlone Press.
- Hoyle, E. (1980). Professionalization and deprofessionalization in education. In E. Hoyle & J. Megarry (Eds.), *World Yearbook of Education 1980* (pp. 42-57). London: Kogan Page.
- Jongmans, C.T. (1996). In-service teacher training for agricultural education in the Netherlands: from supply to demand. *International Journal of Educational Development*, 16(1), 89-95.
- Jongmans, C.T., & Beijaard, D. (1995). *Schoolbeleid en professionalisering van leraren*. Wageningen: Landbouwwuniversiteit.
- Knoers, A.M.P. (1994). De leraar in de autonome school. *Pedagogische Studiën*, 71, 187-199.
- Marx, E. (1987). Vermogen van scholen tot het voeren van bestuurlijk beleid. In L. Genemans (red.), *Autonomie van scholen en deregulering* (pp. 7-27). Nijmegen: ITS.

Marx, E. (1988). De uitkomsten van de OTO-gevalstudies, organisatorisch bezien. *Meso*, 8(38), 20-32.

Marx, E., Vries, O.A. de, Veenman, S.A.M., & Slegers, P. (1995). School en klas. In J. Lowyck & N. Verloop (red.), *Onderwijskunde: een kennisbasis voor professionals* (pp. 43-77). Groningen: Wolters-Noordhoff.

Schuit, H., & Slegers, P. (1990). De schoolorganisatiemodellen van Marx als alternatief voor de dichotome benadering van scholen. In J. Giesbers & F. van der Krogt (red.), *Werken met modellen in organisaties* (pp. 75-86). Alphen a/d Rijn: Samsom/ H.D. Tjeenk Willink.

Slegers, P.J.C. (1991). *School en beleidsvoering. Een onderzoek naar de relatie tussen het beleidsvoerende vermogen van scholen en het benutten van de beleidsruimte door scholen*. Nijmegen: Katholieke Universiteit.

Slegers, P., & Bergen, Th. (1992). De professionele oriëntatie van docenten en de beleidsvoering door scholen. *VELON Tijdschrift*, 13(2), 5-9.

Manuscript aanvaard 11-4-1996

Auteurs

C.T. Jongmans is als onderzoeker werkzaam bij de vakgroep Agrarische Onderwijskunde van de Landbouwniversiteit Wageningen.

D. Beijaard is als universitair hoofddocent werkzaam bij het Interfacultair Centrum voor Lerarenopleiding, Onderwijsontwikkeling en Nascholing (ICLON) van de Rijksuniversiteit Leiden.

Correspondentie-adres: Landbouwniversiteit, vakgroep Agrarische Onderwijskunde, Postbus 8130, 6700 EW Wageningen

Abstract

Teachers' professional orientation and their involvement in school policy-making.

C.T. Jongmans & D. Beijaard. *Pedagogische Studiën*, 1997, 74, 97-107.

School's policy-making capacity appears to be highly determined by teachers' involvement in school policy. Research has shown that teachers' involvement in policy-making differs. Dutch literature suggests that teachers' involvement in school policy-making and their professional orientation are related. This research study examines this supposition. For this purpose, 1079 teachers from 98 schools for junior and secondary agricultural education filled in a questionnaire. The study reveals two main types of teachers: teachers with an extended professionalism ($N=639$) and teachers with a less extended professionalism ($N=418$). The study shows that extended professionals are more involved in school policy-making than less extended professionals. Both types of teachers, however, are not very engaged in school policy-making. These days schools need extended professionals. Extended professionalism, however, only leads to higher involvement of teachers in school policy-making, when school management creates the necessary conditions for this.

Samenvatting

Uit longitudinaal onderzoek naar de determinanten van (exacte) vakkenkeuzen blijkt dat voor jongens en meisjes dezelfde factoren samenhangen met de keuze voor exacte vakken. Het gaat daarbij met name om de vraag of men deze vakken leuk, interessant en gemakkelijk vindt en om de geschiktheid volgens relevante anderen. In het vervolgonderzoek waarover dit artikel rapporteert, is nagegaan of er in de tijd tussen de vakkenkeuze en het eindexamen seksspecifieke ontwikkelingen zijn waar te nemen. Meisjes laten – indien zij na de keuze van het vakkenpakket nog veranderingen in dat pakket aanbrengen – niet vaker exacte vakken uit hun pakket vallen dan jongens; als ze dat wel doen geven ze er dezelfde redenen voor op: het gaat vooral om de moeilijkheidsgraad. Toekomstperspectief speelt op oudere leeftijd geen grotere rol dan in de eerste fase van het voortgezet onderwijs. Er is geen sekseverschil in de tevredenheid over het gekozen pakket. Bij de invoering van profielen in plaats van vakkenpakketten zullen sekseverschillen onverminderd blijven bestaan.

1 Inleiding

Alhoewel het Nederlandse onderwijsbeleid eerder gericht is op indamming van de toeloop naar hogere vormen van onderwijs en met name het opleidingsniveau van meisjes de laatste decennia enorm is gestegen, vinden onderzoekers en beleidsmakers (o.a. via het onderwijsmancipatiebeleid) het nog steeds nodig om achterstanden van die groep onder de aandacht te brengen. Eensgezind wordt vastgesteld dat meisjes in minstens gelijke mate als jongens deelnemen aan hogere vormen van onderwijs; zij vormen momenteel bijvoorbeeld bijna de helft van de universitaire studentenpopulatie (Ministerie van Sociale Zaken en Werkgele-

genheid, 1995). Tegelijkertijd echter wordt de inhoud van het door meisjes genoten onderwijs aan de kaak gesteld: doordat zij zich beperken in de vakken en opleidingen die ze kiezen, beperken ze ook hun kansen op een groot aantal economische en maatschappelijke posities. Het nog steeds in veel mindere mate kiezen van vakken als wiskunde B en natuurkunde lijkt op zich een plausibele keuze; niet iedereen heeft dezelfde interesses. Als echter door dergelijke keuzen een groot deel van perspectiefvolle vervolgopleidingen en beroepen niet meer tot de mogelijkheden behoort, moet toch worden gesproken van een achterstand in behaald onderwijsniveau. Op lagere onderwijsniveaus geldt overigens hetzelfde: het systematisch niet kiezen van meisjes voor technische beroepsopleidingen beperkt hun kansen op de arbeidsmarkt.

Er is veel (met name internationaal) onderzoek gedaan naar de achtergronden van seksspecifieke keuzen in het onderwijs. In de eerste paragraaf van dit artikel worden kort de resultaten van dergelijk onderzoek beschreven. In de volgende paragrafen wordt verslag gedaan van een vervolg op een van die onderzoeken, waarin is nagegaan of er tussen de vakkenkeuze en het eindexamen sprake is van een seksetypisch verloop van de afronding van de VO-opleiding. We beperken ons tot het AVO/VWO en laten beroepsopleidingen hier buiten beschouwing.

2 Determinanten van seksspecifieke vakkenkeuzen

Bij de keuze van exacte vakken in het eindexamenpakket treden aanzienlijke verschillen op tussen meisjes en jongens (zie o.a. Dekkers, 1993). Onder meisjes is de belangstelling voor deze vakken beperkt. Zij kiezen weliswaar iets vaker dan jongens voor wiskunde A, maar nog

steeds kiezen jongens in het VWO tweemaal en in het HAVO ongeveer driemaal zoveel wiskunde B en natuurkunde als meisjes. Voor scheikunde liggen de percentages wat dichter bij elkaar. Cijfers uit vele Westeuropese en Amerikaanse onderzoeken tonen in grote lijnen hetzelfde beeld (Walberg, 1991).

Wat zijn de oorzaken van die zo verschillende vakkenkeuzepatronen van jongens en meisjes? Zoals gezegd is dat bepaald geen onderzoeksvraag die de laatste jaren niet aan bod is geweest. Vanuit vele disciplines is gezocht naar verklaringen voor het typische beeld dat de seksen vertonen. Biologische verklaringen, psychologische, sociologische en onderwijskundige verklaringen wisselden elkaar af, en ook modellen die recht deden aan een veelheid van elkaar beïnvloedende factoren (m.n. Eccles c.s., 1985) zijn in onderzoek getoetst. In het longitudinale onderzoek naar determinanten van vakkenkeuzen waarop dit artikel voortbouwt, is een relevante selectie gemaakt uit de literatuur en is getracht in een model de samenhangende invloed van de verschillende factoren te toetsen. Voor de theoretische verantwoording van de variabelen alsook de operationalisatie ervan verwijzen we naar eerdere publicaties over het onderzoek (o.a. Dekkers, 1993; Bosker & Dekkers, 1994; Smeets & Dekkers, 1996). We volstaan hier met een samenvatting van de resultaten.

Noch het ouderlijk milieu, in casu beroep en opleiding van de ouders, noch de samenstelling van het gezin waaruit de leerlingen komen blijkt van invloed op de vakkenkeuze van de leerlingen in het AVO/VWO; vanzelfsprekend zijn er op deze factor ook geen sekseverschillen. Tegen de in de literatuur gewekte verwachtingen in is ook het toekomstperspectief zoals gemeten in plannen voor (de combinatie van) toekomstig werk en gezin geen factor die van invloed is op het wel of niet kiezen van exacte vakken. Wel is hier duidelijk sprake van sekseverschillen, in de zin dat meisjes al op 14/15-jarige leeftijd van plan zijn later minder dan jongens te werken om gezinstaken op zich te kunnen nemen en dat jongens een daarop aansluitend beeld van hun arbeidstoekomst geven. Hetzelfde geldt voor het oordeel dat wiskunde een noodzakelijk of geschikt vak

voor de toekomst is: meisjes vinden dat minder dan jongens, maar het heeft voor beide groepen geen invloed op hun vakkenkeuze. Ook de attributieve variabelen spelen geen rol; wel verklaren meisjes hun falen in het vak wiskunde vaker door de moeilijkheid ervan, en jongens door gebrek aan inspanning, maar ze kiezen het vak er niet meer of minder om. Datzelfde geldt tenslotte voor de verwachte samenhang tussen toetsscores (speciale wiskundetoetsen uit het onderzoek) en vakkenkeuze: er zijn wel sekseverschillen, maar er is minder samenhang dan verwacht met vakkenkeuze. De samenhang van rapportcijfers met vakkenkeuze is groter, maar valt ook weg als een aantal hierna besproken variabelen in het model wordt opgenomen.

Factoren die *wel* invloed lijken te hebben op de keuze van vakken als wiskunde en natuurkunde zijn: een vak leuk, interessant en makkelijk vinden (attitudes), de door de leerling gepercipieerde geschiktheid van het vak volgens relevante anderen (ouders, docenten etc.) en de interesse in een technische of exacte vervolgopleiding (m.n. natuurkunde). Er is geen sprake van sekseverschillen in de samenhang tussen deze predictoren en vakkenkeuze; voor jongens en meisjes geldt die in dezelfde mate. Wel zijn er sekseverschillen in de factoren op zich, zowel in de predictoren als de vakkenkeuze.*De samenhangen zien er dus wel hetzelfde uit voor jongens en meisjes, maar ze liggen op een ander scoreniveau; jongens scoren hoger op de voorspellers en op de keuze voor wiskunde en natuurkunde. Overigens gelden deze bevindingen andersom voor de predictoren en keuze van het vak Frans; meisjes scoren daar hoger, maar de relaties zijn ook hier niet sekse-specifiek.

Aan bovenstaande bevindingen moet nog worden toegevoegd dat ze ook gelden voor veranderingen in de loop van de jaren waarin de leerlingen zijn gevolgd, vanaf de brugklas tot aan het moment waarop het vakkenpakket werd samengesteld: er bleek een samenhang tussen veranderingen op de genoemde predictoren en veranderingen in (voorgenomen) vakkenkeuze. Een vak interessanter gaan vinden bleek bijvoorbeeld samen te hangen met duidelijker plannen om het vak te kiezen.

3 Vraagstelling vervolgonderzoek

Uit de hierboven summier weergegeven resultaten van het longitudinale onderzoek naar determinanten van (exacte) vakkenkeuze blijkt dat een aantal factoren zowel voor jongens als voor meisjes samenhangt met de keuze van exacte vakken. Enkele andere factoren spelen tegen de verwachting in nauwelijks een rol. Bovendien blijkt dat er in de eerste jaren van het voortgezet onderwijs sprake is van ontwikkelingen op zowel de predictoren als de voorgenomen vakkenkeuze; de ontwikkeling bij beide gaat gelijk op. Om na te gaan of dergelijke ontwikkelingen ook na de samenstelling van het vakkenpakket optreden, is vervolgonderzoek uitgevoerd (zie Smeets & Dekkers, 1996). Daarbij is onder meer aandacht besteed aan eventuele veranderingen in predictoren en/of in het vakkenpakket, eindexamenresultaten en de (voorgenomen) keuze van een vervolgonopleiding. De exacte vakken en eventuele sekseverschillen staan hierbij centraal. In verband met de beoogde vervanging van de vele vakkenpakketten in de tweede fase van het voortgezet onderwijs door vier profielen (Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs, 1994), is ook dat thema in het onderzoek aangesneden.

De onderzoeksvragen waarop in dit artikel wordt ingegaan, luiden als volgt:

- 1 Hebben de leerlingen na de samenstelling van hun vakkenpakket nog veranderingen in dat pakket aangebracht? Zo ja, welke factoren en/of personen hebben daarbij een rol gespeeld?
- 2 In hoeverre zijn er sinds de derde klas veranderingen opgetreden in de houding van de leerlingen ten aanzien van het vak wiskunde?
- 3 In hoeverre zijn er sinds de derde klas veranderingen opgetreden in de toekomstverwachtingen van de leerlingen?
- 4 Hoe zijn de eindexamenresultaten van de leerlingen bij de vakken wiskunde (A en B), natuurkunde en scheikunde?
- 5 Hoe oordelen de leerlingen achteraf over het vakkenpakket waarin zij eindexamen doen of hebben gedaan?
- 6 Hoe zou de profielkeuze van de leerlingen eruit hebben gezien?

4 Opzet vervolgonderzoek

4.1 Onderzoeksgroep

In het oorspronkelijke onderzoek is een cohort leerlingen gevolgd vanaf de brugklas (1990) tot aan het moment van vakkenpakketkeuze (eind tweede klas voor MAVO, eind derde klas voor HAVO, eind vierde klas voor VWO). Deze leerlingen hebben tijdens drie achtereenvolgende schooljaren een vragenlijst ingevuld en een wiskundetoets gemaakt. De HAVO-leerlingen uit het onderzoek hebben – voor zover zij geen vertraging hebben opgelopen – in 1994 eindexamen gedaan, de onvertraagde VWO-leerlingen in 1995. Ten behoeve van het vervolgonderzoek (in 1995) zijn twee vragenlijsten ontwikkeld: een vragenlijst voor schoolverlaters en een vragenlijst voor scholieren. Er is getracht alle HAVO- en VWO-leerlingen waarvan de definitieve vakkenkeuze bekend was en waarvan bovendien de andere gegevens uit eerdere metingen beschikbaar waren, te benaderen met het verzoek een vragenlijst in te vullen. Het ging daarbij om 808 leerlingen, verdeeld over achttien scholen. Bij de schooladministraties is nagegaan of deze leerlingen nog op de school in kwestie zaten of de school inmiddels hadden verlaten. Zeventien scholen verleenden hun medewerking aan het onderzoek. Hierdoor konden 748 leerlingen op hun (laatst bekende) huisadres worden aangeschreven. De vragenlijst voor schoolverlaters is aan 171 oud-leerlingen opgestuurd. De respons bedroeg 114, oftewel 69 procent. De respons bij de scholieren bedroeg 76 procent: hier werden 439 van de 577 enquêteformulieren ingevuld geretourneerd. Over het geheel genomen werd driekwart van de verzonden vragenlijsten ingevuld.

Bij de aan het onderzoek deelnemende scholen zijn bovendien de eindexamenresultaten opgevraagd van alle HAVO- en VWO-leerlingen die in 1994 of 1995 examen hebben gedaan. Op deze manier konden de vakkenkeuze van de desbetreffende jaargroepen en de eindexamenresultaten van de leerlingen uit het onderzoek worden geïnventariseerd. Vijftien van de zeventien scholen hebben deze gegevens beschikbaar gesteld.

4.2 Instrumenten

Zoals reeds aangegeven, zijn er twee vragen-

lijsten ontwikkeld: een vragenlijst voor schoolverlaters en een vragenlijst voor leerlingen die nog op school zaten op het moment van enquêteren. Deze vragenlijsten komen grotendeels met elkaar overeen. Een deel van de opgenomen vragen is gelijk aan of vormt een bewerking van vragen uit de enquête die eerder is gebruikt en die in het schooljaar 1991/92 voor het laatst aan de leerlingen is voorgelegd.

De vragenlijsten omvatten de volgende groepen van vragen:

- achtergrondgegevens (geslacht, klas en schoolsoort/eventuele vervolgopleiding);
- attitudes ten aanzien van wiskunde en Frans (negen items per vak, te scoren op negenpuntsschalen; zie Dekkers, 1993);
- laatste rapportcijfer/cijfer schoolonderzoek voor o.a. de exacte vakken;
- veranderingen in het vakkenpakket *na de pakketkeuze* (vakken die men heeft laten vallen en vakken die men eventueel daarvoor in de plaats heeft gekozen);

- factoren die hebben geleid tot het laten vallen van wiskunde, natuurkunde en/of scheikunde (niet nodig voor vervolgopleiding; niet nodig voor toekomstig beroep; moeilijk vak; vak niet leuk; was slecht in dat vak; op advies van anderen laten vallen. Te scoren op vier-puntsschalen);
- personen die invloed hebben gehad op het laten vallen van wiskunde, natuurkunde en/of scheikunde (ouders; broers/zussen; vrienden/vriendinnen; leraar/leraren; schooldecaan. Te scoren op drie-puntsschalen);
- verwachtingen en plannen voor de toekomst (ten aanzien van de vervolgopleiding; aantal dagen werken; combinatie van werk en gezin);
- oordeel over het gekozen vakkenpakket (weer hetzelfde pakket kiezen? zo nee, waarom niet? welke vakken niet meer kiezen en welke in plaats daarvan?);
- keuze van profielen (te scoren op driepuntsschalen plus categorie 'weet niet').

Vraag	Afhankelijke variabele	Onafh. variabelen	Analyse
1a	exacte vakken laten vallen	geslacht verandering attitude wisk. laatste rapportcijfer verandering toekomstverwachting (werk)	correlaties
1b	redenen om exacte vakken te laten vallen	geslacht	variantie-analyses
1c	invloed van personen bij laten vallen exacte vakken	geslacht	variantie-analyses
2	verandering attitude t.o.v. wiskunde (tussen '92 en '95)	geslacht schoolsoort eindexamencijfer of laatste rapp.cijfer wiskunde A in pakket wiskunde B in pakket	multiple regressie-analyses
3	verandering toekomstverwachtingen (tussen '92 en '95)	geslacht schoolsoort	variantie-analyses
4	eindexamenresultaten wisk. A/B, natuurk., scheik.	geslacht schoolsoort	variantie-analyses
5	oordeel over vakkenpakket	geslacht scholier/schoolverlater schoolsoort	variantie-analyses
6	profielkeuze	geslacht scholier/schoolverlater schoolsoort	Chi ²

4.3 Analyse

Ten behoeve van de analyses zijn verschillende databestanden aan elkaar gekoppeld. Allereerst is op basis van de antwoorden op de vragenlijsten voor schoolverlaters en de vragenlijst voor scholieren een verzamelbestand gemaakt. Hierdoor werd het mogelijk de met elkaar corresponderende vragen gezamenlijk in de analyses te betrekken. Uit de van de scholen verkregen eindexamenresultaten zijn de resultaten van de respondenten bij de exacte vakken en Frans geselecteerd en aan het vragenlijstbestand gekoppeld. Daarnaast is een koppeling gemaakt met het in 1992 opgebouwde bestand met resultaten van de toen ingevulde vragenlijsten en met informatie over de samenstelling van de vakkenpakketten.

Op pagina 111 staat per onderzoeksvraag een overzicht van de afhankelijke en onafhankelijke variabelen en de uitgevoerde analyses.

Van significante samenhangen wordt slechts melding gemaakt indien zij minimaal drie procent van de variantie in de desbetreffende afhankelijke variabele verklaren.

5 Resultaten

5.1 Veranderingen in het vakkenpakket

Een verhoudingsgewijs groot deel van de onderzoeksgroep, ruim een kwart, zegt na de oorspronkelijke samenstelling van het vakkenpakket nog *wijzigingen* te hebben aangebracht in dat pakket. Op dit punt treedt geen noemenswaardig verschil tussen meisjes en jongens op. De vakken die men het meest heeft laten vallen, zijn natuurkunde (6 procent), scheikunde (5 procent) en Frans (5 procent). Wiskunde A blijkt relatief vaak in een later stadium in het vakkenpakket te worden opgenomen, al is dat vaak ter vervanging van wiskunde B, dat men dan laat vallen. De veranderingen die achteraf in het vakkenpakket worden aangebracht, laten geen wezenlijke sekseverschillen zien. Dit houdt tevens in dat de verschillen op dat gebied die al bij de oorspronkelijke samenstelling van het vakkenpakket bestonden, gehandhaafd blijven.

De belangrijkste *reden* om wiskunde (A of B), natuurkunde of scheikunde te laten vallen uit het oorspronkelijk samengestelde vakkenpakket, is – zoals de leerlingen zelf aangeven –

de moeilijkheidsgraad van deze vakken. Dat geldt zowel voor jongens als voor meisjes. Ook het 'slecht zijn' in deze vakken, dat met de moeilijkheidsgraad samenhangt, speelt volgens de leerlingen een belangrijke rol bij het besluit om exacte vakken te laten vallen. Noch de vraag of men het desbetreffende vak leuk vindt, noch de opvattingen over het nut van dat vak voor een beoogde vervolgopleiding of voor een toekomstig beroep, spelen een wezenlijke rol bij het na de 'definitieve' vakkenpakketkeuze laten vallen van de bèta-vakken. Hetzelfde geldt voor de invloed van relevante anderen. Ouders, leraren en de schooldecaan hebben doorgaans slechts een kleine invloed op die beslissing. Bij meisjes telt alleen het advies van anderen om scheikunde te laten vallen significant zwaarder mee dan bij jongens ($F=5.37$; $p<.05$). Vooral de invloed van leraren is hier groter ($F=4.60$; $p<.05$).

De resultaten van de analyses geven geen aanleiding tot de veronderstelling dat veranderingen in eerder gevonden predictoren voor wiskundekeuze samenhangen met veranderingen – na het eigenlijke keuzemoment – in het vakkenpakket ten aanzien van de exacte vakken. Veranderingen in de verwachting omtrent het in de toekomst gaan werken of veranderingen in het gewenste aantal dagen dat men wil gaan werken, correleren evenmin significant met veranderingen in het vakkenpakket wat de exacte vakken betreft. Deze verwachtingen correleerden overigens evenmin met de oorspronkelijk gemaakte keuzen, bij de samenstelling van het vakkenpakket.

5.2 Attitude ten opzichte van wiskunde

In 1992, aan het eind van de derde klas HAVO of VWO, werden al belangrijke verschillen tussen meisjes en jongens geconstateerd in de *houding* ten opzichte van wiskunde. Zo vonden jongens wiskunde gemakkelijker, interessanter en belangrijker met het oog op de toekomst dan meisjes. De in 1995 bij dezelfde leerlingen verzamelde gegevens laten hetzelfde beeld zien.

Uit multiple regressie-analyses komt naar voren dat het geslacht geen significante bijdrage levert aan het voorspellen van verschillen in attitudes ten aanzien van wiskunde die in de tweede fase van het voortgezet onderwijs ontstaan. Factoren die wel van belang zijn, zijn de aanwezigheid van wiskunde (A en/of B) in het

vakkenpakket, de prestaties in wiskunde en – in een enkel geval – de schoolsoort. Zo beoordeelen leerlingen die wiskunde A in hun pakket hebben en leerlingen die hogere cijfers voor wiskunde halen, wiskunde in 1995 doorgaans als gemakkelijker dan in 1992 ($R^2=.14$, respectievelijk .28). Daar staat tegenover dat VWO-leerlingen en leerlingen die wiskunde B hebben gekozen, wiskunde gemiddeld moeilijker zijn gaan vinden ($R^2=.06$, respectievelijk .03). Als leerlingen hogere cijfers voor wiskunde halen en/of wiskunde B in hun vakkenpakket hebben, vinden zij wiskunde in 1995 gemiddeld leuker dan in 1992 ($R^2=.14$, respectievelijk .07). Leerlingen die wiskunde A hebben gekozen, zijn wiskunde interessanter gaan vinden ($R^2=.03$). Een hoger cijfer en de aanwezigheid van zowel wiskunde A als wiskunde B in het vakkenpakket hangen samen met de opvatting dat wiskunde de meer 'echt iets voor mij' is geworden (met respectievelijke R^2 -waarden van .15, .04 en .08). Een hoger cijfer voor wiskunde, tot slot, gaat samen met een minder zenuwachtige houding ten aanzien van wiskunde ($R^2=.13$).

5.3 Toekomstverwachtingen

Zowel bij de groep die het HAVO inmiddels met een diploma heeft verlaten, als bij de leerlingen die nog HAVO- of VWO-onderwijs volgen, zijn aanzienlijke verschillen tussen meisjes en jongens te constateren op het punt van de keuze of de voorgenomen keuze van een *vervolgopleiding*. Jongens kiezen aanzienlijk vaker voor technische opleidingen, terwijl meisjes meer belangstelling hebben voor opleidingen in de gezondheidszorg, verzorging/dienstverlening en lerarenopleidingen. De helft van de jongens die na het behalen van het HAVO-diploma een vervolgopleiding zijn gaan doen, heeft gekozen voor een technische richting. Bij de meisjes heeft bijna niemand voor een dergelijke richting gekozen. De economische en administratieve studierichtingen genieten zowel bij jongens als bij meisjes relatief veel belangstelling.

Bij ruim de helft (57 procent) van de groep die in 1992 al een idee had over het type en de soort vervolgopleiding, is dat idee in 1995 nog steeds hetzelfde of is dat idee inmiddels ten uitvoer gebracht. De belangstelling van jongens voor opleidingen in de sfeer van informatica/wiskunde en laboratorium/scheikunde in 1995

blijft aanzienlijk achter bij hetgeen op basis van de in 1992 gegeven antwoorden zou kunnen worden verwacht. Bij meisjes geldt hetzelfde voor jeugd/verzorging, talen, toerisme en horeca. Bij meisjes in het HAVO blijkt er wel meer belangstelling te zijn voor het leraarschap dan men in 1992 had kunnen verwachten.

Driekwart van de geënquêteerde leerlingen verwacht in de toekomst zeker te gaan *werken*. Een vijfde denkt waarschijnlijk een (betaalde) baan te krijgen. Op dit punt zijn er geen wezenlijke sekseverschillen. Bij de vraag naar de gewenste omvang van de toekomstige baan zijn die verschillen er evenmin. Dat beeld verandert echter als men kinderen zou krijgen. Meisjes zouden dan aanzienlijk minder tijd aan betaald werk willen besteden dan jongens. De trend die in 1992 al te zien was, is in 1995 nog steeds aanwezig. Tot het moment waarop zij kinderen krijgen, willen meisjes ongeveer even veel werken als jongens: gemiddeld ongeveer vier dagen per week. Na de geboorte van een kind willen meisjes ruim twee dagen per week werken en verwachten zij van hun partner dat deze drieëneenhalve dag werkt. Jongens geven inderdaad aan dat zij, indien zij kinderen zouden hebben, gemiddeld drieëneenhalve dag zouden willen werken, terwijl ook zij van mening zijn dat hun partner dan ruim twee dagen per week zou moeten werken.

5.4 Eindexamenresultaten

In de eindexamenresultaten bij de bèta-vakken is sprake van sekseverschillen. Tabel 1 toont de gemiddelde eindexamencijfers van de onderzoeksgroep.

De prestaties van de HAVO-leerlingen wijken niet noemenswaard af van de resultaten die de VWO-leerlingen hebben behaald. De verschillen in prestaties tussen meisjes en jongens zijn significant bij wiskunde A ($F=21.43$; $p<.001$), wiskunde B ($F=9.79$; $p<.005$) en scheikunde ($F=16.29$; $p<.001$). Meisjes die eindexamen doen in deze vakken, halen daarvoor gemiddeld een lager cijfer dan jongens. Ook bij natuurkunde is een dergelijke trend te zien.

Dat meisjes ondanks hun geringere deelname aan de exacte vakken in het voortgezet onderwijs ook nog lagere resultaten behalen, werd al eerder aangetoond, onder andere in een

Tabel 1

Eindexamenresultaten in 1994 en 1995, in gemiddeld eindcijfer per vak; naar geslacht en schoolsoort (570 leerlingen)

	--- HAVO 1994 en 1995 ---			----- VWO 1995 -----		
	gemidd.	std.afw.	N	gemidd.	std.afw.	N
<i>meisjes</i>						
Wiskunde A	6.2	0.9	94	6.2	1.2	70
Wiskunde B	5.8	0.9	31	6.2	1.1	48
Natuurkunde	6.0	0.9	24	6.3	1.1	47
Scheikunde	5.9	1.1	33	6.5	1.1	48
<i>jongens</i>						
Wiskunde A	6.8	1.1	53	6.9	1.4	56
Wiskunde B	6.5	0.9	54	6.5	1.3	82
Natuurkunde	6.3	1.0	57	6.6	1.2	78
Scheikunde	6.8	0.9	43	7.0	1.1	63

onderzoek waarin gezocht werd naar reserve aan talent voor technische universiteiten bij meisjes uit VWO-B opleidingen (Dekkers, 1990).

5.5 Oordeel over vakkenpakket

Ruim een derde van de leerlingen en oud-leerlingen zou niet meer hetzelfde vakkenpakket kiezen indien zij nogmaals voor die keuze zouden worden gesteld. Op dit punt treden noch significante verschillen op tussen meisjes en jongens, noch tussen scholieren en schoolverlaters, noch tussen de HAVO- en VWO-groep. De meerderheid daarvan zou één vak uit het huidige pakket willen vervangen door een ander vak. Uiteindelijk zou dat leiden tot een sterkere positie van de vakken wiskunde A (vier procent meer kiezers) en aardrijkskunde (vijf procent meer), terwijl wiskunde B juist (vier procent) minder kiezers zou trekken. Zes procent van de respondenten zou meer exacte vakken kiezen bij een hernieuwde samenstelling van het vakkenpakket. Het gaat hierbij met name om scheikunde en wiskunde A. Daarbij spelen vooral eisen van een gewenste vervolgopleiding en/of uitbreiding van de mogelijkhe-

den een rol. Tegenover deze groep staat een ongeveer even grote groep die juist minder exacte vakken in het pakket zou opnemen. Vooral de moeilijkheidsgraad van deze vakken zou daartoe aanleiding geven.

5.6 Profielkeuze

De leerlingen is – aan de hand van een overzicht op basis van de voorstellen van de Stuurgroep Profiel Tweede Fase (1994) – gevraagd welk profiel zij (waarschijnlijk) zouden hebben gekozen indien de vakkenpakketten zouden zijn afgeschaft. Tabel 2 geeft een overzicht van de antwoorden. Daarbij is de (kleine) groep die niet wist welk profiel te kiezen, buiten beschouwing gelaten.

Uit Tabel 2 blijkt dat er ook na de invoering van profielen duidelijke verschillen in keuze tussen meisjes en jongens zullen blijven bestaan. Bij 'Natuur en techniek' zijn de verschillen het grootst ($Chi^2=106.45$; $p<.001$; $Cramers'V=.45$). Bij veel jongens geniet dit profiel de voorkeur: 43 procent geeft aan zeker voor 'Natuur en techniek' te kiezen als zij in die positie zouden verkeren. Bij de meisjes zou slechts acht procent zeker voor 'Natuur en

Tabel 2

Verwachte keuze van profielen, naar geslacht; 513 respondenten

	----- meisjes -----			----- jongens -----		
	zeker niet	misschien	zeker	zeker niet	misschien	zeker
Natuur en techniek	67 %	25 %	8 %	30 %	26 %	43 %
Natuur en gezondheid	49 %	29 %	22 %	34 %	51 %	15 %
Economie en maatschappij	25 %	47 %	28 %	40 %	34 %	26 %
Cultuur en maatschappij	35 %	40 %	25 %	72 %	19 %	9 %

techniek' kiezen, terwijl twee derde daar zeker niet voor zou kiezen. Bij het profiel 'Natuur en gezondheid' zijn er eveneens belangrijke verschillen naar geslacht ($Chi^2=25.65$; $p<.001$; $Cramers V=.22$), maar het is moeilijk daaruit een bepaalde voorkeur af te leiden. De helft van de jongens zou dit profiel misschien kiezen. Bij de meisjes is de houding ten opzichte van dit profiel meer uitgesproken dan bij de jongens. Zowel de groep die zeker niet voor dit profiel zou kiezen, als de groep die daar zeker wel voor zou kiezen, is bij de meisjes groter dan bij de jongens. 'Economie en maatschappij' geniet wat meer voorkeur bij meisjes, al zijn de verschillen hier niet al te groot. Bij 'Cultuur en maatschappij' zijn de verschillen groter ($Chi^2=69.62$; $p<.001$; $Cramers V=.37$). Dit profiel wordt door de jongens massaal afgewezen. Bijna driekwart zou hier zeker niet voor kiezen. Bij de meisjes zijn de meningen sterk verdeeld.

De variabelen scholier/schoolverlater en schoolsoort vertonen geen noemenswaardige samenhang met de voorkeur voor bepaalde profielen.

6 Tot slot

Is er sprake van seksespecifieke ontwikkelingen in de laatste jaren van het voortgezet onderwijs, met name met betrekking tot het vakkenpakket? Deze vraag lag ten grondslag aan de onderzoeksvragen waarop in dit artikel antwoorden zijn geformuleerd.

Een kwart van de leerlingen sleutelt in de laatste jaren nog aan het vakkenpakket. Natuurkunde, scheikunde en Frans zijn de vakken die men het meest laat vallen. Wiskunde A dient regelmatig als vervanger voor wiskunde B. Sekseverschillen in de patronen van wijziging treden echter niet op.

Behalve het feit dat meisjes aanzienlijk minder exacte vakken in hun eindexamenpakket hebben dan jongens, zowel op het moment van keuze als bij het eindexamen, zijn ook de eindexamencijfers die ze voor die vakken behalen gemiddeld lager; dit geldt zowel voor wiskunde A en wiskunde B als voor scheikunde, en zowel in het HAVO als in het VWO. Ondanks die lagere cijfers laten meisjes tussen het moment waarop zij hun vakkenpakket samen-

stellen en het eindexamen niet vaker exacte vakken uit dat pakket vallen dan jongens.

Toch wordt als belangrijkste reden om wiskunde, natuurkunde of scheikunde wel uit het pakket te laten vallen zowel door jongens als meisjes meestal de moeilijkheidsgraad van het desbetreffende vak opgegeven. Daarmee gaat meestal ook samen dat men zegt 'slecht' te zijn in dat vak. Of men het vak al dan niet leuk en al dan niet nodig voor een vervolgopleiding of beroep vindt, is daarbij in het algemeen van weinig belang. Ook de invloed van 'relevante' anderen of toekomstplannen ten aanzien van de combinatie van werk en gezin hebben geen wezenlijke invloed op het besluit om alsnog een of meer exacte vakken te laten vallen. Ook eventuele veranderingen in houdingen die samenhangen met de oorspronkelijke keuze van exacte vakken in het vakkenpakket zijn niet de oorzaak van wijzigingen. Hetzelfde geldt voor de invloed van schoolse factoren, al lijkt het erop dat meisjes zich bij hun besluit om na de samenstelling van het vakkenpakket een bèta-vak te laten vallen, wat sterker laten leiden door leraren dan jongens.

Er zijn geen belangrijke verschillen tussen meisjes en jongens in tevredenheid over het gekozen vakkenpakket. Bij de invoering van profielen in plaats van vakkenpakketten, zullen sekseverschillen naar verwachting onverminderd blijven bestaan.

De conclusie kan tenslotte worden getrokken dat meisjes, naarmate de keuze van de vervolgopleiding nadert, in het algemeen geen spijt hebben gekregen van het doorgaans beperkte aantal exacte vakken in hun examenpakket. Slechts een klein groepje had liever meer exacte vakken gekozen. Daaruit moet worden afgeleid dat de beperkte aanwezigheid van meisjes in bèta-studies in het algemeen niet kan worden geweten aan een door hen als onjuist gepercipieerde vakkenpakketkeuze. Indien het stimuleren van de keuze van exacte vakken door meisjes op school al tot resultaten leidt, dan houdt dat dus nog niet in dat daarmee de belangstelling voor bèta-studies bij deze groep in aanzienlijke mate zal groeien. Dat laat overigens onverlet dat 'spijt' later in de onderwijs-respectievelijk beroepsloopbaan alsnog kan optreden.

- Bosker, R.J., & Dekkers, H. (1994). School differences in producing gender related subject choices. *School Effectiveness and School Improvement*, 5, 178-195.
- Dekkers, H. (1990). *Seksespecifieke studiekeuzen in het wetenschappelijk onderwijs*. Nijmegen: ITS.
- Dekkers, H. (1993). Determinanten van seksespecifieke vakkenkeuzen in HAVO en VWO. *Pedagogische Studiën*, 70, 445-459.
- Eccles, J., Adler T.F., Futterman, R., Goff, S.B., Kaczala, C.M., Meece, J.L., & Midgley, C. (1985). 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.
- Ministerie van Sociale Zaken en Werkgelegenheid (1995). *Emancipatie in cijfers, 1995*. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.
- Smeets, E., & Dekkers, H. (1996). *Van vakkenkeuze naar eindexamen. Verslag van vervolgonderzoek naar determinanten van seksespecifieke vakkenkeuzen in HAVO en VWO*. Nijmegen: ITS.
- Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs (1994). *De Tweede Fase vernieuwt. Scharnier tussen basisvorming en hoger onderwijs, deel 2*. Den Haag: Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs.
- Walberg, H.J. (1991). Improving school science in advanced and developing countries. *Review of Educational Research*, 61 (1), 25-69.

Behalve de hier vermelde literatuur verwijzen wij voor literatuur m.b.t. de theoretische achtergrond van dit onderzoek naar o.a. Dekkers, 1993.

Manuscript aanvaard 11-8-1996

Auteurs

H. Dekkers is lid van het managementteam van het Instituut voor Toegepaste Sociale wetenschappen en werkzaam bij de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen

E. Smeets is onderzoeker bij het Instituut voor Toegepaste Sociale wetenschappen Nijmegen.

Adres: Instituut voor Toegepaste Sociale wetenschappen, Postbus 9048, 6500 KJ Nijmegen.

Abstract

Between subject choice and final examination: gender differences?

H. Dekkers & E. Smeets. *Pedagogische Studien*, 1997, 74, 108-116.

From longitudinal research of factors determining subject choices, it is concluded that for girls the same factors correlate with the choosing of mathematics and science as school examination subjects as for boys. These subjects are chosen in particular by pupils who like these subjects, and who think these subjects are interesting and easy. Another important factor is the estimation by 'relevant persons' that these subjects are suited for the pupils concerned.

In this article, the results of a study are described, in which we looked into gender specific developments concerning mathematics and science between the moment the examination subjects were chosen, and the moment of the final examination. From this study, conclusions are drawn that girls who alter examination subjects during this period, don't abandon the subjects mathematics or science more than boys do. If girls do abandon these subjects, they do it for the same reasons boys do: the main reason is the complexity of the subjects. When the final examination is approaching, 'future perspective' doesn't play a larger role in this decision than during the first grades in secondary education.

No gender differences were found with regard to the satisfaction with the examination subjects. The planned introduction of four obligatory combinations of subjects instead of separate subjects, won't diminish gender differences in the choosing of mathematics and science.

W. van Elsäcker en L. Verhoeven

Samenvatting

Onderzocht wordt in hoeverre het voorlezen in groepen van 5 kleuters in vergelijking met klassikaal voorlezen leidt tot meer interactie en tot een groter begrip en een betere retentie van het verhaal. Aan het onderzoek deden 19 Nederlandse, 31 Turkse en Marokkaanse en 22 kinderen uit ex-koloniën mee. Deze kinderen kregen een verhaal voorgelezen in twee condities: klassikaal en in groepen van 5 kleuters. Na elk verhaal is een effectmeting gedaan die bestond uit het individueel navertellen en het bevragen van de kinderen na het voorlezen. De resultaten laten zien dat kleuters in kleine groepen meer met elkaar in interactie zijn, hetgeen leidt tot meer begrip en retentie. De gevonden effecten verschillen niet bij allochtone en autochtone kinderen. In de slotparagraaf worden de resultaten van het onderzoek van commentaar voorzien en worden enkele aanbevelingen gegeven met betrekking tot de onderwijspraktijk.

1 Inleiding

Het belang van voorlezen, zowel thuis als op school, wordt algemeen onderkend. Een aantal vaak gehanteerde doelstellingen is:

- uitbreiden van de kennis van de wereld;
- bevorderen van de cognitieve ontwikkeling;
- kennis maken met waarden en normen;
- gevoelens, gedachten en plezier delen die een boek oproepen;
- vergroten van de betrokkenheid en motivatie bij bepaalde onderwerpen.

Uit onderzoek komt naar voren dat de manier waarop kinderen in aanraking komen met (prenten)boeken en andere vormen van geschreven taal van grote invloed is op de verwerving van geletterdheid (Sulzby & Teale, 1991; Verhoeven, 1994). Vooral de kwaliteit van de interactie die tijdens het voorlezen

plaatsvindt tussen kind en volwassene, blijkt een belangrijke voorspellende factor te zijn voor de leesontwikkeling op de basisschool. Voorlezen heeft zowel een gunstige invloed op de taalontwikkeling als op beginnende geletterdheid en leren lezen (vgl. Teale, 1984; Snow et al., 1991; Bus, Van IJzendoorn, Pellegrini & Terpstra, 1994).

Door middel van voorlezen maken kinderen in feite kennis met geschreven taal. Geschreven taal wijkt in veel opzichten af van gesproken taal. Bepaalde woorden en uitdrukkingen worden bijvoorbeeld in gesproken taal vaak vermeden. Ook passieve zinnen komen in spreektaal minder voor. Hiernaast ontbreekt bij een voorleesverhaal de context die er bij oog-in-oog interacties wel is. De schrijver communiceert immers gescheiden in ruimte en tijd met de lezer of toehoorder. Door deze verschillen hebben jonge kinderen in het begin veel hulp nodig om voorleesverhalen te kunnen begrijpen. Tijdens het voorlezen verschaffen opvoeders deze hulp doorgaans door onbewust diverse strategieën toe te passen die het begrip bevorderen. De interactie rondom het verhaal blijkt dus van groot belang voor het leren begrijpen en onthouden van het verhaal. Interactief voorlezen vormt als het ware een brug tussen mondeling en schriftelijk taalgebruik. In de interactie worden tevens de communicatieve vaardigheden geoefend, wat ook weer een positieve invloed heeft op de ontwikkeling van de schriftelijke vaardigheden; mondelinge en schriftelijke vaardigheden hangen immers nauw samen. Vragen als *'Hoe weet je dat?'* en *'Heb jij dat ook wel eens gedaan?'* verhogen niet alleen de betrokkenheid en motivatie van de kinderen maar zijn ook belangrijk voor het leren lezen. Op deze wijze leren kinderen actief met teksten bezig te zijn (Cazden, 1982). Een gevolg hiervan is dat kinderen onder invloed van interactief voorlezen niet alleen hun receptieve vaardigheid in tekstbegrip maar ook hun productieve tekstvaardigheid vergroten. Dit

blijkt onder meer uit een toenemende competentie in het navertellen van een voorgelezen verhaal (Bus, 1990; Notari, O'Connor & Vadasy, 1993).

Kinderen maken in de voorschoolse periode al kennis met diverse aspecten van geletterdheid. In de meeste gezinnen komen wel schriftelijke materialen voor. Wat de kinderen hiervan leren, hangt echter nauw samen met de manier waarop deze materialen in sociale interacties worden gebruikt. Veel onderzoek leidde tot de conclusie dat de ontwikkeling van de geletterdheid positief beïnvloed wordt door zowel de kwantiteit als de kwaliteit van het voorlezen (Ninio, 1980; Teale, 1981; Heath, 1982; Teale & Sulzby, 1987). Teale (1980) kwam op basis van een inventarisatie van eerder onderzoek tot een viertal factoren die de ontwikkeling van de geletterdheid bevorderen: de aanwezigheid van kinderboeken en kindertijdschriften, de leesfrequentie in het gezin, de toegankelijkheid van schrijfmaterialen en de wijze waarop ouders de spontane lees- en schrijfpogingen van het kind ondersteunen. Wells (1985) benadrukte dat met name de waarde die thuis aan geletterdheid wordt gehecht en de stappen die ouders ondernemen om het kind bij zijn/haar ontdekkingen te steunen, cruciaal zijn. Ook een tolerante, waarderende sfeer blijkt van groot belang.

Volwassenen blijken vaak onbewust diverse strategieën te hanteren om bij het voorlezen de taalontwikkeling van het kind te stimuleren, zoals het stellen van open vragen, het vragen om verduidelijking bij de uitingen van het kind en het voortdurend positief waarderen en uitbreiden van de uitingen van het kind. Deze 'scaffolding-strategieën' blijken naast de taalverwerving in het algemeen ook de verwerving van geletterdheid te stimuleren (Bruner, 1977; Snow, 1983). Verder blijkt de interactie tijdens het voorlezen flexibel en in de loop der tijd aan verandering onderhevig. Volwassenen passen hun interactieve stijl aan als de respons van hun kinderen verandert (Heath, 1982; Morrow, 1988; Teale & Sulzby, 1987). Ook in een cross-sectionele studie van Bus en Van IJendoorn (1988) met één-, drie- en vijfjarige kinderen bleek dat moeders hun interactieve stijl aanpassen. Naarmate de kinderen meer weten over geschreven taal en het verhaal dus beter begrijpen, wordt minder vaak ingegaan op de tekst en

op de illustraties en wordt meer aandacht besteed aan de meer formele aspecten van taal.

Naar het voorleesgedrag van allochtone moeders/ouders is nog vrij weinig onderzoek gedaan. Wel kan gesteld worden dat er bij allochtone gezinnen sprake is van grote onderlinge verschillen. Veertig procent van de Turkse en zeventig procent van de Marokkaanse moeders in Nederland is analfabeet (Pels, 1991). Alleen in gezinnen waar de moeder onderwijservaring heeft, lezen de moeders soms voor en tekenen of schrijven ze soms met hun kinderen. Pels stelt in haar onderzoek dat er van een leescultuur in de meeste Marokkaanse gezinnen geen sprake is. Wel is er een orale literaire cultuur. Deze sluit echter minder goed aan bij het Nederlandse onderwijssysteem. Kurvers en Lemmens (1992) vonden dat dit bij 'tweede-generatie moeders' geheel anders ligt. In verreweg de meeste van deze gezinnen zijn boeken aanwezig en leest de moeder regelmatig voor. Voorts vonden Leseman et al. (1995) dat de effecten van gezinsinteractie op de ontwikkeling van allochtone kinderen in belangrijke mate afhankelijk zijn van de gehanteerde thuistaal, de pedagogische denkbeelden en het cultureel-educatief kapitaal in de context van het gezin.

Samenvattend kunnen we stellen dat door middel van interactief voorlezen kennis en vaardigheden bij kinderen worden gestimuleerd die vooral voor het latere lezen en schrijven van belang zijn. De kwaliteit van de interactie tijdens het voorlezen blijkt daarbij een factor van betekenis. Met name binnen allochtone gezinnen blijkt sprake van een grote mate van variatie in voorleesgedrag. Een belangrijke vraag is nu op welke wijze in het kleuteronderwijs door middel van voorlezen optimaal op de taalontwikkeling van kinderen kan worden ingespeeld. De positie van allochtone kleuters die in hun thuissituatie met een andere taal worden geconfronteerd dan het Nederlands verdient daarbij speciale aandacht.

2 Voorlezen in het kleuteronderwijs

Het voorlezen neemt in het dagelijkse programma op de kleuterschool een belangrijke

plaats in. In de meeste groepen wordt tenminste één keer per dag voorgelezen. Bij een onderzoek onder 22 leerkrachten las 86 procent eens per dag voor en 16 procent zelfs twee maal per dag (Bus, 1990). Meestal werd voorgelezen aan de hele groep tegelijk. In slechts vier klassen werd wel eens in kleinere groepjes voorgelezen, maar dit gebeurde slechts incidenteel. Eén leerkracht gaf aan een enkele keer wel eens individueel voor te lezen.

De effecten van voorlezen op school zijn nog niet systematisch onderzocht, evenmin als de effecten van de voorleesstijl van de leerkracht. Er is echter weinig reden om aan te nemen dat de kwaliteit van de interactie op school een minder belangrijke rol zou spelen dan in de thuissituatie. Bus (1990) observeerde het voorleesgedrag van acht leerkrachten in de kleuterbouw. Hierbij bleek dat de leerkrachten globaal in twee groepen waren in te delen: leerkrachten die interactief voorlezen en leerkrachten die dat niet deden. Leerkrachten met een interactieve stijl deden veel meer moeite om de kinderen het verhaal te laten begrijpen. Dit gebeurde bijvoorbeeld door middel van uitnodigen en uitlokken, samenvatten en informeren. De leerkrachten met een niet-interactieve stijl stelden veel minder vragen en als de kinderen vragen stelden, werd hier vaak niet op gereageerd. Oorzaak-gevolg relaties, voorspellen, middel-doel relaties, etc. werden niet geoefend. De vragen die gesteld werden, hadden vaak betrekking op feiten of details en niet op grotere verbanden of op de plot van het verhaal. De meer complexe aspecten van het verhaal kwamen bij deze leerkrachten dus niet aan de orde. Bus constateerde dat het voorleesgedrag van de leerkrachten tamelijk stabiel was; er veranderde niet veel tijdens opeenvolgende leessessies. Het lijkt er dus op dat leerkrachten een bepaalde stijl van voorlezen hebben. Bus acht het niet uitgesloten dat een niet-interactieve manier van voorlezen met name nadelig is voor kinderen met een laag niveau, omdat de leerkracht niet ingaat op de problemen die zij met het begrijpen van de tekst hebben.

Een belangrijke vraag bij het voorlezen op school is de vraag welke groepsomvang het meest ideaal is. Ook hiernaar is tot nu toe nog maar weinig onderzoek verricht. Eén van de weinige onderzoeken naar de effecten van groepsomvang op de kwaliteit van de interacties

tijdens het voorlezen is gedaan door Morrow en Smith (1990). Zij observeerden 27 kinderen en leerkrachten in verschillende kleutergroepen in drie verschillende settings, namelijk tijdens:

- a. het voorlezen aan de hele groep (tenminste 15 kinderen),
- b. het voorlezen in groepjes van drie,
- c. het individueel voorlezen.

Na het voorlezen werd aan de kinderen een aantal vragen gesteld over het verhaal en daarna werd hun gevraagd het verhaal na te vertellen. De uitkomst was opmerkelijk. De kinderen die in groepjes van drie voorgelezen was, waren de prestaties significant beter dan bij het klassikaal voorlezen en bij het individueel voorlezen. In de klassikale setting werd het laagst gescoord. Men zou kunnen verwachten dat de score bij het individueel voorlezen het hoogst zou zijn, maar dit was niet het geval. Deze uitkomst is in overeenstemming met de ideeën van Slavin (1995) dat coöperatief leren een belangrijke bron vormt voor de cognitieve ontwikkeling van kinderen.

Deze uitkomsten hoeven niet per se te betekenen dat voorlezen aan grotere groepen geen zin zou hebben. Bepaalde instructie-strategieën bij het voorlezen aan de hele klas kunnen eveneens leiden tot een beter begrip van het verhaal (Morrow, 1984, 1985). Bovendien hoort ook de ontwikkeling van de vaardigheid om in grotere groepen kennis te verwerven een onderdeel van de opvoeding te zijn. Morrow en Smith (1990) pleiten er dan ook voor om kinderen regelmatig in wisselende settings voor te lezen, zowel klassikaal als in kleinere groepen.

Er is nog geen systematisch onderzoek gedaan naar het effect van het voorlezen aan allochtone kleuters. Both-de Vries en Bus (1992) stellen evenwel dat de gangbare aanpak van het voorlezen in klassen met allochtone kleuters niet toereikend is en dat gezocht moet worden naar andere mogelijkheden. Damhuis (1988) deed onderzoek naar de gelegenheid tot het leren van het Nederlands door Turkse en Marokkaanse kleuters. Zij komt tot de conclusie dat interactie tussen leerkracht en kleuters in kleine groepjes bevorderlijk is voor het leren van het Nederlands als tweede taal. Het onderzoek van Damhuis had geen betrekking op voorlezen maar op andere activiteiten. Het is

evenwel zeer waarschijnlijk dat interactief voorlezen aan kleine groepen allochtone kinderen zowel gunstig is voor de tweede-taalverwerving als voor de verwerving van geletterdheid.

Wanneer we de beschikbare literatuur samenvatten, kunnen we stellen dat voorlezen een positief effect heeft op de verwerving van geletterdheid en dat het effect groter is naarmate er meer interactie plaatsvindt tussen kind en volwassene. In het kleuteronderwijs in Nederland wordt over het algemeen klassikaal voorgelezen. Of dit op interactieve wijze gebeurt of niet, hangt af van de persoonlijke stijl van de leerkracht. Toch zijn niet alle manieren van voorlezen even effectief. Een belangrijke vraag is bijvoorbeeld welke setting de meeste garanties biedt tot interactie met de kinderen. Een andere vraag is of een bepaalde aanpak bij tweedetaalleerders hetzelfde effect heeft als bij eerstaalleerders. Hiernaar is tot nu toe nog geen onderzoek verricht.

Het doel van het onderhavige onderzoek is inzicht te verkrijgen in het effect van de groeps grootte bij het voorlezen aan allochtone en autochtone kleuters in groep 2 van de basisschool. In dit onderzoek is gekozen voor een vergelijking van de klassikale setting met een setting in groepen van 5. In de experimentele conditie is om pragmatische redenen gekozen voor een groeps grootte van 5. Een kleinere groeps grootte lijkt in de huidige onderwijspraktijk moeilijk te realiseren. In het onderzoek is uitgegaan van de volgende hypothesen:

1. Interactief voorlezen aan groepen van 5 leidt zowel bij allochtone als bij autochtone kleuters tot meer interactie dan interactief voorlezen aan de hele groep.
2. Als wordt voorgelezen in groepjes van 5 begripen zowel allochtone als autochtone kinderen het verhaal beter en onthouden ze er meer van dan wanneer klassikaal wordt voorgelezen.

Op voorhand mag worden verwacht dat allochtone kinderen minder hoog zullen uitkomen op taken als tekstbegrip en navertellen. In het licht van genoemde hypothesen rijst echter de vraag in hoeverre de effecten van het lezen in kleine groepen voor allochtone en autochtone kleuters even groot zullen zijn. Enerzijds

lijkt het aannemelijk dat allochtone kinderen, gezien hun geringere vaardigheid in het Nederlands minder vaardig zullen zijn in de verbale interactie dan autochtone kinderen. Anderzijds zouden allochtone kleuters meer baat kunnen hebben van de interactie in kleine groepen vanwege het feit dat ze van huis uit minder ervaring hebben met interactief voorlezen. Uit eerder onderzoek blijkt dat interactie tijdens het voorlezen minder belangrijk wordt naarmate kinderen op dit punt meer ervaren zijn (vgl. Bus et al., 1994; Dickinson & Smith, 1994).

3 Opzet van het onderzoek

De kern van het onderzoek wordt gevormd door twee effectmetingen bij 72 kinderen in groep 2 van de basisschool; één meting na het klassikaal voorlezen van een verhaal en een tweede meting na het voorlezen van een verhaal in kleine groepjes. De effectmetingen bestonden uit het individueel navertellen en het bevragen van de leerlingen na het voorlezen; eerst vertelden de kinderen het verhaal na in hun eigen woorden en daarna beantwoordden ze een aantal begripsvragen naar aanleiding van het verhaal.

3.1 Proefpersonen

Aan het onderzoek namen vier scholen uit de regio Den Haag deel. Deze scholen waren van te voren twee aan twee gematcht op de volgende kenmerken:

- samenstelling van de schoolbevolking,
- sociaal-economische status van de ouders.

School 1 en 3 staan in Rijswijk, beide op een paar honderd meter van de gemeentegrens met Den Haag. Op deze scholen is 45 à 50 procent van de leerlingen allochtoon. Er zijn met name veel Surinaamse kinderen. De meeste allochtone leerlingen en ook een aantal autochtone zijn afkomstig uit de nabijgelegen Haagse Molenvijk. Op school 1 en 3 wordt geen O.E.T.-onderwijs gegeven. De kinderen komen uit gezinnen met een zeer verschillende achtergrond, zowel uit middle-class milieus als uit de lagere sociaal-economische klassen.

School 2 en 4 liggen in de binnenstad van Den Haag, respectievelijk in de wijk Transvaal en in de Schilderswijk. Beide scholen hebben

Tabel 1

De verdeling van de verschillende etnische groepen over de vier scholen

Etn. groep	School 1	School 2	School 3	School 4	Totalen
Turks	1	11	2	11	25
Marokkaans	1	3	0	2	6
Ex-koloniën	7	5	8	2	22
Nederlands	10	0	9	0	19
Totalen:	19	19	19	15	72

99 à 100 procent allochtone (voornamelijk Turkse) leerlingen. De ouders van deze kinderen behoren over het algemeen tot de lagere sociaal-economische klasse. Op beide scholen wordt zowel O.E.T.-Turks als O.E.T.-Arabisch gegeven.

Van de vier geselecteerde proefscholen namen in eerste instantie 80 kinderen uit groep 2 deel aan het onderzoek. Alleen kinderen die minstens een jaar kleuteronderwijs op dezelfde school hadden gevolgd, kwamen in aanmerking. De leeftijd van de kinderen varieerde van 5 tot 6,5 jaar. Zittenblijvers namen niet deel aan het onderzoek. De scores van de kinderen die bij de tweede meting niet aanwezig waren, werden bij de statistische analyses buiten beschouwing gelaten. Uiteindelijk bleven er 72 proefpersonen over; 28 meisjes en 44 jongens, gelijkelijk verdeeld over de vier scholen en over etnische groepen. De verdeling van de verschillende etnische groepen over de vier scholen is af te lezen in Tabel 1.

3.2 Instrumenten

Veel zorg werd besteed aan het selecteren van twee geschikte prentenboekverhalen. Uiteindelijk viel de keus op twee verhalen uit de prentenboeken van het kleuterprogramma 'Schatkist' (Zwijssen, Tilburg):

- Verhaal A: Het Pannekoekemantje (door Frank Herzen)
- Verhaal B: Een politie-auto voor Boy (door Paul van Loon)

Vooraf was geverifieerd of deze verhalen op de proefscholen in het verleden al eens aan de kinderen waren voorgelezen. Dit bleek niet het geval.

Verhaal A:

In het eerste verhaal bakt een oud vrouwtje voor een klein meisje een pannekoekemante-

tje. Dat mannetje wil echter niet opgegeten worden en gaat er vandoor. Hij komt een bakker, een houthakker en een wolf tegen die hem allemaal achtervolgen, samen met het meisje. Een vos biedt aan hem naar de overkant van de rivier te brengen maar halverwege eet hij het pannekoekemantje op, hoewel hij had beweerd geen pannekoekemantjes te lusten.

Verhaal B:

Het tweede verhaal gaat over het jongetje Boy dat met zijn vader in de stad loopt, een mooie politieauto in een etalage ziet en dan met zijn vader het warenhuis binnen gaat om de auto te kopen. Tijdens het spelen op de speelgoedafdeling komt Boy op de roltrap terecht en gaat naar beneden. Hij ziet z'n vader nergens meer en gaat stil in een hoekje zitten huilen. Uiteindelijk brengt een vriendelijke juffrouw hem terug naar zijn vader die op de speelgoedafdeling met treintjes zit te spelen en Boy niet eens gemist heeft.

Het was de bedoeling om tijdens het voorlezen in de klassikale setting en in de kleine groep in gelijke mate gelegenheid te bieden tot interactie. Daartoe werden bij beide verhalen van te voren 9 standaardvragen geformuleerd die in beide condities op vaste plaatsen in de tekst gesteld werden. Daarbij ging het om informatieve open vragen over de setting en het verloop van het verhaal. Daarnaast werd in beide condities gereageerd op initiatief van de kinderen.

Het programma 'Schatkist' biedt bij elk voorleesverhaal tevens vier losse platen met hierop dezelfde illustraties als in de prentenboeken. Deze platen werden bij het navertellen aan de kinderen voorgelegd, bij wijze van geheugensteun.

Ten behoeve van het onderzoek zijn verder twee scoreformulieren (voor verhaal A en verhaal B) ontwikkeld om de kwaliteit van het na-

vertellen te kunnen beoordelen. Beide scoreformulieren bestaan uit 20 items die zijn geconstrueerd op basis van een propositionele analyse van de beide verhalen (vgl. Heesters & Verhoeven, 1991). De eerste 16 items vormen een inventarisatie van de belangrijkste centrale betekenisinhouden. In de items 17 t/m 20 gaat het om de relaties tussen de betekenisinhouden. Aan de hand van deze items wordt beoordeeld of het kind er in slaagt om relaties uit te drukken als: middel-doel, oorzaak-gevolg, probleem-oplossing, veronderstelling-conclusie. De maximum score op deze naverteltaak was 20. Bij een random steekproef van 10 kinderen is de interbeoordelaarsbetrouwbaarheid met twee beoordelaars voor beide teksten vastgesteld. Deze bleek hoog (respectievelijk .94 en .97).

Daarnaast zijn voor elk van beide verhalen vragenformulieren ontwikkeld om na te kunnen gaan in hoeverre de kinderen in staat zijn het verhaal te begrijpen. De vragenformulieren bestaan beide uit 16 items met vragen naar aanleiding van de belangrijkste constituerende elementen uit het verhaal (setting, plot, ontknoping). In totaal konden op deze taak 32 punten behaald worden (2, 1 of 0 punten per item). Ook voor deze taken is bij een steekproef van 10 kinderen op vergelijkbare wijze als bij de naverteltaken de interbeoordelaarsbetrouwbaarheid bepaald. Deze kwam in beide gevallen uit op .93.

3.3 Procedure

Bij het klassikaal voorlezen is uitgegaan van de volgende condities:

1. Aan het begin van de ochtend wordt de onderzoeker door de groepsleerkracht aan de kinderen voorgesteld. De leerkracht legt de leerlingen uit wat deze 'juf' komt doen.
2. Na het afwerken van de 'ochtendrituelen', zoals het invullen van de presentielijst e.d., wordt het verhaal door de onderzoeker geïntroduceerd en voorgelezen. De kinderen zitten tijdens het voorlezen in de kring. Op vaste momenten worden standaardvragen gesteld om interactie over het verhaal tot stand te brengen. Na elke pagina loopt de onderzoeker de kring rond om de illustraties te tonen. De gehele voorleessessie wordt op cassetteband opgenomen.
3. De kinderen hervatten het normale pro-

gramma van die dag. Na een half uur begint het toetsen van de leerlingen. De kinderen worden hiertoe in willekeurige volgorde één voor één meegenomen naar een aparte ruimte.

4. De vier platen die bij het voorgelezen verhaal horen, worden voor het kind neergelegd met de volgende instructie:

'Ik heb jullie vanmorgen een verhaaltje voorgelezen, hè? Weet je dat nog? Weet je wat we nu gaan doen? Nu doen we net of ik het verhaaltje nog nooit gehoord heb. Jij mag het mij vertellen. Dit zijn de platen van het verhaal. Daar mag je naar kijken bij het vertellen. Dat is gemakkelijk als je het verhaal niet meer zo goed weet. Vertel maar alles wat je nog weet van het verhaal.'

Het navertellen wordt op cassetteband opgenomen. Als er niets (meer) gezegd wordt door het kind, wordt het navertellen gestimuleerd met 'Dat was goed! Weet je nog meer?'

5. Aansluitend aan het navertellen worden de 16 vragen van het vragenformulier gesteld. De instructie hierbij luidt als volgt:

'Dat heb je goed gedaan! Nu doen we de plaatjes weg. En dan heb ik nog een paar vragen over het verhaal. Die weet je vast ook well'

De gegeven antwoorden worden letterlijk op het vragenformulier genoteerd. Als een kind het antwoord niet weet, wordt dit op ongedwongen wijze door de onderzoeker gegeven:

- Wat wilde Boy graag hebben?
- (het kind zegt niets of geeft een verkeerd antwoord)
- Hij wilde graag een politie-auto, hè?

Het bevragen heeft meer het karakter van een ongedwongen conversatie dan van een ondervraging, om te vermijden dat de kinderen zich onzeker gaan voelen.

Bij het voorlezen in kleine groepjes is uitgegaan van de volgende condities:

1. Vijf kinderen worden meegenomen naar een aparte ruimte. De samenstelling van de groepjes is heterogeen; met zoveel mogelijk kinderen uit verschillende etnische groepen bij elkaar. Hier wordt het verhaal voorgelezen. Op vaste momenten worden de van tevoren opgestelde vragen gesteld. Het voor-

Tabel 2

Opzet van experimenteel onderzoek naar effecten van klassikaal en groepsgewijs voorlezen

Metingen in tijdsvolgorde	Betrokken school	Volgorde verhalen	Wijze van voorlezen
dag 1 - meting 1	school 1	verhaal A	klassikaal
dag 2 - meting 1	school 2	verhaal A	klassikaal
dag 3 - meting 1	school 3	verhaal B	klassikaal
dag 4 - meting 1	school 4	verhaal B	klassikaal
dag 5 - meting 2	school 1	verhaal B	in groepjes
dag 6 - meting 2	school 2	verhaal B	in groepjes
dag 7 - meting 2	school 3	verhaal A	in groepjes
dag 8 - meting 2	school 4	verhaal A	in groepjes

lezen wordt op cassetteband opgenomen. De kinderen zitten om de onderzoeker heen en kunnen tijdens het voorlezen de illustraties zien.

2. De kinderen gaan terug naar de klas en hervatten het normale programma van die dag.
3. Het verhaal wordt voorgelezen aan het volgende groepje van vijf.
4. De eerste vijf kinderen worden na elkaar getoetst en daarna volgen de kinderen van het tweede groepje. Dit gebeurt in dezelfde volgorde en op dezelfde manier als bij het klassikaal voorlezen.
5. Het verhaal wordt achtereenvolgens voorgelezen aan het derde en het vierde groepje.
6. De kinderen van groepje 3 en 4 worden na elkaar getoetst.

Alle voorleessessies werden gedaan door één en dezelfde persoon met ervaring met onderwijs aan kleuters. Om de interactie binnen de sessies achteraf te kunnen bepalen werden alle sessies met behulp van een cassetterecorder opgenomen. Tussen de twee voorleessessies per school was steeds een periode van twee weken aangehouden. De kinderen begonnen steeds met de klassikale voorleessessie. Twee weken nadien volgde de sessies in kleine groepjes. Deze vaste volgorde was gekozen, omdat de klassikale voorleessessie nauw aansluit bij de gangbare voorleespraktijk in het kleuteronderwijs en er dus minimale transfer naar de tweede sessie mag worden verwacht.

Om een effect van verschil in moeilijkheidsgraad van de twee voorleesverhalen zoveel mogelijk uit te bannen werden de twee verhalen in gelijke mate aan de twee onderzoekscondities toebedeeld. Op school 1 en school 2 werd begonnen met verhaal A in de eerste sessie en

vervolgd met verhaal B in de tweede sessie, terwijl de volgorde van verhalen bij school 3 en school 4 juist andersom was. Tabel 2 geeft het onderzoeksdesign schematisch weer.

Om de afnamecondities binnen de onderscheiden voorleessessies zoveel mogelijk te standaardiseren werden de volgende maatregelen genomen:

- Alle voorleessessies en metingen werden verricht door de onderzoeker, die zelf acht jaar in het basisonderwijs werkzaam is geweest.
- Alle metingen vonden plaats binnen het tijdsbestek van een maand. De tweede meting op een school volgde steeds veertien dagen na de eerste meting, op dezelfde dag van de week.
- De tijd tussen het voorlezen en het toetsen werd per kind zoveel mogelijk constant gehouden. Dat wil zeggen dat de kinderen de tweede keer in dezelfde volgorde bevraagd werden en dat kinderen die bij de eerste meting vrij snel na het voorlezen aan de beurt waren, dit de tweede keer zoveel mogelijk ook weer waren.

3.4 Verwerking van de gegevens

Ter beantwoording van de eerste onderzoeksvraag zijn de cassettebanden met de verhalen en de bijbehorende interacties aan een nadere analyse onderworpen. Per school zijn alle uitingen van kinderen die betrekking hadden op het verhaal opgeteld en ondergebracht in de categorie 'relevante uitingen'. De interbeoordelaarsbetrouwbaarheid (twee beoordelaars) bleek hoog (.98). Opmerkingen van kinderen die niets met het verhaal te maken hadden, vielen onder de 'niet-relevante uitingen'. Niet meegeteld werden:

- opmerkingen en vragen van de onderzoeker,
- opmerkingen als: *Juf! Ja! Nee! Ik weet het! Hier!* e.d. van de kinderen.

Vervolgens is met het oog op de tweede onderzoeksvraag nagegaan in hoeverre de scores op tekstbegrip en navertellen van de kinderen in de twee voorleescondities significant verschillen. Daaraan voorafgaand is om de interne consistentie van de items op de gehanteerde taken te controleren per taak Cronbachs Alpha berekend. Eveneens ter controle is bepaald in hoeverre de volgorde van de twee verhalen een differentieel effect had op de resultaten van de kinderen op de maten voor tekstbegrip en navertellen.

Met behulp van variantie-analyse (within-subjects design) is nagegaan in hoeverre het effect van het voorlezen in kleine groepjes significant was en of dit effect voor verschillende deelpopulaties van kinderen vergelijkbaar was. Daartoe werden de kinderen onderverdeeld in drie herkomstgroepen die nagenoeg gelijklijk over de twee voorleescondities zijn verdeeld:

1. autochtone Nederlandse kinderen (19);
2. kinderen uit ex-koloniale groepen (22);
3. Turkse en Marokkaanse kinderen (31).

4 Resultaten

4.1 Interactie tijdens het voorlezen

Voorlezen aan de hele klas lokte weinig spontane reacties uit. Kinderen begrepen heel goed dat het geen zin had om in een grote groep spontaan te reageren; als iedereen tegelijk iets uitroept, is er niemand meer te verstaan. Dit betekende dus dat er beurten gegeven moesten worden, waardoor veel van de spontaniteit verloren ging en waardoor ook lang niet alle kinderen aan bod konden komen. In de klassikale setting bleef er een duidelijke rolverdeling tussen de leerkracht (in dit geval de onderzoeker) die de vragen stelde en de kinderen die antwoord gaven. In deze setting kwam het niet één keer voor dat de kinderen met elkaar of met de leerkracht in discussie gingen.

In de 16 voorleessessies in kleine groepjes bleken kinderen over het algemeen actief te reageren tijdens het voorlezen. De verteller van het verhaal werd daarbij gezien als een gesprekspartner met wie de loop van het verhaal

kan worden bediscussieerd. Ter illustratie volgen hier twee fragmenten van discussies die tijdens het voorlezen in kleine groepen ontstonden:

Fragment 1: Discussie 'Hoe vang je een pannekoekemannetje'

- Hendrik: Maar dan moet je hem ook maar wat bloed geven met een mes!
- Onderzoeker: Waarom?
- Hendrik: Dan kan die niet meer opstaan.
- Onderzoeker: Oh, jij wilt hem gelijk doodsteken?
- Hendrik: Ja.
- Aron (hoofdschuddend): Nee, maar hij heb geen bloed!

Fragment 2: Discussie 'Hoe gemeen is een vos'

- Onderzoeker: Vinden jullie de vos lief of gemeen?
- In koor: Gemeen!
- Onderzoeker: Waarom dan?
- Hans: Want als hij, als hij pannekoekje, eh, panne ... jongetje opeet!
- Ritchy: In het echt zijn ze lief.
- Onderzoeker: De vossen?
- Ritchy: Ja!
- Kimberly: Echt niet!
- Ritchy: Echt wel!
- Onderzoeker: De vos zei: Ik lust helemaal geen pannekoekemannetjes!
- Hans: Nee, hij lust alleen maar kindertjes.
- Onderzoeker: Kindertjes?
- Kimberly: Echt niet. Hij lust alleen maar konijntjes en vogeltjes.

Om een beter zicht te krijgen op de wijze van interactie tijdens de twee onderscheiden voorleescondities zijn de interacties tijdens het voorlezen aan een nadere analyse onderworpen. De resultaten van deze kwantitatieve analyse van de interactie tijdens de beide voorleescondities zijn af te lezen in Tabel 3.

De verschillen tussen de klassikale setting en de setting in kleine groepjes zijn opmerkelijk. Gemiddeld zijn er in de klassikale situatie per verhaal 27 relevante uitingen gedaan. Gemiddeld per leerling komt dit neer op nagenoeg één uiting per kind. In de setting met groepjes van 5 bedroeg het aantal uitingen gemiddeld 49, hetgeen neerkomt op ongeveer 10 uitingen

Tabel 3

Aantal relevante en niet-relevante uitingen van kinderen tijdens het klassikaal voorlezen en het voorlezen in groepjes van 5

	klassikaal		groepje 1		groepje 2		groepje 3		groepje 4	
	relevant	niet relevant	relevant	niet relevant	relevant	niet relevant	relevant	niet relevant	relevant	niet relevant
School 1	32	0	46	3	45	2	28	3	36	2
School 2	24	0	36	0	45	9	26	1	74	9
School 3	30	0	59	2	75	0	36	0	72	4
School 4	23	0	29	2	52	0	67	6	60	4

per individueel kind. Deze resultaten kunnen worden opgevat als een eenduidige ondersteuning van de eerste onderzoekshypothese: Interactief voorlezen aan groepen van 5 leidt zowel bij allochtone als bij autochtone kleuters tot beduidend meer interactie dan interactief voorlezen aan de hele groep. Tabel 3 laat eveneens zien dat de interactiepatronen binnen de klassikale voorleesconditie op de vier scholen nauwelijks verschillen. Zowel binnen klassen met gemengde etnische groepen (school 1 en 3) als binnen klassen met overwegend allochtone kinderen (school 2 en 4) blijkt sprake van een overduidelijk effect van groeps-groote op de mate van interactie tijdens het voorlezen.

4.2 Retentie en begrip van het verhaal

Om de interne consistentie van de items op de taken voor het meten van de vaardigheid in tekstbegrip en navertellen naar aanleiding van de twee voorleesverhalen te meten is Cronbachs Alpha berekend. De waarde van Alpha kwam voor de tekstbegriptaak bij verhaal A uit op .81, voor de tekstbegriptaak bij verhaal B op .80. Voor de naverteltaken bij de beide verhalen kwam de numerieke waarde van Alpha uit op respectievelijk .76 en .79. In alle gevallen kan de interne consistentie van de geconstrueerde taken redelijk worden genoemd.

Verder is met behulp van variantie-analyse vastgesteld in hoeverre de volgorde van aan-

bieding van de twee voorleesverhalen een effect had op de scores van kinderen op beide typen taken. Dit bleek niet het geval. Eveneens is nagegaan of de volgorde van de twee verhalen een effect had op het niveau van zowel de Tekstbegriptaak als de Naverteltaak. Variantie-analyse wees uit dat dit niet het geval was. Dit houdt dus in dat er geen significante verschillen waren in moeilijkheidsgraad van de bij de twee verhalen behorende vragenformulieren en naverteltaken.

Tabel 4 geeft een overzicht van de gemiddelden en standaarddeviaties op de onderdelen Tekstbegrip en Navertellen van kinderen aan wie klassikaal, dan wel in kleine groepjes is voorgelezen.

In Tabel 4 valt af te lezen dat er beduidend hoger gescoord wordt na het voorlezen in groepjes van 5. Het hoofdeffect 'Wijze van voorlezen' is significant ($F(70,1)=42.74$, $p<.001$). Hypothese 2 is hiermee dus bevestigd: Als wordt voorgelezen in groepjes van 5 begrijpen kinderen het verhaal beter en onthouden ze er meer van dan wanneer klassikaal wordt voorgelezen. De interactie tussen 'Wijze van voorlezen' en 'Taak' was niet significant, hetgeen erop duidt dat het effect van de wijze van voorlezen voor beide taken even groot is.

Om een mogelijk differentieel effect te kunnen vaststellen in effecten voor allochtone en autochtone kleuters zijn enkele aanvullende

Tabel 4

Gemiddelde scores en standaarddeviaties op de taken Tekstbegrip en Navertellen na klassikaal voorlezen en voorlezen in groepjes van 5

	Tekstbegrip (klassikaal)	Tekstbegrip (groepjes van 5)	Navertellen (klassikaal)	Navertellen (groepjes v. 5)
Gemiddelde scores:	14.63	17.13	6.5	9.1
Standaarddeviaties:	6.82	6.46	3.7	3.5

Tabel 5

Gemiddelde scores en standaarddeviaties van de drie herkomstgroepen op de Tekstbegripstaak en de navertel-taak na klassikaal voorlezen en na voorlezen in groepjes van 5

	Tekstbegrip klassikaal	Tekstbegrip groepjes v. 5	Navertellen klassikaal	Navertellen groepjes v. 5
Nederlandse kinderen	18.90 (5.8)	22.05 (6.0)	8.42 (4.2)	10.74 (4.1)
Kinderen uit (ex-)koloniën	15.77 (6.0)	17.73 (4.9)	6.36 (3.4)	9.05 (2.9)
Mar./Turkse kinderen	11.19 (6.4)	13.68 (5.7)	5.48 (3.2)	8.07 (3.1)

analyses verricht. Tabel 5 geeft de gemiddelden en standaarddeviaties op de tekstbegrip-taak van de kinderen uit de drie onderscheiden deelpopulaties.

Multivariate variantie-analyse laat zien dat de onderscheiden hoofdeffecten 'Groep' ($F(69,2) = 12.81, p < .001$) en 'Wijze van voorlezen' ($F(69,1) = 39.89, p < .001$) beide significant zijn en dat de interactie tussen beide factoren dat niet is. We kunnen dus concluderen dat er sprake is van substantiële verschillen in scores tussen de drie groepen: Nederlandse kinderen scoren het hoogst, gevolgd door kinderen van (ex-)koloniën en Turkse/Marokkaanse kinderen. Verder blijkt sprake van een positief effect van het voorlezen in kleine groepjes dat voor de drie onderscheiden groepen vergelijkbaar is.

Bij het vragen beantwoorden en ook bij het navertellen bleek dat vooral de Turkse en Marokkaanse kleuters, die ten tijde van het onderzoek toch al minstens een jaar op de Nederlandse school zaten, nog erg weinig woorden tot hun beschikking hadden. Ze begonnen enthousiast te vertellen maar stuiten steeds op moeilijkheden bij het zoeken naar woorden. Ondanks deze problemen bleken alle kinderen wel de clou van beide verhalen begrepen te hebben, hoe weinig ze er soms ook van konden navertellen. Een belangrijk deel van de verklaring van de verschillen tussen de etnische groepen moet dus zeker gezocht worden in het feit dat veel allochtone kleuters het Nederlands nog slechts zeer gebrekkig beheersen.

Gegeven het feit dat in de beide uitgevoerde variantie-analyses de interactie tussen herkomst en aanpak niet significant is mogen we concluderen dat de verschillende herkomstgroepen in dezelfde mate profiteren van het interactief voorlezen in kleine groepen. Dit betekent dat allochtone kinderen evenzeer profiteren van het voorlezen in kleine groepjes als

autochtone kinderen. Hypothese 2 is hiermee dus bevestigd: Het positieve effect van het voorlezen in groepen van 5 ten opzichte van klassikaal voorlezen is voor allochtone en autochtone kleuters even groot.

5 Conclusies en aanbevelingen

Het uitgevoerde onderzoek laat zien dat kinderen bij het voorlezen van een verhaal in kleine groepen in vergelijking met klassikaal voorlezen meer met elkaar in interactie zijn en een verhaal beter begrijpen en onthouden. Waar in de klassikale situatie het aantal interacties per kind gemiddeld ongeveer 1 bedroeg, was dit in de kleine groepen per kind gemiddeld 10. Het lijkt aannemelijk dat het groter aantal interacties ertoe leidt dat kinderen in kleine groepen meer actief betrokken zijn bij het verwerken van het verhaal, hetgeen kan worden afgeleid uit de hogere scores voor tekstbegrip en navertellen.

Een belangrijke conclusie is verder dat het effect van het voorlezen in kleine groepjes bij allochtone kinderen even sterk is als bij autochtone kinderen. Tussen de onderscheiden deelpopulaties in het onderzoek bleek sprake van aanzienlijke verschillen in scores op zowel het onderdeel Tekstbegrip als het onderdeel Navertellen: autochtone kleuters scoorden hoger dan hun leeftijdgenoten uit ex-koloniën die op hun beurt weer hoger scoorden dan kleuters van Turkse en Marokkaanse herkomst. Alle groepen bleken echter in dezelfde mate te profiteren van het voorlezen in kleine groepen. Als zodanig kan het interactief voorlezen in kleine groepen ook als belangrijke werkvorm binnen het onderwijs Nederlands als tweede taal worden gezien. Omdat er in kleine groepjes meer gelegenheid is voor interactie en discussie zullen moeilijkheden en onduidelijkhe-

den met betrekking tot het verhaal in een kleine groep eerder opgelost worden dan in een klas-sikale setting. Het belang van het werken met allochtone kleuters in kleine groepen is eerder benadrukt door Damhuis (1988).

Een kwestie die nader onderzoek nood-zaakt, betreft de optimale groeps-grootte bij in-teractief voorlezen. Morrow en Smith (1990) suggereerden in hun onderzoek naar het effect van groeps-grootte dat het naar alle waarschijn-lijkheid niet veel uitmaakt of er aan groepjes van 3, van 5 of van 7 wordt voorgelezen. De in-druk die tijdens dit onderzoek is ontstaan, is echter dat een groepje van 5 of 6 het maximum is. Als er meer kinderen deelnemen, kunnen ze niet allemaal gelijktijdig in het boek meekij-ken. Tevens wordt de afstand tot de leerkracht een stuk groter voor de kinderen die aan de bui-tenkant zitten. Dit zou wel eens een negatief ef-fect kunnen hebben op de betrokkenheid en de concentratie en als gevolg daarvan ook op de interactie. Nader onderzoek hiernaar zou wen-selijk zijn.

De indruk bestaat dat er op de basisschool nog te weinig bekend is over de vele positieve effecten van interactief voorlezen. Dat voorle-zen een positief effect heeft op de taalontwik-keling in het algemeen is bekend, maar dat de interactie tijdens het voorlezen dit effect ver-groot en dat er meer interactie plaatsvindt in kleine groepjes, is veel minder bekend. Van be-lang daarbij is de constatering binnen het on-derhavige onderzoek dat een ontspannen sfeer en een goede concentratie bij kinderen belang-rijke randvoorwaarden vormen voor het suc-cesvol verloop van de interactie tijdens het voorlezen in kleine groepjes. Een en ander stelt hoge eisen aan de professionalisering van de leerkracht op dit terrein.

De positieve waardering van het voorlezen in kleine groepjes is geen pleidooi om het klas-sikaal voorlezen af te schaffen. De verschil-lende vormen van voorlezen hebben elk hun eigen intrinsieke waarde. Wel wordt gepleit voor een meer gedifferentieerde aanpak van het voorlezen, met verschillende soorten prenten-boeken, waarbij afwisselend in kleine groepjes en klassikaal wordt voorgelezen en waarbij be-paalde boeken meerdere malen worden her-haald. Zo komen kinderen aan wie thuis weinig (in het Nederlands) is voorgelezen beter aan hun trekken. Een betere (na)scholing van leer-

krachten op het terrein van de verschillende vormen van voorlezen is gewenst. Het ultieme doel van scholing zou moeten zijn dat leer-krachten leren te reflecteren op hun eigen voorleesgedrag, zodat zij geleidelijk in staat worden hun onderwijsstrategieën te optimali-seren (vgl. Gaffney & Anderson, 1991).

Literatuur

- Both-de Vries, A.C., & Bus, A.G. (1992). De verhalen-man: Een vertelproject met video. *Jeugd in School en Wereld*, 77, 12-15.
- Bruner, J. (1977). Early social interaction and lan-guage development. In H.R. Schaffer (Ed.), *Stu-dies in mother-child interaction* (pp. 21-37). Lon-don: Academic Press.
- Bus, A.G., & IJzendoorn, M.H. van (1988). Mother-child interactions, attachment and emergent literacy: A cross-sectional study. *Child Develop-ment*, 59, 1262-1272.
- Bus, A.G., IJzendoorn, M.H. van, Pellegrini, A.D., & Terpstra, W. (1994). Van voorlezen naar lezen. *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs*, 10, 157-174.
- Bus, A.G. (1990). Van voorlezen naar lezen. *Moer*, 7, 320-327.
- Cazden, C.B. (1982). Contexts for literacy: In the mind and in the classroom. *Journal of Reading Behavior*, 14, 413-427.
- Damhuis, R. (1988). *Tweede-taalverwerving in kleu-tergroepen. Een onderzoek naar de gelegenheid tot het leren van het Nederlands door Turkse en Marokkaanse kleuters*. SCO-rapport nr. 164, Am-sterdam: S.C.O.
- Dickinson, D.K., & Smith, M.W. (1994). Long-term ef-fects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly*, 29 (2), 104-123.
- Gaffney, J.S., & Anderson, R.C. (1991). *Two-tiered scaffolding: Congruent processes of teaching and learning*. Technical Report, 523. Champaign: Center for the Study of Reading.
- Heath, S.B. (1982). What no bedtime story means: Narrative skills at home and at school. *Language in Society*, 11, 49-76.
- Heesters, K., & Verhoeven, L. (1991). Ontwikkeling van tekstvaardigheid in het Nederlands als eer-ste en tweede taal. *Tijdschrift voor Taalbeheer-sing*, 13 (3), 169-184.

- Kurvers, J., & Lemmens, G. (1992). *Leergeld van een generatie. Tweede generatie Marokkaanse en Turkse moeders over taalstimulering van kleuters*. Rijswijk: Ministerie van WVC.
- Leseman, P., Sijlsing, F., Jap-A-Joe, S., & Şahin, S. (1995). Gezinsdeterminanten van de cognitieve ontwikkeling van vierjarige Nederlandse, Surinaamse en Turkse kleuters. *Pedagogische Studiën*, 72, 186-205.
- Morrow, L.M. (1984). Reading stories to young children: Effects of story structure & traditional questioning strategies on comprehension. *Journal of Reading Behavior*, 16, 273-288.
- Morrow, L.M. (1985). Retelling stories: Strategies for improving children's comprehension, concept of story structure, and oral language complexity. *Elementary School Journal*, 85, 647-661.
- Morrow, L.M. (1988). Young children's responses to one-to-one readings in school settings. *Reading Research Quarterly*, 23, 89-107.
- Morrow, L.M., & Smith, J.K. (1990). The effects of group size on interactive storybook reading. *Reading Research Quarterly*, 25, 213-230.
- Ninio, A. (1980). Picture-book reading in mother-infant dyads belonging to two subgroups in Israel. *Child Development*, 51, 587-590.
- Notari, A., O'Connor, R., & Vadasy, P. (1993). *Ladders to literacy: an activity book (draft version)*. Seattle: Washington Research Institute.
- Pels, T. (1991). *Marokkaanse kleuters en hun culturele kapitaal. Opvoeden en leren in het gezin en op school*. Amsterdam/Lisse: Swets & Zeitlinger.
- Slavin, R.E. (1995). *Cooperative learning*. Needham Heights, MA: Allyn & Bacon.
- Snow, C.E. (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review*, 53, 165-189.
- Snow, C.E., Barnes, W., Chandler, J., Goodman, I., & Hemphill, L. (1991). *Unfulfilled expectations: Home and school influences on literacy*. Boston: Harvard University Press.
- Sulzby, E. & Teale, W.H. (1991). Emergent literacy. In R. Barr & D. Pearson (Eds.), *Handbook of reading research*. (pp. 727-758). New York: Longman.
- Teale, W.H. (1980). *Early reading: An annotated bibliography*. Newark, DE: International Reading Association.
- Teale, W.H. (1981). Parent's reading to their children: What we know and need to know. *Language Arts*, 58, 902-911.
- Teale, W. (1984). Reading to young children: Its significance in the process of literacy development. In H. Goelman, A. Oberg & F. Smith (Eds.), *Awakening to literacy* (pp. 110-121). Exeter, NH: Heinemann.
- Teale, W.H. & Sulzby, E. (1987). Literacy Acquisition in early childhood: The roles of access and mediation in storybook readings. In D.A. Wagner (Ed.), *The future of literacy in a changing world* (pp. 111-130). New York: Pergamon Press.
- Verhoeven, L. (1994). *Ontluikende geletterdheid*. Lisse: Swets & Zeitlinger.
- Wells, G. (1985). Preschool literacy-related activities and success in school. In D.R. Olson, N. Torrance & A. Hildyard (Eds.), *Literacy, language and learning* (pp. 229-255). Cambridge, England: Cambridge University Press.

Manuscript aanvaard 28-12-1996

Auteurs

W. van Elsäcker is momenteel uitvoerder van een onderzoeksproject naar de ontwikkeling van leesvaardigheid in relatie met leesmotivatie dat wordt uitgevoerd binnen de vakgroep Orthopedagogiek van de Katholieke Universiteit Nijmegen.

L. Verhoeven is hoogleraar Orthopedagogiek aan de Katholieke Universiteit Nijmegen.

Adres: Instituut voor Orthopedagogiek, KUN, Montessorilaan 3, 6525 HR Nijmegen.

Abstract

Children in Kindergarten learn more from interactive storybook reading in small groups

W. van Elsäcker & L. Verhoeven. *Pedagogische Studiën*, 1997, 74, 117-129.

Storybook reading in Kindergarten turns out to play a significant role in children's early literacy development. The responsivity of teachers plays an important role in helping children to gain insight into the functions and structure of both oral and written language. Scaffolding can be seen as a crucial concept associated with teacher assistance in children's language and literacy learning. In the present study the effect of group setting on interactive storybook reading was explored in multilingual kindergarten classes. The children in these classes heard stories read in two settings: traditional class vs small groups. The results made clear that children in small groups learn more from storybook reading than the children in the traditional class setting. The children in the small groups showed much more interaction than the children in the whole-class setting. Moreover, on free recall and oral comprehension tests children who heard stories in the small-group setting performed better than children who heard stories in the whole-class setting. The effect was equally strong for native Dutch and immigrant children.

Courseware in onderzoekend natuurkunde-onderwijs: een implementatie perspectief

J.M. Voogt

Samenvatting

De integratie van de computer in het onderwijs is een complexe innovatie. In dit onderzoek is courseware voor computer ondersteund natuurkunde-practicum ontwikkeld en geëvalueerd. Kenmerkend voor de ontwikkelde courseware is dat de computer wordt ingezet als hulpmiddel om onderzoekend natuurkunde-onderwijs te realiseren. Ter ondersteuning van de implementatie van de courseware werd de docent-handleiding voorzien van handelingsaanwijzingen voor docenten. Via een case-study onderzoek (5 docenten/6 klassen) is nagegaan of de implementatie van de courseware succesvol is verlopen.

Uit de resultaten van het onderzoek blijkt dat gestructureerd leerlingmateriaal de implementatie van de vernieuwing op een globaal niveau bevordert. Echter, alleen die docenten die ook gebruik maakten van de docent-handleiding (twee van de vijf docenten) hanteren een didactisch aanpak die noodzakelijk wordt geacht voor onderzoekend natuurkunde-onderwijs. Voor anderen was de overgang van gangbare routines naar onderzoekend leren te groot.

Inleiding

De inzet van de computer als hulpmiddel bij de uitvoering van metingen en het be- en verwerken van meetgegevens wordt wel Computer Ondersteund Practicum genoemd. Het is een zeer geëigende toepassing van de computer in het natuurkunde-onderwijs, zeker als practicum een belangrijk onderdeel vormt van het natuurkunde curriculum.

Uit onderzoek naar de effecten van het gebruik van Computer Ondersteund Practicum is gebleken dat leerlingen beter grafieken leren interpreteren, dat de samenwerking tussen leerlingen wordt bevordert en dat leerlingen geconcentreerd werken en zeer gemotiveerd zijn

(Thornton & Sokoloff, 1990; Nakhleh, 1994). De studies tonen evenwel ook aan dat het niet vanzelfsprekend is dat leerlingen door middel van Computer Ondersteund Practicum onderzoeksvaardigheden (Striley, 1987) en natuurwetenschappelijke begrippen leren (Linn & Songer, 1991). Uit een literatuuroverzicht van Nakhleh (1994) blijkt dat Computer Ondersteund Practicum pas effect heeft op het leren van onderzoeksvaardigheden en het leren van natuurwetenschappelijke begrippen als het is ingebed in een zorgvuldig ontworpen curriculum.

In het natuurwetenschappelijk onderwijs wordt onderzoekend leren vaak in verband gebracht met 'de natuurwetenschappelijke methode'. Leerlingen dienen niet alleen te worden onderwezen in het natuurwetenschappelijk kennisdomein, maar ook in de onderzoeksbenadering die in de natuurwetenschappen wordt gehanteerd. Sinds de jaren zestig hebben de curricula voor het onderwijs in de natuurwetenschappen uit met name de Engel-Saksische landen 'de natuurwetenschappelijke methode' als belangrijk organisatieprincipe (Voogt, 1996). Het project Science - A process approach (Gagné, 1966) is van grote invloed geweest op curricula die expliciet aandacht schenken aan het leren van onderzoeksvaardigheden. In dit project wordt een onderscheid gemaakt tussen basisvaardigheden (zoals observeren, meten en voorstellen) en geïntegreerde vaardigheden (zoals hypothese formuleren en modellen ontwikkelen en toetsen).

In de loop der jaren is uiteraard ook kritiek geformuleerd op de eenzijdige belangstelling voor de natuurwetenschappelijke methode als kern van het natuurwetenschappelijk curriculum. De bezwaren richten zich met name op de suggestie dat natuurwetenschappelijk onderwijs ingericht op deze wijze in principe contextonafhankelijk zou zijn (Driver et al., 1994). Driver et al. pleiten voor een constructivistische

benadering van het natuurwetenschappelijk onderwijs, waarbij leerlingen actief participeren in de constructie van een raamwerk dat betekenis geeft aan de natuurwetenschappelijke wijze waarop de werkelijkheid wordt geïnterpreteerd. Het opzetten en uitvoeren van eigen experimenten is dan expliciet bedoeld om bij te dragen aan de ontwikkeling van een, voor leerlingen, betekenisvol theoretisch raamwerk en het op basis daarvan afleiden van natuurkundige regels (Millar, 1987; Linn, 1987). De ontwikkelingen in Nederland staan meer in dit laatste perspectief. De toenemende aandacht voor het belang van onderzoekend leren in het natuurwetenschappelijk onderwijs wordt uitdrukkelijk ingebed in voor de leerlingen herkenbare samenhangende contexten, zoals blijkt uit de kerndoelen Natuur- en scheikunde voor de basisvorming (Besluit kerndoelen en advies-urentabel basisvorming 1993-1998).

Samenvattend kan worden gesteld dat onderzoekend leren in het natuurwetenschappelijk onderwijs zich kenmerkt door uitdrukkelijk te streven naar een samenhang tussen onderzoeksvaardigheden (tot uiting komend in aandacht voor onderzoeksopzet en -uitvoering) en de ontwikkeling van een betekenisvol theoretisch raamwerk (tot uiting komend in de opbouw van een begrippenkader en de afleiding van natuurkundige regels).

In een dergelijke curriculumopvatting heeft de docent de taak leerlingen te ondersteunen in hun eigen leerproces door geschikte aanwijzingen te geven gericht op de zone van naaste ontwikkeling (Linn & Burbulus, 1993). Onderzoeken (Shafto & Capper, 1987; Roberts, 1988) op het gebied van Computer Ondersteund Practicum laten zien dat juist deze rol van de docent een belangrijk obstakel vormt voor de implementatie van het curriculum. De docent wordt in zijn nieuwe rol niet zozeer aangesproken op zijn kennis, maar op zijn organisatorische, begeleidende en stimulerende capaciteiten. Teneinde docenten bij te staan in de uitvoering van hun nieuwe rol stelt Van den Akker (1988) voor om docent-ondersteunend curriculummateriaal te ontwikkelen met daarin concrete handelingsaanwijzingen ('procedurele specificaties'), vooral voor die onderdelen van de innovatie die kwetsbaar zijn bij de implementatie.

Het curriculum dat in het kader van het onder-

havige onderzoek is ontwikkeld beoogt leerlingen te leren experimenten op te zetten en uit te voeren en de resultaten van die experimenten in verband te brengen met een theoretisch raamwerk op basis waarvan elementaire natuurkundige regels worden afgeleid. Het curriculum is bedoeld voor het natuurkunde-onderwijs in de onderbouw. De courseware die in het kader van de lessenserie is ontwikkeld bevat zowel educatieve programmatuur als schriftelijk leerling- en docentmateriaal.

Omdat verwacht werd dat met name docenten weinig bekend zijn met onderzoekend natuurkunde-onderwijs is de docentenhandleiding voorzien van procedurele specificaties: aanwijzingen en suggesties gericht op de rol van begeleider, en technische en organisatorische aanwijzingen gericht op planning en uitvoering van de lessen. Het evaluatie-onderzoek richt zich met name op de wijze waarop de courseware door docenten in de praktijk is geïmplementeerd.

Om te beoordelen of de implementatie van courseware succesvol verloopt, is het van belang na te gaan wat de kwaliteit van het gebruik ervan is. In het onderzoek is gebruik gemaakt van de drie dimensies die Van der Grift (1987) onderscheidt voor de kwaliteit van een innovatie: theoretische, empirische en praktische kwaliteit. De *theoretische kwaliteit* heeft betrekking op de theoretische en vakdidactische principes waarop de innovatie is gebaseerd. De theoretische kwaliteit van de courseware is geoptimaliseerd door middel van expert-oordelen. In een pre-test posttest design is de *empirische kwaliteit* vastgesteld. Na afronding van de lessenserie met de courseware blijken de leerlingen een aanzienlijke leerwinst te hebben geboekt. Een uitgebreide beschrijving en rapportage omtrent de theoretische en empirische kwaliteit is te vinden in Voogt (1993). De *praktische kwaliteit* verwijst naar de haalbaarheid en de uitvoerbaarheid van de innovatie in de praktijk. Door middel van een case-study onderzoek is nagegaan in welke mate de vernieuwing geïmplementeerd is door docenten. In dit artikel wordt met name ingegaan op die praktische kwaliteit van de vernieuwing, en staat dus vooral het docent-perspectief centraal.

Ter beoordeling van de praktische kwaliteit van de courseware zijn de volgende onderzoeksvra-

Introductie

1. Meten met de computer (les 1)
2. Temperatuur regelen (les 2)

Introductie van IP-Coach en van basisvaardigheden in het uitvoeren van experimenten, zoals lezen van grafieken

Inwendige energie en warmtetransport

3. Opwarmen van verschillende hoeveelheden (les 3)
4. Opwarmen van verschillende vloeistoffen (les 4)

Onderzoek gericht op de relatie tussen hoeveelheid en soort vloeistof, temperatuursverandering en verandering in inwendige energie; onderzoeksvraag formuleren (4), voorspellen van temperatuurverloop (3) en interpretatie van grafieken (3+4); meetfouten (4).

Intermezzo: theorieles (les 5)

5. Afkoelen: blazen of ijs (les 6)

Onderzoek naar de effecten van verschillende manieren van afkoelen; extrapoleren; samenwerking tussen onderzoeksgroepjes.

6. Smelten en stollen (les 7)

Onderzoek naar de relatie tussen temperatuursverandering en fase-overgang; zelfstandig noteren van meetresultaten

Intermezzo: theorieles (les 8)

7. Belemmering van warmtetransport (les 9)

Vaststellen van de mate van warmte-isolatie van verschillende materialen in een klassikaal experiment; proefopstelling vaststellen

Samenvatting en overzicht

8. Kijken als een natuurkundige (les 10)

Een terugblik op de natuurkundige regels die in de lessen zijn afgeleid

Figuur 1. Overzicht van de lessenserie

gen geformuleerd:

- welke kenmerken van de courseware bevorderen of belemmeren het gewenste plannings- en begeleidingsgedrag van docenten?
- verandert het plannings- en begeleidingsgedrag van docenten tijdens het gebruik van de courseware?
- waarom tonen docenten het gewenste plannings- en begeleidingsgedrag al dan niet?
- welke neveneffecten belemmeren de implementatie van de courseware?

1 De ontwikkelde lessenserie

Er is courseware ontwikkeld voor het onderwerp 'opwarmen en afkoelen' bedoeld voor een lessenserie van 10 lessen. Figuur 1 geeft een overzicht van de lessenserie.

De courseware bestaat uit educatieve programmatuur¹ met daarbij begeleidend materiaal voor leerlingen en docenten.

Het onderwerp 'opwarmen en afkoelen' wordt aangeboden aan de hand van door de leerlingen uit te voeren experimenten. Er wordt gewerkt in vaste groepjes van vier leerlingen. Binnen de groepjes zijn de volgende taken te vervullen: experiment uitvoeren, computer bedienen, aantekeningen maken en coördineren. Deze taken wisselen per les. In het leerlingmateriaal staat aangegeven wanneer welke taken moeten worden uitgevoerd. Ook staat aangegeven wanneer de groep met elkaar moet overleggen. Kenmerkend voor de lessenserie is dat de leerlingen tijdens hun experiment het temperatuurverloop in de tijd kunnen volgen via het beeldscherm van hun computer. Met behulp van de resultaten van hun experimenten leiden de leerlingen een theoretisch begrippenkader en natuurkundige regels af. De leerlingen worden daarbij gestuurd door gerichte vragen in het leerlingmateriaal.

Het docentmateriaal bestaat uit een algemene inleiding over de opzet van de lessenserie en de voorbereiding van de lessen. Daarnaast krijgt

de docent, geïntegreerd in het leerlingmateriaal (op de linkerpagina), concrete organisatorische en didactische aanwijzingen.

2 Onderzoeksopzet

Omdat in dit onderzoek de beoordeling van de praktische kwaliteit van courseware centraal staat, is gekozen voor een onderzoeksopzet in de daadwerkelijke gebruikssituatie: de klas. Er is een exploratieve case-study uitgevoerd (Yin, 1989). Als cases worden de natuurkunde curricula in de onderbouw van de twee bij het onderzoek betrokken scholen beschouwd. Het onderzoek naar de praktische kwaliteit van de courseware impliceert dat de *docent in zijn klas* beschouwd wordt als de onderzoekseenheid in deze studie. In navolging van Yin (1989) zijn criteria geformuleerd om de onderzoeksresultaten van de case-study te kunnen beoordelen.

Criteria voor een positieve beoordeling van de praktische kwaliteit van de courseware zijn:

- de docenten zijn positief over de 'instrumentaliteit' van het materiaal, d.w.z. a) de programmatuur is makkelijk te bedienen, b) het leerlingmateriaal is duidelijk, c) de aanwijzingen en suggesties in de docentenhandleiding zijn bruikbaar en d) de docenten maken daadwerkelijk gebruik van de docentenhandleiding bij hun lesvoorbereiding en -uitvoering;
- de docenten voeren de voorgestelde lessen daadwerkelijk uit;
- de docenten tonen een plannings- en begeleidingsgedrag dat onderzoekend leren bevordert, d.w.z. ze bevorderen bij leerlingen vaardigheden die bijdragen aan onderzoeksopzet en -uitvoering, stimuleren het leggen van een relatie met een theoretisch kader en bevorderen het afleiden van natuurkundige regels.

2.1 Innovatie- en praktijkprofiel

Om een beoordeling te kunnen geven van de praktische kwaliteit van het materiaal is het niet voldoende om docenten daarover te bevragen. Immers de wijze waarop docenten de bedoelingen van de courseware interpreteren, kan verschillen met de wijze waarop docenten de courseware in de concrete lespraktijk realiseren. Omdat met name de realisatie in de praktijk in het onderzoek belangrijk werd gevonden heeft

het accent in de gegevensverzameling op de observaties van lessen gelegen.

De geobserveerde lessen zijn geanalyseerd met behulp van een zogenaamd innovatieprofiel (Van den Akker, 1988; Van den Akker & Voogt, 1994). Een innovatieprofiel geeft (1) een operationele beschrijving van die componenten in het curriculum die door de ontwerpers als essentieel worden beschouwd en (2) een beschrijving van de mate waarin verschillende realisaties van die componenten overeenkomen met de bedoelingen van de ontwerpers.

In navolging van de kenmerken van onderzoekend natuurkunde-onderwijs zijn de volgende componenten in het profiel onderscheiden: *onderzoeksopzet* (bijvoorbeeld leerlingen stimuleren een onderzoeksvraag te formuleren), *onderzoeksuitvoering* (bijvoorbeeld leerlingen stimuleren een voorspelling te doen), *ontwikkelen van een begrippenkader* (bijvoorbeeld leerlingen stimuleren om resultaten in verband te brengen met theoretische begrippen) en *afleiden van natuurkundige regels* (bijvoorbeeld leerlingen stimuleren verbanden te leggen tussen nieuwe regels en eerder afgeleide regels).

Daaraan is toegevoegd de component *algemene vaardigheden*. Deze component heeft betrekking op samenwerking en planning van leerlingen en is van belang voor practicum-onderwijs dat in groepjes leerlingen plaatsvindt.

Elke component bestaat uit een leselement (de drempel) dat beschouwd wordt als een minimale vereiste voor het zinvol doceren van de lessenserie. Daarnaast zijn leselementen geformuleerd, die de implementatie van de lessenserie versterken of verzwakken. Aan zowel het leselement dat als drempel wordt beschouwd, als aan de leselementen die de implementatie van de lessenserie versterken of verzwakken worden punten toegekend.

Door aan docenten punten toe te kennen voor de aanwezigheid van de verschillende leselementen uit het innovatieprofiel in hun lespraktijk ontstaat een maat voor de implementatie van de lessenserie: het praktijkprofiel. In het praktijkprofiel wordt door middel van een percentage het plannings- en begeleidingsgedrag van de docent uitgedrukt t.o.v. het ideaal gewenste plannings- en begeleidingsgedrag.

Via de drempelscore zijn maximaal slechts 30 (16,7%) van de in totaal 180 te behalen punten te verwerven. Door deze lage drempelscore

is het mogelijk verschillen tussen docenten in plannings- en begeleidingsgedrag vast te stellen.

Bovenop de drempelscore kunnen punten worden verkregen voor leselementen die bijdragen aan het verwachte plannings- en begeleidingsgedrag. Om recht te doen aan de samenhang tussen onderzoeksvaardigheden en de opbouw van een theoretisch raamwerk zijn in het profiel ongeveer evenveel punten toegekend aan de componenten onderzoeksopzet (22 punten, 12,2%) en -uitvoering (54 punten, 30,0%) als aan de componenten theoretisch raamwerk (35 punten, 19,4%) en natuurkundige regels (36 punten, 20,0%). Met de component algemene vaardigheden kunnen maximaal 33 punten (18,3%) worden verworven.

Voor leselementen die het plannings- en begeleidingsgedrag verzwakken scoren docenten negatief. De maximale negatieve score is gelijk aan de drempelscore (-30 punten, -16,7%).

Het innovatieprofiel is ontwikkeld door de onderzoeker en gevalideerd door vijf deskundigen, waaronder curriculumspecialisten en vakdidactici natuurkunde.

2.2 Betrokken scholen en docenten

Bij het onderzoek zijn twee scholen voor voortgezet onderwijs betrokken. Van elke school namen alle derde klassen HAVO, VWO of HAVO/VWO deel. Dit zijn zes klassen, waaraan door vijf verschillende docenten les wordt gegeven. Tabel 1 geeft een overzicht van de betrokken klassen en docenten.

School I kan als volgt worden getypeerd:

School I is een school voor MAVO, HAVO en VWO. De drie derde klassen die in het onderzoek participeren zijn HAVO/VWO klassen. Het natuurkunde-curriculum in de onderbouw van deze school bestaat uit een diepgaande

behandeling van een beperkt aantal onderwerpen. Meestal wordt een onderwerp geïntroduceerd in een practicumles, gevolgd door een theorieles, waarin enkele toepassingen aan de orde komen. Kuiper (1993) spreekt in dit verband van participatiegericht natuurkunde-onderwijs. In school I zijn de volgende docenten bij het onderzoek betrokken:

Docent A heeft 9 jaar ervaring en is bevoegd voor wiskunde en natuurkunde. Gewoonlijk geeft docent A voornamelijk wiskunde. Door een tekort aan natuurkunde docenten geeft docent A dit jaar meer natuurkundelessen dan wiskundelessen. Docent A heeft enige ervaring in het gebruik van computers in zijn onderwijs, met name bij wiskunde. Hij is niet overtuigd dat de computer het onderwijs efficiënter maakt.

Docent B heeft 15 jaar ervaring en is bevoegd voor biologie en natuurkunde. Biologie is zijn hoofdvak, slechts af en toe geeft hij enkele lessen natuurkunde. Het schooljaar, voorafgaand aan het onderzoek, heeft hij zelfs helemaal geen natuurkunde gegeven. Dit jaar geeft docent B ongeveer evenveel lessen biologie als natuurkunde. Docent B gebruikt in zijn biologielessen regelmatig een computersimulatieprogramma.

Docent C heeft 13 jaar ervaring en is bevoegd voor natuurkunde en wiskunde. De laatste jaren geeft hij alleen nog maar natuurkunde. In het schooljaar waarin het onderzoek plaatsvindt geeft hij daarnaast een uur per week informatica. Docent C heeft zelf programma's ontwikkeld voor het natuurkunde-onderwijs. Docent C gebruikt regelmatig computers tijdens zijn lessen, maar heeft niet eerder gewerkt met Computer Ondersteund Practicum. Docent C was ook betrokken bij de formatieve evaluatie van de courseware.

School II wordt als volgt gekarakteriseerd:

School II is een school voor HAVO, atheneum en gymnasium. De drie derde klassen die

Tabel 1
Betrokkenen bij het onderzoek – een overzicht

	SCHOOL I			SCHOOL II		
	A	B	C	D	E(1)*	E(2)
Docent						
Schooltype	HAVO/VWO	HAVO/VWO	HAVO/VWO	HAVO	Gymnasium	Atheneum
Klasgrootte	30	30	30	21	22	26

* Docent E participeert met 2 klassen in het onderzoek

in het onderzoek participeren zijn een HAVO-, een atheneum- en een gymnasiumklas. Het natuurkunde-curriculum in de onderbouw van deze school heeft een concentrische opbouw. Een veelheid aan onderwerpen wordt oppervlakkig behandeld in de tweede klas en iets diepgaander in de derde klas. De natuurkundelessen zijn voornamelijk theoretisch van aard. De docent demonstreert proeven met behulp van de Technisch Onderwijs Assistent. Incidenteel krijgen de leerlingen practicum. Het practicum is vooral bedoeld als een illustratie van de theorie. Kuiper (1993) spreekt in dit verband van vakstructuurgericht natuurkunde-onderwijs. In school II participeren de volgende docenten in het onderzoek:

Docent D heeft 15 jaar ervaring en is bevoegd voor natuurkunde. Docent D heeft het voorafgaande schooljaar in de bovenbouw lesgegeven over databases. Soms gebruikt hij de computer voor demonstratieproeven. Hij heeft zelf enkele computerprogramma's geschreven.

Docent E heeft 5 jaar ervaring. Hij geeft 4 jaar natuurkunde. Dit jaar voor het eerst op school II. Daarnaast geeft hij les aan een andere school in de regio. Docent E heeft nooit eerder computers gebruikt in zijn onderwijs. Ongeveer een jaar heeft hij thuis een computer. Hij zegt niet met computers te kunnen omgaan.

2.3 Faciliteiten

Beide scholen zijn vanuit het onderzoeksproject voorzien van tien computers met interface, meetpaneel en temperatuursensor, twee printers, en de benodigde programmatuur. Als bijdrage aan het project investeerde School I in verrijdbare karren, zodat de computers makkelijk in en uit het natuurkundelokaal konden worden gereden. Daarnaast maakte School I ook zelf temperatuursensoren.

2.4 Dataverzameling en -analyse

Lesobservaties. Door onderzoeksassistenten zijn alle lessen ($N=80$) die in het kader van de lessenserie werden gegeven geobserveerd en opgenomen op geluidsband. De lesobservaties richtten zich op het verzamelen van gegevens over het feitelijk lesverloop, de tijd die aan elk afzonderlijk hoofdstuk is besteed, eventuele technische problemen met de apparatuur en programmatuur en het gebruik door de docenten van de aanwijzingen en suggesties uit de docen-

tenhandleiding. Door de lessen op geluidsband op te nemen is het mogelijk om de interventies van de docent met groepjes leerlingen te volgen, tijdens de uitvoering van de onderzoekjes. De uitgeschreven geluidsbanden zijn gebruikt bij de analyse van de gegevens. De lesobservaties zijn geanalyseerd met behulp van het innovatieprofiel. De toekenning van scores is uitgevoerd door twee getrainde onafhankelijke beoordelaars. Met behulp van Cohens Kappa (Cohen & Manion, 1980) is per docent de interbeoordelaarsbetrouwbaarheid bepaald. Deze varieerde tussen 0,85 en 0,98, en is daarmee zeer bevredigend. Nadat de interbeoordelaarsbetrouwbaarheid was vastgesteld, overlegden beide beoordelaars daar waar zij verschilden in de toekenning van scores om tot de definitieve score per docent te komen. Als de beoordelaars niet tot overeenstemming kwamen besliste de onderzoeker over het al dan niet toekennen van de score, rekening houdend met de argumentatie van beide beoordelaars. Dit was noodzakelijk voor 10 van de 63 verschillen tussen de beoordelaars.

Half-gestructureerd interview. Na afloop van de lessenserie is een interview afgenomen met elke docent. In het interview zijn de volgende onderwerpen besproken: welke doelen worden volgens de docent met de lessenserie nagestreefd, in welke mate komt de lessenserie overeen met de wijze waarop de natuurkundelessen gewoonlijk worden gegeven, hoe wordt de les voorbereid, hoe ervaart de docent de lesuitvoering en welke opbrengst heeft de lessenserie volgens de docent. Daarnaast is in het interview achtergrondinformatie verzameld, zoals aantal jaren ervaring in het onderwijs, ervaring met computers en bevoegdheid.

Met behulp van de dimensies voor de praktische kwaliteit van een innovatie (Doyle & Ponder, 1977-78): 'instrumentality', 'congruence' en 'cost', is het interview met de docenten geanalyseerd. De operationalisering van deze dimensies is ontwikkeld in eerder onderzoek (Voogt, 1990). 'Instrumentality' heeft betrekking op het oordeel van docenten omtrent de duidelijkheid van de programmatuur en het leerling- en docentmateriaal met het oog op lesvoorbereiding en gebruik in de klas. 'Congruence' hangt samen met de mate waarin de lessenserie past in de lespraktijk van de docent en

in het natuurkunde-curriculum in de onderbouw. 'Cost' is de afweging tussen de inspanning die geleverd moet worden om de lessenserie uit te voeren en de opbrengsten, in de vorm van betere realisatie van beoogde doelstellingen en gemotiveerde leerlingen.

3 Resultaten

3.1 Gebruik van courseware in de les: 'opportunity to learn'

Om vast te stellen of de docenten de leerlingen daadwerkelijk in de gelegenheid stellen de lessenserie te volgen, is nagegaan hoeveel tijd de docenten aan de courseware hebben besteed en of ze de courseware in de aanbevolen volgorde hebben onderwezen. Het blijkt dat de tijd die docenten aan de lessenserie besteed hebben, aanzienlijk meer is dan door de ontwikkelaars van de courseware is gepland. Deze tijdsoverschrijding komt vooral door de introductie- en theorielessen. De practicumlessen hebben ongeveer de geplande tijd gevergd. Uit de observaties wordt duidelijk dat de lessen inderdaad zijn uitgevoerd in de gewenste volgorde. Bij sommige docenten, met name docenten B en C is de afwisseling tussen practicum- en theorielessen, met name in de tweede helft van de lessenserie, minder dan gewenst. Voor docent B geldt dat met name in het begin enkele lessen uitvielen; hij gaf er de voorkeur aan in elk geval alle practicumlessen door te laten gaan, zodat de leerlingen alle experimenten konden uitvoeren. Docent C vindt het belangrijk de leerlingen in eigen tempo te laten werken, waardoor hij aan het eind van de lessenserie problemen kreeg met de organisatie van zijn theorielessen. De resultaten laten zien dat op een globaal niveau zeker sprake is van implementatie van de lessenserie. In de volgende paragraaf zal echter worden aangetoond dat de mate van implementatie voor de onderscheiden componenten en per hoofdstuk van het lesmateriaal sterk verschilt voor de bij het onderzoek betrokken docenten.

3.2 Verschil in gebruik tussen docenten

Zoals reeds vermeld is het praktijkprofiel een maat voor de implementatie van de courseware. Twee docenten (A en D) halen op het praktijkprofiel een score van rond de 50%. De score van de overige drie docenten varieert van ongeveer

23% tot 39%. Bij een drempelscore (het percentage interventies dat moet worden uitgevoerd voor een minimale uitvoering van de courseware) van 16.7% scoren alle docenten dus boven deze minimale uitvoeringsvariant.

Ook is nagegaan wat de score op het praktijkprofiel per onderscheiden component is, te weten *algemene vaardigheden, onderzoeksopzet, onderzoeksuitvoering, ontwikkelen van een begrippenkader en afleiden van natuurkundige regels*. In Figuur 2 worden de resultaten grafisch weergegeven.

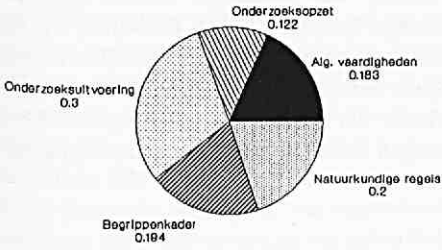
Uit deze gegevens blijkt dat de verdeling van de score van docenten A en D over de verschillende componenten van het curriculum min of meer overeenkomt met de ideale scoreverdeling. Docenten B, C en E echter blijken relatief veel interventies te gebruiken voor algemene vaardigheden (met name docent B). Docenten B en E (met name in klas E2) hebben relatief te weinig interventies gericht op opzet en uitvoering van onderzoek. Docenten C en E gebruiken te weinig interventies voor het afleiden van natuurkundige regels. Bovendien scoren ze op sommige componenten lager dan de drempelscore voor de betreffende component.

De score van docenten A en D is gezien de complexiteit van de innovatie acceptabel te noemen, met name omdat deze graad van implementatie bereikt is door alleen schriftelijk materiaal ter ondersteuning aan te bieden. Docenten A en D vertonen een begeleidings- en planningsgedrag dat ondersteunend is t.a.v. onderzoekend leren. Voor de score van de overige docenten moet worden vastgesteld dat deze laag is, afneemt gedurende de uitvoering van de lessenserie en niet overeenkomt met de ideale verdeling per component. Het begeleidings- en planningsgedrag van deze docenten laat met het oog op onderzoekend leren nog te wensen over.

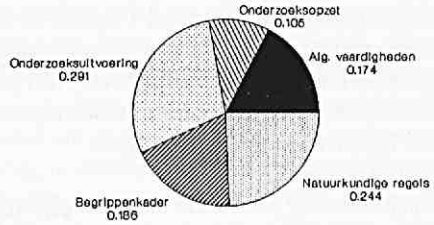
3.3 Factoren die de implementatie beïnvloeden

Technische problemen. Tijdens de lesobservaties blijkt dat met name school I te kampen heeft met technische problemen tijdens de uitvoering van de courseware, samenhangend met de door hen zelf ontwikkelde temperatuursensoren. Docenten B en C ervaren dat hun lessen hierdoor zeer verstoord worden. Alhoewel docent A met dezelfde problemen geconfronteerd wordt, geeft hij te kennen zich daaraan in het begin weliswaar gestoord te hebben, maar gedurende

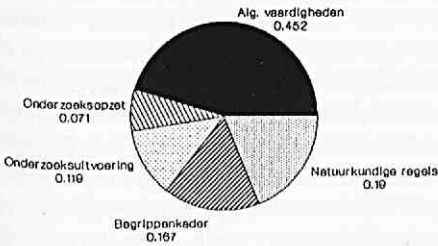
**mate van implementatie
componenten - ideaal**



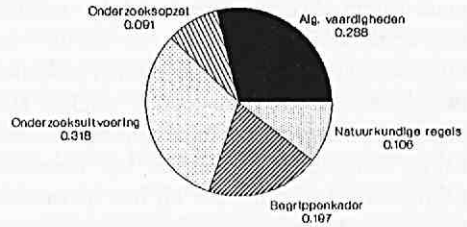
**mate van implementatie
componenten - docent A**



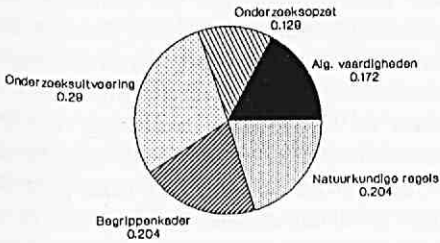
**mate van implementatie
componenten - docent B**



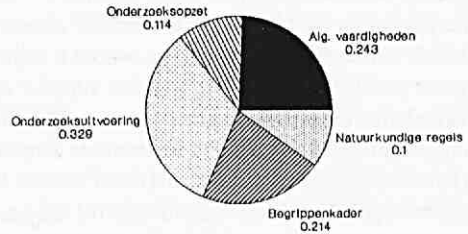
**mate van implementatie
componenten - docent C**



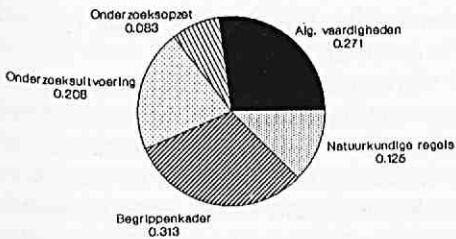
**mate van implementatie
componenten - docent D**



**mate van implementatie
componenten - docent E(1)**



**mate van implementatie
componenten - docent E(2)**



Figuur 2. Mate van implementatie per component

het verdere verloop van de lessenserie de problemen te hebben geaccepteerd. Gezien de technische problemen is de wisseling van Technisch Onderwijs Assistent in school I tijdens de uitvoering van de lessenserie een extra handicap. Op school II heeft men slechts af en toe met problemen van technische aard te maken gehad. Deze bleken makkelijk te verhelpen.

Perceptie van de docent. De interviewgegevens leveren informatie over de perceptie van de docenten omtrent de praktische bruikbaarheid van de courseware op de dimensies 'instrumentality', 'congruence' en 'cost' (Doyle en Ponder, 1977-78).

'Instrumentality' heeft betrekking op de mate waarin de courseware duidelijk is met het oog op lesvoorbereiding en -uitvoering. Het blijkt dat vier van de vijf docenten het programmapakket eenvoudig in gebruik vinden. Het begeleidend materiaal voor de leerlingen is duidelijk. Drie (B,C, E) van de vijf docenten gebruiken het leerlingboek bij hun lesvoorbereiding en niet de docentenhandleiding. Slechts twee docenten (A en D) gebruiken ook de docentenhandleiding bij hun lesvoorbereiding. Deze docenten vinden de aanwijzingen in de docentenhandleiding praktisch en bruikbaar.

'Congruence' houdt verband met de mate waarin de courseware past bij de lespraktijk van de docent en past bij het gebruikelijke natuurkunde curriculum. Docent E van school II blijkt grote problemen te hebben met het verschil in benadering en aanpak van de courseware (participatiegericht) en zijn eigen vertrouwde aanpak (vakstructuurgericht). Docent D van school II daarentegen vindt de begeleidende rol die van hem verwacht wordt plezierig. Docenten B en C van school I geven aan problemen te hebben met de organisatie van hun lessen. Beide docenten refereren aan hun eigen practicumlessen en vinden dat tijdens practica leerlingen zelfstandig moeten werken, wat voor hen inhoudt dat de docent weinig aan gerichte begeleiding doet.

'Cost' houdt verband met de beoordeling van de opbrengsten ten opzichte van inspanningen die de docent moet leveren om de courseware uit te voeren. Alle docenten zijn het erover eens dat de courseware motiverend is voor leerlingen. Ze verschillen van mening wat betreft de andere baten. Vier van de vijf (behalve A) vinden dat leerlingen zelfstandig leren werken, ze

betwijfelen echter of leerlingen na de lessenserie een systematische benadering van opzet en uitvoering van onderzoek kunnen hanteren. Sommige docenten (D en E) wijzen op het belang van een andere manier van kijken naar en omgaan met grafieken. Docenten A en C benadrukken het inzicht dat leerlingen hebben gekregen in de rol van de computer bij natuurkunde.

4 Discussie

In dit onderzoek is nagegaan welke kenmerken van courseware bijdragen aan de implementatie van een natuurkunde-curriculum waarin het leren van onderzoeksvaardigheden centraal staat. Met name de praktische kwaliteit van de courseware is onderzocht. Er is vastgesteld dat docenten die met een dergelijke dubbele innovatie te maken krijgen, niet alleen voor de opgave staan om een nieuw medium in hun onderwijs in te passen, maar dat van hen plannings- en begeleidingsgedrag wordt verwacht dat de invoering van de courseware ondersteunt.

De resultaten van het onderzoek laten zien dat op een globaal niveau de courseware is geïmplementeerd: docenten voeren de lessenserie in de gewenste volgorde uit. Kijken we echter nauwkeuriger naar het plannings- en begeleidingsgedrag van docenten dan zien we dat voor drie van de vijf docenten dit te wensen overlaat. Alhoewel het oordeel van de docenten omtrent de verschillende onderdelen van de courseware betrekkelijk positief is, blijkt dat de docenten aanzienlijk verschillen in het gebruik van de docentenhandleiding. Met name deze handleiding is, door concrete handelingsaanwijzingen te geven, bedoeld als ondersteuning voor de begeleidende en stimulerende rol die van docenten wordt verwacht. De resultaten van dit onderzoek suggereren dat docenten blijkbaar niet gewend zijn om een docentenhandleiding te gebruiken bij hun lesvoorbereiding (zie ook Van der Geest, 1991).

Van de drie docenten die de handleiding niet (nauwelijks) gebruiken, blijkt dat docent E van school II grote problemen heeft met het verschil in benadering en aanpak van de courseware (participatiegericht) en zijn eigen vertrouwde aanpak (vakstructuurgericht). Die problemen blijken onoverbrugbaar. Daarnaast heeft deze

docent ook grote moeite met de bediening van het programma, alhoewel hij zegt daar tijdens de concrete lesuitvoering geen last van te hebben ('de leerlingen wisten zelf wel hoe het programma bediend moest worden'). Docenten B en C van school I ervaren de technische problemen waarmee ze te kampen hebben als een grote belemmering voor hun lesuitvoering. Van groter belang lijkt echter de schijnbare overeenkomst tussen de benadering van natuurkunde-onderwijs in de courseware en het gebruikelijk natuurkunde-curriculum in school I. In beide curricula staan practicumlessen centraal. Deze ogenschijnlijke gelijkheid lijkt verhullend te zijn voor de herkenning van essentiële verschillen, zoals de noodzakelijkheid van gerichte begeleiding van de leerlingen door de docent bij de courseware lessen. Beide docenten vinden dat tijdens practica leerlingen zelfstandig moeten werken, wat voor hen inhoudt dat de docent weinig aan gerichte begeleiding doet. Eén van de docenten drukt dat uit door te zeggen 'tijdens het practicum moet ik bij wijze van spreken zo weg kunnen lopen'. De twee docenten die de docentenhandleiding gebruikt hebben, hebben tijdens hun les-uitvoering ook met technische problemen te kampen gehad (docent A) of te maken gehad met een zwakke overeenkomst van de courseware met de gebruikelijke lespraktijk (docent D). Desalniettemin slagen zij er tot zekere hoogte in die problemen te overbruggen. Dit hangt waarschijnlijk samen met het feit dat zij de docentenhandleiding waardevol achten om: (1) de lesvoortgang te kunnen bewaken, dan wel (2) de didactische aanwijzingen in de handleiding.

Docenten die de afstand tussen de vernieuwing met hun vertrouwde lespraktijk als erg groot ervaren, of docenten die de verschillen tussen de vernieuwing en hun praktijk niet herkennen, zijn niet geneigd van hun vertrouwde strategie van lesvoorbereiding af te wijken. Docenten die wel gebruik maken van een handleiding zien voor zichzelf voordelen bij het gebruik van de handleiding. Zij zijn bereid af te wijken van de hun vertrouwde aanpak bij de voorbereiding van lessen.

Dit onderzoek laat zien dat voor de integratie van courseware in onderzoekend natuurkunde-onderwijs naast zorgvuldig ontworpen curricula (Nakhleh, 1994, Linn & Songer, 1991) ook expliciet aandacht nodig is voor begeleiding

van docenten in hun nieuwe rol. Het lijkt wenselijk om naast concrete handelingsaanwijzingen (procedurele specificaties) in een docentenhandleiding, een op implementatie gerichte vorm van nascholing aan te bieden. Steeds vaker wordt duidelijk dat traditionele vormen van nascholing onvoldoende mogelijkheden bieden om docenten te ondersteunen bij verandering in didactisch gedrag en verandering van opvattingen (Van den Berg, 1996). Meer dan traditionele vormen van nascholing lijken docentnetwerken, met daarbij de mogelijkheid tot directe communicatie via e-mail, mogelijkheden te bieden voor docenten tot reflectie op onderwijsopvattingen en didactisch handelen (DiMauro & Gal, 1994; Bos, Krajcik & Patrick, 1995).

De resultaten van dit onderzoek ondersteunen de opvatting dat in welke vorm van nascholing dan ook de aandacht niet primair op de technologie dient te worden gericht (zie ook Kerr, 1991; Owen, 1992). Procedureel gespecificeerd materiaal kan een rol spelen als demonstratie- en oefenmateriaal (Roes, 1994), maar ook dienen als een gezamenlijke activiteit van een docentnetwerk, waarbinnen collega-docenten elkaar feedback geven en coachen. Het ontwikkelde innovatieprofiel kan worden gebruikt om de reflectie op de opvattingen van docenten over hun onderwijs en de vernieuwing waar het om gaat, te expliciteren.

Noot

- 1 Er is gebruik gemaakt van het programma IP-COACH, ontwikkeld door CMA, Nieuwe Achtergracht 170, 1018 WV Amsterdam.

Literatuur

- Akker, J.J.H. van den (1988). The teacher as learner in curriculum implementation. *Journal of Curriculum Studies*, 20, 47-55.
- Akker, J.J.H. van den, & Voogt, J.M. (1994). The use of innovation and practice profiles in the evaluation of curriculum implementation. *Studies of Educational Evaluation*, 20 (4), 503-512.
- Berg, E. van den (1996). *Effects of inservice education on implementation of elementary science*. Academisch proefschrift. Enschede: Universiteit Twente.

- Bos, N.D., Krajcik, J.S., & Patrick, H. (1995). Telecommunications for teachers: supporting reflection and collaboration among teaching professionals. *Journal of Computers in Mathematics and Science Teaching*, 14 (1/2), 187 - 202.
- Cohen, L., & Manion, L. (1980). *Research methods in education*. London: Croom Helm.
- DiMauro, V., & Gal, S. (1994). The use of telecommunications for reflective discourse of science teacher leaders. *The Journal of Science and Technology*, 3 (2), 123-135.
- Doyle, W., & Ponder, G.A. (1977-1978). The practicality ethic in teacher decision making. *Interchange*, 8 (3), 1-12.
- Driver, R., Asoko, H., Leach, J. Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, 23, 5-12.
- Gagné, R.M. (1966). Elementary science: A new scheme of instruction. *Science*, 155, 49-53.
- Geest, T.M. van der (1991). *Tools for teaching writing as a process. Design, development, implementation and evaluation of computer-assisted writing instruction*. Academisch proefschrift. Enschede: Universiteit Twente, Faculteit Wijsbegeerte en Maatschappij-wetenschappen.
- Grift, W. van der (1987). *De rol van de schoolleider bij onderwijsvernieuwingen*. Den Haag: VUGA.
- Kerr, S.T. (1991). Lever and Fulcrum: Educational technology in Teachers' thought and practice. *Teacher College Record*, 93 (1), 114 - 136.
- Kuiper, W.A.J.M. (1993). *Curriculumvernieuwing en lespraktijk*. Academisch proefschrift. Enschede: Universiteit Twente.
- Linn, M.C. (1987). Establishing a research base for science education: Challenges, trends and recommendations. *Journal of Research in Science Teaching*, 24 (3), 191-216.
- Linn, M.C., & Burbulus, N.C. (1993). Construction of knowledge and group learning. In K.Tobin (Red), *The practice of constructivism in science education* (pp. 91-121). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Linn, M.C., & Songer, N.B. (1991). Teaching thermodynamics to middle school students: What are appropriated cognitive demands? *Journal of Research in Science Teaching*, 28(10), 885-918.
- Millar, R. (1987). Towards a role for experiment in the science teaching laboratory. *Studies in Science Education*, 14, 109-118.
- Nakhleh, M.B. (1994). A review of Microcomputer-Based Labs: How have they affected science learning? *Journal of Computers in Mathematics and Science Teaching*, 13 (4), 368-381.
- Owen, M.(1992). A teacher-centred model of development in educational use of computers. *Journal of Information Technology for Teacher Education*, 1 (1), 127 - 137.
- Roberts, N. (1988). Computers and science instruction. In N. Roberts, N.C. Carter, S.N. Friel & M.S. Millen (Eds.), *Integrating computers into the elementary and middle school* (pp. 123-151). Englewood Cliffs, NJ: Prentice-Hall.
- Roes, M. (1994). *Het ontwerpen van nascholing aan de hand van curriculummaterialen*. Paper gepresenteerd op de Onderwijs Research Dagen, Utrecht.
- Shafto, S., & Capper, J. (1987). Microcomputer-based Laboratories. *Teaching, Learning and Technology*, 1 (2), 4-9.
- Striley, J. (1987). The computer as lab partner: Classroom experience gleaned from one year microcomputer-based laboratory use. *Journal of Educational Technology Systems*, 5 (3), 225-236.
- Thornton, R.K., & Sokoloff, D.R. (1990). Learning motion concepts using real-time microcomputer-based laboratory tools. *American Journal of Physics*, 58 (9), 858-867.
- Voogt, J. (1990). Courseware evaluation by teachers - an implementation perspective. *Computers in Education*, 14 (4), 299-307.
- Voogt, J.M. (1993). *Courseware for an inquiry-based science curriculum. An implementation perspective*. Doctoral dissertation. Enschede: Universiteit Twente.
- Voogt, J.M. (1996). New information technology in science education. In Tj. Plomp & D.P. Ely (Eds.), *International Encyclopedia of Educational Technology* (pp. 551-555). Oxford: Pergamon Press.
- Yin, R.K. (1989). *Case study research. Design and methods*. Beverly Hills, CA: Sage.

Manuscript aanvaard 23-1-1997

Auteur

J.M. Voogt. Universitair Docent, Vakgroep Curriculumtechnologie, Faculteit Toegepaste Onderwijskunde.

Adres: Universiteit Twente, Postbus 217, 7500 AE Enschede.

Abstract

Courseware for an inquiry-based science curriculum: an implementation perspective

J.M. Voogt. *Pedagogische Studiën*, 1997, 74, 130-141.

The integration of courseware in the science curriculum can be considered as a complex educational innovation, especially for teachers. Courseware using a Microcomputer-based Laboratory environment was designed. To promote to the implementation of the courseware special attention was paid to a teacher guide with concrete suggestions.

A case-study in two schools (five teachers/six classes) aimed at assessing whether the implementation of the courseware was successful. The results of the study indicated that the teachers differed profoundly to the extent in which they implemented the courseware. All teachers used the student textbook in planning and executing their lessons. Only two teachers used the teacher guide. These teachers benefited from the guide which resulted in implementation of the courseware more closely to the 'ideal' profile of implementation as intended by the designers. Reasons for the way the courseware was implemented are inferred and discussed.

Congres Dyslexie '96

Sinds in 1984 de Stichting Dyslexie Nederland (SDN) werd opgericht door Prof. dr. J.J. Dumont, hebben zich in Nederland (en één maal in Antwerpen) elf vrij grootschalige congressen Dyslexie afgespeeld. Het doel van dit soort congressen is om bij een breed publiek begrip te kweken voor leerlingen met ernstige lees- en spellingproblemen in basis- en voortgezet onderwijs in Nederland en Vlaanderen, en om de deelnemers op de hoogte te brengen van de laatste ontwikkelingen op onderzoeks- en praktijkgebied. Het laatste congres (Dyslexie '96) vond op 5 juni 1996 plaats in de Groninger Martinihal. Het congres was in samenwerking met de Stichting Dyslexie Nederland georganiseerd door de afdeling Orthopedagogiek van de Vakgroep Pedagogiek en Onderwijskunde van de Rijksuniversiteit Groningen en de afdeling KNO van het Academisch Ziekenhuis te Groningen; er was één sponsor: SONY Nederland. Er waren meer dan 1200 deelnemers, in meerderheid uit het onderwijsveld. Er werden 14 lezingen gehouden waarvan 2x3 in parallelsessies. Daarnaast bood een themamarkt een overzicht van bestaande en nieuwe behandelmateriaal en -methoden voor ernstige lees- en spellingproblemen.

Zoals aangegeven in de openingstoespraak (Van den Bos, RUG) waren de lezingen te ordenen langs de volgende theoretische en praktische denklijnen.

Definities

Al heel lang – d.w.z. vanaf het begin van deze eeuw – is er internationale wetenschappelijke en praktische belangstelling voor het verschijnsel 'dyslexie'.

Alternatieve termen zijn congenitale of aangeboren legasthenie, woordblindheid, en (specifieke) ontwikkelingsdyslexie. Wat willen deze termen in de kern alle aanduiden? Een ernstig probleem met het vlot en accuraat leren identificeren of produceren (spellen) van gedrukte/geschreven woorden. Met 'probleem met het *leren* ..' wordt aangegeven dat de persoon met dyslexie wel degelijk in een normale

instructiesituatie verkeert of heeft verkeerd, en dat het *niet* gaat om het verlies (tengevolge van een ongeval of ziekte) van een reeds verworven vaardigheid, of om analfabetisme in de zin van 'nooit onderricht gehad in lezen en schrijven'. Naast een kern bevatten de definities van dyslexie doorgaans allerlei toevoegingen, variërend van vrij algemene oorzakelijke hypothesen tot nogal gerichte uitspraken over verstoringen in bepaalde processen, bijvoorbeeld:

Dyslexie berust niet op defecten van zintuigen of van de algemene intelligentie, maar op een van de norm afwijkende ontwikkeling van bepaalde cognitieve processen. Deze van de norm afwijkende ontwikkeling van cognitieve processen die aan lezen en spellen ten grondslag liggen, is meer bepaald door neurologisch-constitutionele invloeden dan door omgevingsfactoren.

"Dyslexia is a specific language based disorder of constitutional origin, characterized by difficulties in single word decoding, usually reflecting insufficient phonological processing abilities" (Deze definitie is in 1995 geformuleerd door een ledencomité van de Amerikaanse 'Orton Dyslexia Society').

Welke definitie men ook kiest, steeds zijn er discussies over geweest. Een punt dat nogal eens vergeten wordt, is dat de functie van die definities voor de diverse betrokken partijen (onderwijsambtenaren, schoolleiders/leerkrachten, wetenschappers) nogal uiteen kan lopen. Ook verschillen wetenschappers onderling nogal vaak van mening. Een voorbeeld van een bijzonder controversieel punt is geweest de achteloze omdraaiing van bepaalde redeneringen in oorspronkelijke definities. Bijvoorbeeld: de aanname dat de stoornis '*niet* berust op defecten van zintuigen of van de algemene intelligentie' betekent niet dat *alleén* van dyslexie gesproken kan worden in geval van leesproblemen bij mensen met een normale intelligentie of intacte zintuigen!

Controverses zijn vooral ook ontstaan doordat veel studies die de bovengenoemde 'problematiese cognitieve processen' specificcerden zo ongeveer het alleenrecht claimden.

Voorbeelden: Dyslexie is een aandachtsstoornis; dyslexie vindt haar oorsprong in een vertraagde, gebrekkige en verstoorde taalontwikkeling. Het is echter bepaald niet zo dat *alle* dyslectici aandachts-, taal- en spraakmoeilikheden hebben of hadden. Omgekeerd is het ook niet zo dat alle kinderen met spraak-taal-moeilijkheden dyslectisch worden of zijn (Goorhuis-Brouwer, AZG, ging hier in haar lezing op in).

Net zoals de relatie tussen spraak-taalontwikkeling en dyslexie vrij complex is, geldt dat ook ten aanzien van dyslexie en het visuele waarnemings- en voorstellingsstelsel. Is er bij (sommige) dyslectische personen sprake van functionele stoornissen in bepaalde modules van deze systemen? Overigens wordt door diverse auteurs niet gesproken in termen van stoornissen maar eerder van een *andere* denk-ontwikkeling. In het Nederlandse werk van Maria Krabbe en Nel Ojemann en het vertaalde werk van de Amerikanen Davis en Braun komen tal van observaties voor die suggereren dat tijdens het lezen opgeroepen betekenisvoorstellingen van de woorden de *woordbeelden* vaak wegdrukken. Davis spreekt van vervormde woordwaarnemingen op grond van desoriëntatie (meerdere perspectieven nemen). Deze desoriëntaties werken in ruimtelijke situaties en beroepen voordelig, maar onvoordelig bij het omgaan met een verbaal symboolstelsel als het alfabet. Vele vragen n.a.v. deze typering zijn echter nog onbeantwoord. Naast sterke kanten van beelddenkers moeten er ook zwakke kanten zijn. Zijn dat precies dezelfde zwakke kanten als die van de niet- of minder duidelijk beelddenkende dyslectici?

Wat kunnen we n.a.v. dergelijke discussiepunten concluderen? De meest veilige conclusie is dat de alomvattende verklarende theorie die bovendien recht doet aan de heterogeniteit van de probleembeelden, er nog niet is. Een voor de praktijk zeer belangrijke stap is een 'beschrijvende' en ook inperkende definitie van dyslexie. De eerste lezing op het congres (door Ruijssenaars, RUL) begon dan ook met het plaatsen van het begrip dyslexie in het perspectief van definitieveranderingen, diagnostiek, behandeling en orthopedagogische aspecten. Impliciet in deze en vele andere lezingen was ook de roep om kwaliteitsverhoging van het lees-spellingonderwijs.

Van jong naar oud

Naarmate leeftijd en scholing toenemen, wordt van het lezende kind steeds meer geëist op schriftelijk gebied. Naarmate eventuele leesproblemen maar niet opgelost worden, lopen kinderen steeds groter risico's op 'onremedieerbaarheid'. De leeftijdslijn begon bij de bijdragen van Bus (RUL) en Van den Broeck (RUL) die het accent legden op de voorschoolse en vroegschoolse periode waarin het kind in contact komt met schriftelijke taal. Door Van den Broeck werd benadrukt dat daarbij vooral het leerproces van de relaties tussen eenheden zoals grafemen en fonemen belangrijk is. Ook Verhoevens (KUN) bijdrage ging over jonge lezers (groep 3) en de effecten van een recent ontwikkeld ondersteuningsprogramma (ELLO) bij jonge probleemlezers. Terwijl Koning (Stichting PRAVO) aandacht vroeg voor een *procesgerichte* aanpak van de spellingvaardigheid van basisschoolleerlingen – een aanpak die bovendien toegepaste spelling i.p.v. 'dictegedrag' moet bevorderen – werd later op de dag de leeftijdslijn verder doorgetrokken naar de oudere leerling in regulier en speciaal basisonderwijs (bijv. Van der Leij, VU; Aarnoutse, KUN). De leerlingen in het voortgezet onderwijs stonden meer centraal in de hoofdstukken van Hacquebord (RUG) en Loykens (Stichting Molendrift).

Leerproblemen en leerstoornissen

Hoeveel zwakke tot zeer zwakke lezers (*waaronder* de dyslectici) zijn er? Uitgaande van recente en landelijk genormeerde tests is dit ongeveer 15%. Hoe komen we aan dat getal? Door de standaardafwijkingen van de tests af te trekken van het landelijk gemiddelde en beneden dat statistische punt te spreken van een subnormale (zwakke tot zeer zwakke) prestatie.

Nu kan op lokaal niveau, op een willekeurige school, het aantal leerlingen dat onder het genoemde statistische punt scoort meer en minder dan 15% bedragen. Wat veroorzaakt deze fluctuatie? Het antwoord is dat de omgevingsfactoren of de leerlingfactoren, of beide soorten factoren, die ten grondslag liggen aan de groepsprestatie en de individuele verschillen, op de bewuste school niet helemaal representatief zijn voor de totale landelijke populatie. Wat de oorzakelijke factoren betreft, zijn mis-

schien de doelen en lesplannen anders dan bij andere scholen; misschien moeten ten gevolge van het WSNS-beleid ineens veel leerlingen binnen boord gehouden worden, of heeft de school te maken met; anders gemotiveerde leerlingen (of -leraren!), andere sociaal-culturele invloeden, of eenvoudigweg met relatief veel 'zwakke' of juist 'begaafde' leerlingen.

Hoe het ook zij, er zal in veel gevallen een kritische blik op het leescurriculum, de lesomstandigheden, het thuismilieu en de communicatie tussen school en thuis nodig zijn. Vooral Stevens (UU) ging hierop in. En in *alle gevallen* is extra onderwijsinspanning nodig om het aantal kinderen met leesproblemen zo vroeg mogelijk en zoveel mogelijk te beperken. Er zijn namen voor een dergelijke aanpak. In Nederland spreekt men van *zorgverbreding* en *zorg op maat*. De Amerikaan Howard Adelman spreekt van speciale onderwijsstrategieën in de vorm van *gepersonaliseerd* ('personalized') onderwijs. Het laatstgenoemde houdt in dat een betere afstemming wordt nagestreefd tussen onderwijsvormen en individuele behoeften op omgevings-, tempo-, capaciteits- en vooral op *motivatiegebied*. Het blijkt (vooral benadrukt in de lezingen van Ruijsseenaars en Verhoeven) dat deze extra 'remediërende' hulp in het stadium van *beginnend lezen* het aantal probleemlezers fors kan terugbrengen. Soortgelijke berichten komen uit Amerika (zie bijv. de verwijzing naar het onderzoek van Vellutino in de Scientific American van augustus 1995 (rubriek Science and the Citizen).

Er wordt geschat dat er na de extra ingezet- te *aanvankelijke* interventieprogramma's nog een groep van 1 tot 4% kinderen met leesproblemen overblijft die niet gunstig reageert op de individu-gerichte extra inspanningen. Voor deze leerlingen is de toepassing van personalisatie van onderwijs of zorgverbreding niet voldoende om het leerprobleem te verminderen of op te heffen. Dan móet bij deze groep – conform de toevoegingen in de oorspronkelijke definities van dyslexie – wel sprake zijn van in het cognitief en neurologisch systeem van de kinderen liggende blokkades. In de traditionele leerstoornissenliteratuur geldt voor deze kinderen pas het label 'leerstoornis' of 'dyslexie'. Ruijsseenaars ging in zijn lezing in op de verschuivende deskundigheden die aan de orde zijn op het continuüm van de begrippen leer-

problemen en leerstoornissen.

Wil de echte dyslecticus opstaan?

Dyslexie bestaat! We zouden bijna zeggen: ondanks definities. Het is voor menig leerkracht in het regulier onderwijs een ervaringsfeit om af en toe een leerling in de klas te hebben die maar niet van z'n leesprobleem af te helpen is. Gelukkig gebeurt dat niet vaak (op LOM- en MLK-scholen echter helaas redelijk vaak), en gelukkig zijn er ook succesverhalen. We moeten natuurlijk erg oppassen om zodra er een succes geboekt wordt, de leerling niet meer dyslectisch te noemen. Dat zou een al te ondiepe interpretatie van de beschrijvende definitie zijn. Evenzeer moeten we ervoor oppassen om *alleen* die leerlingen dyslectisch te noemen bij wie de woordidentificatievaardigheid nog steeds niet verbetert na jaren remediërend of behandelend onderwijs.

In twee lezingen (Van der Leij en Struiksma) werd de problematiek van 'kern'dyslexie uiteengezet. Anders dan dat succes al vroeg in de behandeling blijkt (Struiksma, Pedologisch Instituut Rotterdam), is op gedragsniveau niet goed aan te geven waarin succesvol- en niet-succesvol behandelde kinderen van elkaar verschillen. We kunnen dus nog maar slecht de toekomst van individuele leerlingen voorspellen. Eén ding weten we echter zeker, als er niets gedaan wordt, leren ze er 'geen letter bij'. Het pakket leesgerichte en persoonlijkheidsgerichte hulp dat door deze inleiders beschreven werd, kan daarom niet serieus genoeg genomen worden.

Behandel-dilemma's en andere vragen

Van Bon concentreerde zich in zijn lezing op een dilemma dat in de praktijk vaak voorkomt: Hoe lang moet er doorgegaan worden met decodeergerichte training? Moet op een gegeven moment niet op 'compenserende' instructies worden overgegaan? Terecht stelde Van Bon dat decodeertraining *wezenlijk blijft* in het bestrijden van woordidentificatieproblemen en dat dyslexie *niet* verdwijnt wanneer deze decodeertraining *vervangen* zou worden door onderwijs in begrijpend lezen, tekstgericht lezen e.d.

Daarnaast moet echter, conform modellen met aandacht voor *alle* facetten van lezen, de taalbegripscomponent van het begrijpend lezen

in de gaten gehouden en gevoed worden door onderwijs en training. Hierbij is dus geen sprake van 'of-of' maar eerder van 'en-en'. In Aarnoutse's lezing werd aandacht gevraagd voor de inrichting van het begrijpend leesonderwijs aan probleemlezers van de basisschoolleeftijd. Het lezen van probleemlezers in het voortgezet onderwijs was het onderwerp van Hacquebords lezing. Hacquebord besprak een tekstbegrip-toets die leidt tot een lezerstypologie en aansluitende onderwijsstrategieën in de vorm van hulpprogramma's.

Loykens betoogde vervolgens dat behandeling en ondersteuning van leerlingen met 'problemen met de talen' zich ook vaak dient te richten op het veranderen van de *leerstijl* van de leerling en zij besprak een methodiek tot gedragsverandering.

De twee laatste lezingen van de congresdag gingen respectievelijk over de vraag naar de invloed van dyslexie op de persoonlijkheidsontwikkeling (Van Leeuwen, Paedologisch Instituut Duivendrecht) en over een typisch pedagogisch onderwerp: 'Ouders en leraren, twee werelden' (Stevens). Van Leeuwen trok o.m. de conclusie dat, alhoewel dyslexie op de kinderleeftijd vaker dan verwacht samengaat

met 'externaliserende' gedragsproblemen, dyslectici op latere leeftijd niet vaker delinquent of agressief gedrag vertonen dan niet-dyslectici. Stevens, ten slotte, besprak het onderwerp 'leerlingen opdelen in twee groepen' in de context van verschillende definities en opvattingen die 'school' en 'thuis' hebben over ontwikkeling van het kind, opvoeding en onderwijs op school en opvoeding thuis. Door zijn appèl op: 'vertrouwen van leraren in eigen kunnen, hun vermogen om de leerling in zichzelf te doen geloven, niet toegeven aan de systeemeigen neiging de leerlingen in twee groepen op te delen' verschaft Stevens deze congresdag de nodige balans tussen kijken naar het leren van kinderen, naar de leermiddelen (curricula en materialen), en naar de houding en het denken van hun leraren.

Aan het slot van deze kroniek vermelden we nog dat de meeste op het congres gehouden lezingen hoofdstukken vormen in een in 1996 te verschijnen boek (titel: *Dyslexie '96*; uitgever: Garant) onder redactie van K.P. van den Bos en D.R. van Peer.

K.P. van den Bos en D.R. van Peer

Boekbespreking

L. Vygotkij

Cultuur en ontwikkeling

(Samenstelling, inleiding, vertaling en annotatie door R. van der Veer) Boom, Amsterdam/Meppel 1996

237 pagina's (incl. noten, literatuur, register). f 44,50, ISBN 90 5352 272 7

Een recensie van een vertalingbundel met enkele teksten van Vygotkij stelt de recensent voor een apart probleem. Richt hij zich op een kritische bespreking van de geselecteerde teksten van de oorspronkelijke auteur (Vygotkij in dit geval), of richt hij zich op wat de samensteller/vertaler (Van der Veer in dit geval) daarmee gedaan heeft? Beide natuurlijk, is wellicht de eerste reactie, maar omdat het hier om wat oudere teksten gaat (geschreven in de jaren twintig en dertig) ligt dat hier toch iets ingewikkelder. Het werk van Vygotkij is inmiddels al door velen doorgelicht en op zijn merites, beperkingen en tekortkomingen geëvalueerd (men zie onder andere Wertsch, 1987; Van der Veer & Valsiner, 1990). Het zou niet veel zin hebben die commentaren hier over te gaan doen.

Vygotkij's werk moet gezien worden in de context van de turbulente jaren twintig en dertig in de toenmalige Sovjet-Unie, in relatie tot het intellectuele klimaat van die tijd, de inzichten die toen bediscussieerd werden en – niet te vergeten – de relatief korte loopbaan van Vygotkij. Vygotkij's werk draagt overduidelijk de sporen van die omstandigheden waaronder het ontstaan is.

Het bijzonder aan dit boek is nu dat de samensteller – een Vygotkij specialist bij uitstek – door een zeer minutieus notenapparaat zelf al kritisch commentaar geeft op de teksten van Vygotkij, waardoor onduidelijkheden en soms ook onjuistheden aan het licht komen. De recensie van de zes gekozen teksten van Vygotkij zelf is dus eigenlijk door de samensteller zelf van deze bundel al geschreven. Dit alleen al maakt het boek bijzonder de moeite waard.

Het boek begint met een kort 'Ten geleide', geschreven door de dochter van Vygotkij (Gita Vygodskaja), inmiddels een bejaarde dame maar nog steeds een geziene gast op

vele conferenties. In een beknopte, maar heldere inleiding geeft Van der Veer zelf vervolgens een uiteenzetting over het leven, het werk en de denkwereld van Vygotkij. Daarbij gaat hij ook in op de betekenis van Vygotkij's werk in het licht van huidige discussies over onderwijs en opvoeding. Ten slotte wordt ook kort een legitimering van de teksteuze gegeven.

Het moet niet eenvoudig geweest zijn om uit het omvangrijke oeuvre van Vygotkij zes teksten te kiezen die niet te lang zijn, de veelzijdigheid van Vygotkij's werk illustreren en bovendien ook nog actualiteitswaarde hebben. Het mag gezegd worden dat de samensteller uitstekend geslaagd is in die keuze, hoewel de hier gepresenteerde selectie ook geen compleet beeld geeft van Vygotkij's rijke oeuvre (we missen bijvoorbeeld essays over de emoties, persoonlijkheid, defectologie).

In het eerste hoofdstuk (*Het bewustzijn als probleem voor de gedragspsychologie*) laat Vygotkij zien dat functionaliteit een belangrijke factor is in het ontstaan van het menselijk bewustzijn. Hij maakt duidelijk dat daarin zowel sociale en cultuurhistorische krachten een rol spelen, als neurofysiologische. Vygotkij's betoog laat zien dat voor hem een verklaring van de menselijke ontwikkeling altijd recht moet doen aan de culturele en de biologische dimensie. In het laatste hoofdstuk van het boek (*Hersenen en gedrag*, hoofdstuk 6) werkt hij zijn positie nog verder uit. Door de moderne technologische mogelijkheden is de kennis van de neurologische systemen de laatste decennia weliswaar sterk gegroeid en is er grote vooruitgang geboekt ten opzichte van wat Vygotkij wist. Toch is zijn positie in het debat over bewustzijn, hersenen en menselijke activiteit geenszins verouderd. Zijn probleem van de verklaring voor de samenhang tussen hersenen en bewustzijn is immers nog steeds niet opgelost. Echter, Vygotkij zocht naar een holistische oplossing waarin de functionaliteit van bewustzijn en hersenen ten opzichte van de menselijk activiteit bewaard blijft. Hij verzette zich tegen een atomistische benadering die dit probleem probeert op te lossen door het verknippen in allerlei deelprocessen en mechanismen en de aparte bestudering daarvan. Het belang van deze holistische benade-

ring (waarin het geheel de betekenis van de delen bepaalt) is voor Vygotskij een centraal methodologisch principe dat hij op meerdere van zijn onderzoeksterreinen probeert toe te passen. In feite zijn alle in dit boek opgenomen hoofdstukken voorbeelden van deze aanpak. Juist deze methodologische benadering – hoe problematisch misschien nog – maakt Vygotskij's bijdrage ook in het debat over heren en bewustzijn tot een actuele positie.

In zijn essay over de *Gestaltpsychologie* (hoofdstuk 2) laat Vygotskij zien dat juist het genoemde holistische methodologische principe een reden is waarom hij het redelijk goed kan vinden met de Gestaltpsychologie van zijn dagen. Het streven om een verklaring te construeren van menselijk handelen en ontwikkeling waarbij het dualisme tussen subject en object doorbroken is, lag zowel de Gestaltpsychologen als Vygotskij na aan het hart. Ook deze problematiek is in de sociale wetenschappen nog steeds een actueel probleem. Allerlei meer recente pogingen om bij de benadering van deze problematiek terug te grijpen op het werk van (bijv.) Heidegger hebben ongetwijfeld te maken met dit verzet tegen dualistische verklaringen. In zijn essay over de Gestaltpsychologie beargumenteert Vygotskij dat holistisch-structurele verklaringen van menselijke activiteit en ontwikkeling ook de historiciteit van de structuren zelf onder ogen moeten zien. Juist dat punt dreigt door de Gestaltpsychologen over het hoofd te worden gezien.

De invloed van de Gestaltpsychologie is ook vandaag de dag – in bepaalde cognitivistische en handelingspsychologische theorieën – nog merkbaar. Dat maakt Vygotskij's voorzichtige kritiek op de Gestaltpsychologie zoals hij die toen kende, juist ook zo interessant voor onze tijd. Wie het artikel goed leest en met name let op de uitspraken die gedaan worden over betekenisontwikkeling, zal er ook nog wel argumenten in vinden die relevant zijn voor de huidige discussie over contexten in het onderwijs (als structuren die de betekenisontwikkeling ondersteunen). Dit zelfde thema is op een indringende manier terug te vinden in het hoofdstuk over *De rol van de omgeving* (hoofdstuk 5). Het laat zien dat de psychologische betekenis van de omgeving vooral is gelegen in de manier waarop deze

door het kind wordt beleefd. We kunnen dan ook nooit op grond van oppervlakkig inschatting of zelfs objectieve analyse bepalen of iets een psychologisch relevante omgeving ('context') zal zijn. Dit nog steeds zeer behartenswaardige hoofdstuk laat zien dat de relatie van het kind met de omgeving verandert in de loop van zijn ontwikkeling. Hoewel Vygotskij niet erg specifiek is over de aard van deze ontwikkeling, blijft het voor ons huidige onderwijs een uiterst relevante notie. In feite sluit de hedendaagse discussie over de notie van de leidende activiteiten (Leont'ev/El'konin) hierbij aan.

De hoofdstukken 3 en 4 gaan meer direct over aspecten van kinderlijke ontwikkeling in relatie tot onderwijs en opvoeding. In hoofdstuk 3 (*Het probleem van de taal en het denken van het kind in de theorie van Piaget*) vinden we een tekst terug die misschien bekend is als hoofdstuk twee uit Vygotskij's monografie "Denken en spraak". Vygotskij gaat daar kritisch in op het werk van Piaget. Ook deze tekst is nog bijzonder actueel nu we juist de honderdjarige geboortedag van Piaget en Vygotskij gevierd hebben. In 1996 zijn diverse conferenties gewijd aan dit thema. Het is duidelijk dat de meningen over de verhouding tussen Vygotskij's en Piagets theorieën nog steeds verdeeld zijn. Valsiner besluit zijn editorial in het speciale Piaget-Vygotskij nummer van *Culture & Psychology* (Vol. 2, nr 3) met de opmerking: "All in all, it is my hope that the present Special issue puts to rest the artificially exaggerated contrasts between Lev Vygotsky and Jean Piaget. Despite the obvious differences in their areas of interests, they shared the same systemic, developmental ethos that takes the open-endedness of developmental processes seriously. Whether that ethos becomes phrased in terms of 'dialectical synthesis' or 'progressing equilibration' is largely a matter of personal preferences of the theoreticians in the construction of their terminology".

Valsiner benadrukt dus de verwantschap tussen beide benaderingen, maar dat is geenszins een algemeen gedeelde opvatting. In zijn slotrede op de Vygotsky-conferentie (september '96 in Genève) waarschuwt Bruner juist tegen een al te gemakkelijke fusie tussen beide theorieën, aangezien ze op verschillende

methodologische vooronderstellingen berusten. De Piagetiaanse benadering vertrekt volgens Bruner vanuit een nomothetische opvatting over ontwikkeling, de Vygotskiaanse benadering volgt meer een interpretatieve-narratieve lijn. Bruner pleit ervoor om beide benaderingen volgens hun eigen vooronderstellingen te bestuderen en te ontwikkelen en ze niet ongefundeerd te vermengen. Een soortgelijke conclusie trekt Obuchova in het Vygotskij-nummer van het Russische tijdschrift *Voprosy Psichologii* (1996, nr 5). Ik vermoed dat in de vergelijkende analyse van het werk van Vygotskij en Piaget de door Van der Veer vertaalde en becommentarieerde tekst een belangrijke bron zal blijven.

In hoofdstuk vier (*De dynamiek van de verstandelijke ontwikkeling van de leerling in relatie tot het onderwijs*) lezen we Vygotskij's analyse van de mogelijkheid tot ontwikkelingsstimulering door onderwijs. In dit kader ontwerpt hij het begrip zone van naaste ontwikkeling dat in de huidige onderwijswetenschap zo ongeveer tot kernbegrip geworden is voor grote groepen onderzoekers, ontwikkelaars, opleiders en andere practici. Vygotskij zelf blijft hier nog hangen in allerlei technische oplossingen voor het probleem. De notie is inmiddels (met behulp van begrippen als 'scaffolding', en 'participatie') veel verder ontwikkeld. Hoewel dit hoofdstuk in bepaalde opzichten wellicht wat verouderd is, is het goed om toch weer eens te overdenken wat Vygotskij's eigenlijke drijfveer achter het ont-

werp van dit begrip 'zone van naaste ontwikkeling' was. Zijn ideeën om de diagnostiek van ontwikkeling en leervorderingen te enten op indicaties over ontwikkelingsmogelijkheden (in plaats van toetsprestaties) van leerlingen zijn nog steeds actueel.

Het is een aanwinst dat Van der Veer deze teksten voor het Nederlandstalige publiek toegankelijk gemaakt heeft in een perfect vertaalde én zorgvuldige geannoteerde uitgave. Vygotskij's originele teksten zijn meestal heel toegankelijk, maar niet altijd even bloemrijk geschreven. Van der Veer is erin geslaagd om zelfs dat in de vertalingen te handhaven.

Natuurlijk zijn we zo langzamerhand toe aan een kritische evaluatie van Vygotskij's wetenschappelijke werk, maar daarvoor is het nodig om ook precies te weten wat hij zelf gezegd heeft. Een ieder die Vygotskij's benadering (in onderzoek en opleiding) serieus neemt, zou dit boek onder handbereik moeten hebben.

B. van Oers

Literatuur

- Veer, R. van der, & Valsiner, J. (1990). *Understanding Vygotsky. A Quest for Synthesis*. Oxford: Blackwell.
- Wertsch, J. (1987). *Vygotsky and the social formation of mind*. Cambridge, MA: Harvard University Press.

Mededelingen

Promoties

Universiteit van Amsterdam
Promovendus: Drs. G.J. Caris
Titel proefschrift: Laten praten. Onderzoek naar leidinggevende vaardigheden en kwaliteiten van gespreksvoering op consultatiebureau van Kruiswerk Amsterdam
Promotor: Prof.Dr. J.M.A. Hermanns
Datum: 25 maart 1997

Katholieke Universiteit Nijmegen
Promovendus: H. Biemans
Titel proefschrift: Fostering activation of prior knowledge and conceptual change
Promotor: Prof.Dr. P.R.J. Simons
Datum: 3 maart 1997

Katholieke Universiteit Nijmegen
Promovendus: Drs. J. Visser
Titel proefschrift: Lezen op AVI-niveau
Promotor: Prof.Dr. R.M. van den Berg
Co-promotores: Dr. C.A.J. Aarnoutse en Dr. M.J.M. Voeten
Datum: 20 juni 1997
Tijd: 13.30 uur

Universiteit Twente
Promovendus: Drs. M.G. Roes
Titel proefschrift: Nascholing op basis van lesvoorbeelden in de context van curriculumvernieuwing
Promotor: Prof.Dr. Tj. Plomp
Datum: 21 februari 1997

Universiteit Twente
Promovendus: Drs. M.W.J.A. Braam-Voeten
Titel proefschrift: Van taaltheorie naar taaltherapie: evaluatie en analyse van een congruentie- en capaciteitentherapie bij kinderen met taalstoornissen
Promotor: Prof.Dr. S. Dijkstra
Datum: 14 maart 1997

Universiteit Twente
Promovendus: Drs. B.G. Doornekamp
Titel proefschrift: Probleemoplossen bij techniek: een studie naar effectieve domeinspecifieke ontwerprichtlijnen

Promotor: Prof.Dr. W.J. Nijhof
Datum: 21 maart 1997

Universiteit Utrecht
Promovendus: Drs. B.C. de Jong
Titel proefschrift: Jan Ligthart (1859-1916). Een schoolmeester-pedagoog uit de Schilderswijk
Promotores: Prof.Dr. J.D. Imelman en Prof.Dr. B. Rang
Datum: 26 november 1996

Universiteit Utrecht
Promovendus: Drs. M.C. van der Wende
Titel proefschrift: Internationalising the curriculum in Dutch higher education: An international comparative perspective
Promotores: Prof.Dr. N. Lagerweij en Prof.Dr. Tj. Plomp

Recent gestart onderzoek

Universiteit Twente
Titel project: Toepassing van adviserende 'Intelligent agents' in virtuele studie-omgevingen
Projectleiding: Prof.Dr. J.C.M.M. Moonen en Dr. I.P.F. De Diana

Titel project: Het leren oplossen van logica-problemen
Projectleiding: Prof.Dr. S. Dijkstra

Titel project: Het ontwikkelen van ontwerpvaardigheden voor onderwijskundige instrumentatie via 'scaffolding' in een gehyperlinkte informatie-omgeving
Projectleiding: Prof.Dr. J.C.M.M. Moonen, Dr. B.A. Collis en Dr.Ir. P.W. Verhagen

Titel project: De organisatie van het opleiden en leren van zelfsturende teams
Projectleiding: Prof.Dr. J. Scheerens

Titel project: Optimaal ontwerpen van adaptieve tests en tests voor meerdimensionale vaardigheden
Projectleiding: Prof.Dr. W.J. van der Linden

Oratie

Dr. J.G.L. Thijssen, benoemd tot hoogleraar bij de vakgroep Onderwijskunde van de Universiteit Utrecht, met de leeropdracht 'Bedrijfsopleidingen en beroepsonderwijs' zal op 13 mei 1997 zijn ambt aanvaarden met het uitspreken van een rede, getiteld: *Leren om te overleven*.

Plaats: Academiegebouw van de Universiteit Utrecht, Domplein 29, Utrecht

Tijd: 16.15 uur

Benoemingen

Prof.Dr. J. Terwel is per 1 oktober 1996 benoemd tot hoogleraar Onderwijspedagogiek aan de Vrije Universiteit te Amsterdam. Hij is daarnaast sinds 1994 hoogleraar Onderwijskunde (in het bijzonder curriculum en instructie van exacte vakken) aan de Universiteit van Amsterdam.

Dr. S.J. van Zolingen is per 1 januari 1997 benoemd als Universitair Docent bij de vakgroep Curriculumtechnologie van de Faculteit der Toegepaste Onderwijskunde aan de Universiteit Twente.

Inhoud andere tijdschriften

Pedagogisch Tijdschrift
21e jaargang, nr. 6, 1996

Voor het Nederlands als wetenschappelijke taal, door de redactie

Pedagogisch commentaar, door R. Vandenberghe

Onderzoek naar praktijkkennis en beoordeling van docenten, door D. Beijaard en N. Verloop
Leren oplossen van wiskundige problemen in kleine groepen. Effecten van een didactische interventie, door D. Hoek, P. van den Eeden en J. Terwel

Education gets real – learning from experience, door P. Standish

Tijdschrift voor Onderwijsresearch
21e jaargang, nr. 4, 1996

Inleiding op het thema leerstijlen, door J. Vermunt en J. van der Sanden

Verschillen in leerstijlen: sekse en sekse-identiteit, door S. Severiens en G. ten Dam
Leeromgevingspercepties en leerstrategieën van eerstejaars studenten sociale wetenschappen, door R.F.A. Wierstra en E.P.M. Beerends
Leerstijlen en het opzoeken van informatie in tekst, door E.T. Stoutjesdijk en J.J. Beishuizen
Leerstijlen en zelfstandig leren in het voorportaal van het studiehuis, door A. Roosendaal en J. Vermunt
Gedecentraliseerde besluitvorming binnen MBO-instellingen: een meetmodel voor domeinen van besluitvorming, door P.L.J. Boerman en T.M. Berger

Tijdschrift voor Orthopedagogiek
36e jaargang, nr. 1, 1997

Spelstimulering in het speciaal onderwijs, door P.H.M. van Eijk en A. van der Leij
Nederlandstalige Harterschalen voor het vaststellen van het zelfconcept van kinderen en adolescenten, door J.W. Veerman en P.B. Swennenhuis

De opvoeding van jonge drielingen, door S.M.J. van Hekken, J.T. Cremers-Rijsemus, B. Doornebal-Mulder en J.F. Orlebeke

Tijdschrift voor Orthopedagogiek
36e jaargang, nr. 2, 1997

Uit je zelf. Via ontwikkelend onderzoek naar een leergang woordenschat en communicatieve vaardigheden voor het VSO/praktijkonderwijs, door G.M. Hoogenkamp

Onderzoek naar validiteit Heerlense Ernst Taxatie Schaal (L-HETS) in klinische praktijk, door H.J. Pelzer, W.J.P.J.M. Steerneman en E.E.J. de Bruyn

'Dyslexie. Afbakening en behandeling': een reactie vanuit de praktijk, door W.A. Wassenaar

Dyslexie: van rapport naar praktijk. (Reactie op Wassenaar), door A.J.F.M. Ruijssenaars

Tijdschrift voor Orthopedagogiek
36e jaargang, nr. 3, 1997

Zorgen voor morgen ... Langetermijneffecten bij kinderen met een hersentumor, door B.W. Duijzer-van Dijk
Vroegtijdige interventie van leesproblemen

met nadruk op metacognitie en leesmotivatie,
door L. Verhoeven en H. van de Ven
Stabilisatie deelname aan LOM en MLK onaf-
hankelijk van WSNS-beleid, door Y.J. Pijl

Ontvangen boeken

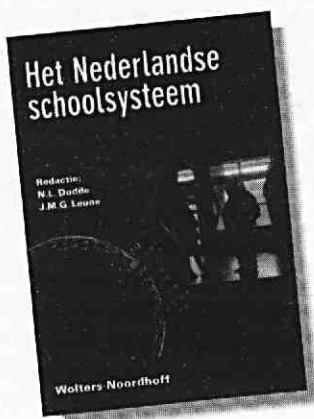
Buijnsters, P.J., & Buijnsters-Smets, L., *Bibliografie van Nederlandse school- en kinderboeken 1700-1800*. Uitgeverij Waanders, Zwolle, 1997, f 125,-.

Dam, G.T.M. ten, Hout, J.F.M.J. van, Terlouw, C., & Willems, J. *Onderwijskunde hoger onderwijs. Handboek voor docenten*. Uitgeverij Van Gorcum, Assen, 1996, f 69,50.

Lange, G. de, *Leren wandelen aan Vaders hand. Christelijke opvoeding: theorie en praktijk*. Uitgeverij Buijten & Schipperheijn, Amsterdam, 1997, f 32,50.

Pijnenburg, H.M., Rijswijk, C.M. van, Ruijsenaars, A.J.J.M., & Veerman, J.W. (red.), *Pedologisch Jaarboek 1996*. Uitgeverij Eburon, Delft, 1996, f 39,50.

Standaardwerken over onderwijs

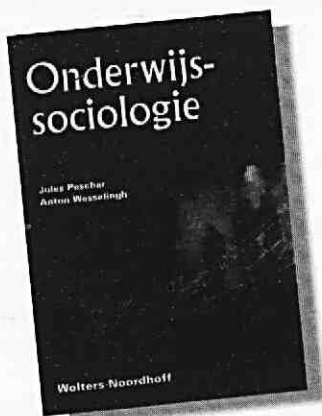


Het Nederlandse schoolstelsel

N.L. Dodde, J.M.G. Leune

De uitgave *Het Nederlandse schoolstelsel* is een bewerking en dé opvolger van *Onderwijs, bestel en beleid* en *Schets van het Nederlands schoolwezen*. Deze actuele uitgave geeft een glashelder overzicht van het Nederlandse schoolstelsel en onderwijsbeleid van dit moment.

Een handzaam naslagwerk met belangrijke vakliteratuur, gebaseerd op wetenschappelijk relevante informatie.
ISBN 90 01 24500 5 ing 300 p f52,50



Onderwijssociologie

J.L. Peschar, A.A. Wesselingh

Een boeiende introductie in de onderwijssociologie. Aan de orde komen belangrijke vraagstukken over onderwijs, opvoeding en maatschappij. Een opmerkelijke uitgave met verrassende inzichten. Inclusief de nieuwste theorieën en recente resultaten uit onderzoek.

ISBN 90 01 70691 6
ing 2e dr 294 p f62,00

Ook verkrijgbaar via de boekhandel

Wolters-Noordhoff
Postbus 58
9700 MB Groningen
Telefoon (050) 522 65 22

**Wolters
Noordhoff**

444/6448e-96

AB-10p: 15 OCT 6441 AB-1C

Het belang van Nederlandstalige wetenschappelijke tijdschriften

Met de onderstaande verklaring sluiten wij ons als redactie van Pedagogische Studiën aan bij een initiatief dat enige tijd geleden door de redactie van Pedagogisch Tijdschrift is genomen. Wij leggen daarbij overigens, zoals zal blijken, een aantal andere accenten.

Pedagogische wetenschappen en onderwijskunde zijn de laatste jaren geïnternationaliseerd. Het aantal buitenlandse publicaties, met name in het Engels, is sterk toegenomen. Onze participatie in internationale congressen en de contacten met buitenlandse collega's zijn duidelijk verbeterd. Kortom, de Nederlandse en Vlaamse pedagogiek en onderwijskunde zijn internationaal goed zichtbaar geworden. Een belangrijke stimulans hierin is uitgegaan van de onderzoekscholen, die buitenlandse publicaties als toelatingseis hanteren.

Deze positieve ontwikkeling heeft echter ook een schaduwzijde: de Nederlandstalige tijdschriften worden langzamerhand steeds meer weggedrukt als medium van wetenschappelijke communicatie. Het onvoorziene neven-effect van de toelatingseis van de Nederlandse onderzoekscholen en van vergelijkbare Vlaamse beoordelingsprocedures is, dat steeds minder in onze eigen taal wordt geschreven. De Nederlandstalige pedagogische en onderwijskundige tijdschriften hebben tot nu toe getracht een rol te vervullen als intermediair tussen wetenschappelijke vakgenoten en als intermediair tussen wetenschap en praktijk. De merkwaardige situatie ontstaat nu, dat het voor veel wetenschappelijke vakgenoten minder interessant wordt om in deze tijdschriften te publiceren, terwijl het duidelijk is dat de kwaliteit van deze tijdschriften minstens even goed is als die van menig wel erkend buitenlands periodiek. Onze verantwoordelijkheid naar de betrokken maatschappelijke velden mag niet op deze manier verwaarloosd worden. Het lijkt wel alsof we over onze eigen problemen nog slechts in buitenlandse tijdschriften met elkaar kunnen com municeren.

Er treden oneigenlijke verschuivingen op in

de wetenschappelijke belangstelling, enkel doordat bepaalde onderwerpen, hoewel ze dringend om onderzoek vragen, gemeden worden omdat ze minder makkelijk in buitenlandse tijdschriften worden opgenomen. Naast de hier reeds aangeduide functie van Nederlandstalige tijdschriften als 'intermediair' tussen wetenschap en praktijk hebben zij dus ook een belangrijke rol bij het rapporteren over dit type onderzoeksonderwerpen.

Wij pleiten daarom voor een herwaardering van de Nederlandstalige tijdschriften als wetenschappelijk forum. Niet door de internationalisering terug te draaien, maar door van elke pedagoog en onderwijskundige óók degelijke bijdragen in onze eigen taal te blijven verwachten. De onderzoekscholen moeten hun toelatingcriteria op dit punt bijstellen door, naast artikelen in de moderne vreemde talen, publicaties van hoog niveau in het Nederlands te eisen. Alleen zo kan het Nederlandstalig wetenschappelijk forum zijn functie naar behoren blijven vervullen.

De redactie

G. Staphorsius en N.D. Verhelst

Samenvatting

De CLIB (Cito leesindex voor het basis- en speciaal onderwijs) is een index voor zowel de leesvaardigheid van lezers als voor de leesbaarheid van teksten. Met behulp van de CLIB kunnen de leesvaardigheid van een lezer en de leesvaardigheid die een tekst vereist met elkaar vergeleken worden. Daardoor is het mogelijk om in de dagelijkse onderwijspraktijk de leesvaardigheid van leerlingen en de moeilijkheid van teksten op elkaar af te stemmen. Omdat de CLIB verwijst naar een domein van teksten die lezers met een bepaald vaardigheidsniveau met begrip kunnen lezen, noemen we de CLIB een domeingerichte index. De CLIB is in de eerste plaats een index voor de begrijpelijkheid en het begrijpen van teksten. Het onderzoek dat deze index opleverde past in een lange traditie. Veel minder traditioneel is het onderzoek naar een index voor de door teksten vereiste leestechniek.

Toch zou een index voor de vereiste leestechniek didactisch gezien van belang kunnen zijn. We denken dan vooral aan de leerlingen die aanzienlijk beter kunnen begrijpen dan decoderen. We hebben met het oog daarop een onderzoek uitgevoerd. Het leverde de CILT (Cito Index voor de leestechniek) op. Met behulp van een multi-pele regressievergelijking ($R = .90$) kan worden bepaald wat de CILT van een tekst is. Deze geeft aan hoe hoog de technische leesvaardigheid is, die een tekst vereist. De predictoren in de regressievergelijking zijn 'woordlengte' en 'woordfrequentie'. Het criterium voor leestechniek is het aantal woorden uit teksten dat in één minuut gemiddeld gelezen wordt door basisschoolleerlingen in groep 5 tot en met 8. Voor de domeingerichte bepaling van de leestechniek van leerlingen kan in principe elke valide toets voor de meting van de leestechniek worden gebruikt. Wij zijn er in geslaagd de scores op de toetsen Leestempo van het Cito te transformeren naar de CILT-schaal.

1 Inleiding: leesbaarheidspredictie

Het zoeken naar een index voor de leesbaarheid van teksten impliceert een zogenaamd leesbaarheidsonderzoek. Het eerste leesbaarheidsonderzoek is dat van Vogel en Washburne (1928). Zij slaagden erin een voorspeller van verschillen in begrijpelijkheid te ontwikkelen. Een leesbaarheidsonderzoek heeft in grote lijnen de volgende opzet. Van teksten in een steekproef uit een gegeven populatie bepaalt men de vereiste leesvaardigheid, ofwel de leesbaarheid. Vervolgens gaat men na door welke tekstkenmerken de vastgestelde verschillen in leesbaarheid tussen teksten verklaard kunnen worden. Men gebruikt daarbij de multi-pele regressie-analyse als techniek. Deze levert een multi-pele regressievergelijking of leesbaarheidsformule op. Daarmee kan de leesbaarheid van teksten worden voorspeld die behoren tot de populatie waaruit het onderzoeksmateriaal getrokken is. We zullen hier ter verduidelijking de grote lijnen van het onderzoek weergeven dat de CLIB opleverde (zie Staphorsius, 1994). In dit onderzoek bestond het onderzoeksmateriaal uit een steekproef uit de populatie non-fictie voor de jeugd. Het begrip 'leesbaarheid' werd nader bepaald met 'begrijpelijkheid', de mate waarin teksten een beroep doen op de vaardigheid begrijpend lezen. Om een maat voor de begrijpelijkheid te kunnen berekenen maakten we van elke tekst in de steekproef met behulp van de cloze-procedure (Taylor, 1953; zie ook Staphorsius, 1994) een cloze-toets. De variant van de procedure die werd gebruikt, schrijft de deletie van elk zevende woord voor. We vroegen aan leerlingen in groep 5 tot en met 8 te proberen de weggelaten woorden weer in te vullen. Het gemiddeld aantal juiste reconstructies, de gemiddelde cloze-score van de teksten in onze steekproef gebruikten we als criterium voor 'leesbaarheid'. Vervolgens gingen we na door welke onafhankelijke variabelen dit criterium het beste voorspeld werd. In verreweg het meeste leesbaarheidsonderzoek

zijn de onafhankelijke variabelen eenduidig te identificeren kenmerken van teksten. Veel gebruikte kenmerken zijn: de woordlengte, de bekendheid en frequentie van woorden, de lengte van zinnen, maar ook de frequentie van complexe syntactische structuren. In het leesbaarheidsonderzoek dat Staphorsius in 1994 rapporteert, vormen het criterium en (uiteindelijk) 4 onafhankelijke variabelen of predictoren de invoer van een multiële regressie-analyse. De predictoren zijn: het percentage frequente woorden (FREQ77; later in dit artikel zal nog blijken waarvoor deze afkorting precies staat), het percentage verschillende woorden (PTYPES), de gemiddelde woordlengte in letters (GWL) en het percentage zinnen (PZW; het aantal woorden is de percenteerbasis) als syntactische maat. De vergelijking ziet er als volgt uit:

$$\begin{aligned} \text{CLIB} = & \\ 46 - 6.603 \times \text{GWL} + 0.474 \times \text{FREQ77} - 0.365 & \\ \times \text{PTYPES} + 1.425 \times \text{PZW} & \end{aligned}$$

In de formule is de CLIB de voorspelde gemiddelde cloze-score (in de uitgaven voor het onderwijs werkt het CITO overigens met een transformatie van die score). Het praktische nut van deze regressievergelijking is, dat de leesbaarheid van andere teksten dan die in de steekproef, niet empirisch bepaald hoeft te worden, althans voorzover die teksten behoren tot de populatie waaruit het onderzoeksmateriaal afkomstig is. De vergelijking wordt in de literatuur ook vaak aangeduid met 'leesbaarheidsformule'.

In het onderzoek waarvan we hier verslag doen, is de leesbaarheid **niet als begrijpelijkheid**, maar als **veronderstelde leestech-niek** gedefinieerd. Het heeft niettemin in grote lijnen dezelfde opzet als het onderzoek naar de CLIB. Daarmee is ook de route voor het verslag van ons onderzoek naar de voorspelbaarheid van de door teksten voor de jeugd vereiste leestech-niek bepaald. In de volgende paragraaf gaan we eerst in op het criterium. We definiëren daar ook de populatie teksten en bespreken de selectie van ons onderzoeksmateriaal daaruit. In de derde paragraaf gaan we in op de predictoren. In paragraaf 4 rapporteren we de resultaten van de multiële regressie-analyse. In de vijfde paragraaf verkennen we een mogelijkheid om de uitkomsten van de regressiever-

gelijking die we in paragraaf 4 presenteren van een domein- of taakgerichte interpretatie te voorzien. Een score die domeingericht geïnterpreteerd kan worden, geeft aan welke taken een leerling al wel en welke taken hij nog niet met succes kan uitvoeren, of andersom, welke teksten geschikt zijn voor welke leerlingen.

2 Criterium voor de vereiste leestech-niek

Meting van de door teksten vereiste leestech-niek

Technisch lezen is het omzetten van grafemen in fonemen. Het belangrijkste aspect van de leestech-niek is – nadat de teken-klank-koppelingen geleerd zijn – de snelheid waarmee het leesproces geschreven woorden herkent ofwel de koppeling tussen woordvorm en de daarmee in het interne lexicon geassocieerde kennis tot stand brengt.

Een bekend en als betrouwbaar en valide geaccepteerd instrument voor de meting van de leestech-niek is de zogenaamde één-minuut-toets of kortweg: de EMT (Ballard, 1920, was voor wij weten de eerste die een één-minuut-toets gebruikte). Dit instrument bestaat in het algemeen uit een reeks van 100 tot 200 woorden. De maat voor de leestech-niek die het levert, is het aantal woorden dat in 60 seconden correct verklankt wordt. In ons land is het gebruik van de test van Brus en Voeten (1973) zeer verbreid, zowel in het basisonderwijs als in onderzoek. PPOON gebruikt een één-minuut-toets (Zwarts, 1990) en het project Leerlingvolg-systeem van het Cito heeft nog onlangs onder de naam 'Drie-Minuten-Toets' drie versies van dit type instrument uitgegeven (Verhoeven, 1995). Deze bestaat uit drie in moeilijkheid verschillende één-minuut-toetsen. Wij maken in ons onderzoek gebruik van een variant van de EMT. De reeksen woorden waaruit ons onderzoeksmateriaal bestaat, zijn afkomstig uit bestaande teksten. We gaan ervan uit dat als van een reeks A in één minuut leestijd meer woorden verklankt worden dan van een reeks B, de gelezen woorden in reeks A 'technisch' leesbaarder zijn dan de gelezen woorden in een reeks B.

Tabel 1
Gemiddelde, SD en bereik van de gemiddelde DMT3-scores per reeks (N = 48)

Gem. LVS-EMT	SD	Laagste waarde	Hoogste waarde
59.421	0.4185	58.580	60.358

Constructie en selectie

In ons onderzoek hebben we gebruik gemaakt van 48 reeksen van 150 woorden. De reeksen zijn afkomstig uit een steekproef van teksten, die op hun beurt weer komen uit een corpus, dat gevormd is ten behoeve van het CLIB-onderzoek (Staphorsius, 1994). Dit corpus bestaat uit een steekproef van 240 teksten 'fictie' en 240 teksten 'non-fictie' (zie voor een uitvoerige verantwoording van deze steekproef Staphorsius, Krom & De Geus, 1988). Uit elk van de subcorpora hebben we voor het onderzoek naar de 'technische' leesbaarheid aselect 24 teksten getrokken. Per tekst hebben we drie steekproeven van 50 woorden getrokken. In elke trekkingsronde begonnen we met een aselect bepaald woord van de eerste zeven woorden (de fragmenten in het corpus hebben een omvang van ongeveer 350 woorden). Vervolgens selecteerden we elk zevende woord. En dat drie keer. De gevolgde procedure – we zullen verder om bondig te kunnen verwijzen ook wel spreken van ETP, een afkorting van EMT-tekst-procedure – leidde uiteindelijk dus tot 48 woordreeksen. Net als bij Ballard (1920), maar anders dan in de hierboven met name genoemde één-minuut-toetsen zijn in deze reeksen ook functiewoorden opgenomen. Een ander gevolg van de ETP is dat dezelfde woordvormen meerdere keren in een reeks kunnen voorkomen (in de ETP is een stap opgenomen die voorkomt dat dezelfde woordvormen direct na elkaar in een reeks voorkomen).

We gaan er op grond van de gevolgde procedure van uit dat het gemiddelde en de spreiding

van de woordkenmerken in de 48 reeksen woorden een afspiegeling zijn van de verdeling van woordkenmerken in de populatie fictie en non-fictie voor de jeugd. Als ons onderzoek inderdaad resulteert in een voorspeller van de door teksten vereiste leestechneek, kunnen we de belasting van de woordherkenner in het leesproces door een gegeven tekst uit de populatie fictie en non-fictie voor de jeugd voorspellen op basis van kenmerken van de woorden in die tekst.

Afname van de 48 reeksen van 150 woorden

We hebben omdat we niet alle 48 woordreeksen aan één proefpersoon of aan één groep proefpersonen konden afnemen en toch de ingezette leestechneek per reeks moesten controleren, het 'randomized blocking' principe toegepast. Dat vereist een voormeting van de leestechneek van proefpersonen. Op 5 scholen in een steekproef die we gebruikten voor eerder onderzoek hebben we aan alle leerlingen in groep 4 tot en met 8 de kaart met de moeilijkste reeks woorden in de Cito-LVS Drie-Minuten-Toets afgenomen (omdat dit de derde kaart in de Drie-Minuten-Toets is, spreken we verder kortweg van DMT3). Op basis van de scores op deze toets hebben we 48 groepen van 36 leerlingen samengesteld met eenzelfde gemiddelde score en standaardafwijking op de DMT3. Elke groep van 36 kreeg een van de 48 reeksen toegewezen.

In Tabel 1 geven we het gemiddelde van de gemiddelde DMT3-scores per verzameling. De gemiddelde DMT3-score per groep van 36 leerlingen komt overeen met de gemiddelde leestechneek in het basisonderwijs aan het einde van groep 5.

Resultaten

In Tabel 2 geven we het gemiddelde en de spreiding van de criteriumwaarden (het aantal in 1 minuut gelezen woorden) van alle 48 reeksen woorden, maar we onderscheiden in de

Tabel 2
Gemiddelden en SD van de criteriumwaarden in fictie en non-fictie samen en op fictie en non-fictie afzonderlijk

Fictie en non-fictie		Fictie		Non-fictie	
Gemiddelde	SD	Gemiddelde	SD	Gemiddelde	SD
81.16	6.73	81.84	5.55	80.48	7.79

tabel ook tussen de reeksen die het resultaat zijn van de toepassing van de EMT-tekst-procedure op de twee in onze steekproef te onderscheiden genres, fictie en non-fictie.

3 De predictoren

Het woordlengte-, het woordfrequentie- en het contexteffect

Welke kenmerken komen als predictoren in aanmerking? Het antwoord op deze vraag is uiteraard afhankelijk van de specificatie van het begrip 'technisch lezen' of 'leestechiek'. Interindividuele verschillen in leestechiek worden, nadat lezers de bestaande 'grafeem-foneem'-koppelingen geleerd hebben, bepaald door het tempo waarin lezers in staat zijn tot het herkennen van woorden. Zijn er kenmerken van woorden die het woordherkennings-tempo tijdens het lezen van 'lopende' teksten beïnvloeden en die voor de predictie van de door teksten vereiste leestechiek relevant kunnen zijn? In de literatuur wordt naar twee robuuste effecten verwezen: het woordlengte- en het woordfrequentie-effect. (zie Thomassen, Noordman & Eling, 1984, 1991; Rayner & Pollatsek, 1989; Just & Carpenter, 1987). Korte en frequente woorden worden sneller herkend. Behalve op het woordlengte- en frequentie-effect wijst de literatuur op het contexteffect. In de discussie over dit effect gaat het er niet in de eerste plaats om of de context het leestempo al dan niet beïnvloedt. Centraal staat de vraag of het contexteffect intra- dan wel postlexicaal is: faciliteert de context of faciliteren kenmerken van organisatieniveaus in de tekst hoger dan het woord (zin, alinea enzovoort) de leestaak tijdens of na het herkennen van woorden? Onderzoek wijst soms op een intra-lexicaal effect, maar vaker op een postlexicale invloed van de context (Rayner & Pollatsek, 1989). Voorzover het contexteffect inderdaad post-lexicaal is, hoeft het bij de selectie van de predictoren, noch bij de keuze van het criterium een belangrijke rol te spelen, want het gaat ons in de eerste plaats om verschillen in tempo voor zover die afhankelijk zijn van variabelen die inwerken op het gemak en de snelheid waarmee lezers decoderen.

Een intra-lexicaal contexteffect is het zogenaamde repetitie-effect. Het blijkt dat een

woord dat in een experiment herhaald wordt aangeboden, bij de herhaalde aanbiedingen sneller herkend wordt. Hudson, Bergman, Houtmans en Nas (1984) maken melding van onderzoek van Scarborough et al. (1977) en Scarborough et al. (1979) waaruit blijkt dat dit effect zelfs na een dag nog meetbaar is. Woordherkenningsmodellen verklaren het repetitie-effect uit een verlaging van de herkenningsdrempel. Dit effect wijst erop dat 'diversiteit van de woordenschat' als predictor in het leesbaarheidsonderzoek bruikbaar zou kunnen zijn. Operationalisering van deze variabele lijkt zinvol omdat door onze EMT-tekst-procedure hetzelfde woord meerdere keren in een reeks kan voorkomen.

Predictoren

De maten voor woordlengte in ons onderzoek duiden we aan met GWL en GWLG, terwijl we naar de maat voor woordfrequentie verwijzen met *FREQ77*. Als maat voor de diversiteit van de woordenschat in teksten gebruiken we *PTY-PES*.

GWL is de gemiddelde woordlengte in letters, het quotiënt van het aantal letters en het aantal woorden. GWLG is de gemiddelde woordlengte uitgedrukt in lettergrepen. Deze twee predictoren zullen elkaar, zo leert ons het leesbaarheidsonderzoek, sterk overlappen. Dat neemt niet weg dat GWL het criterium voor veronderstelde leestechiek nauwkeuriger zou kunnen voorspellen dan GWLG, of andersom. Om te kunnen bepalen welke van de twee de beste voorspeller is, betrekken we beide predictoren in ons onderzoek.

FREQ77 is de door ons gehanteerde maat voor 'woordfrequentie'. Deze is gebaseerd op het onderzoek van Staphorsius, Krom en De Geus (1988). Zij voerden op het zogenaamde P335-corpus, de hierboven al eerder genoemde steekproef van 480 teksten uit de populatie jeugdlectuur, frequentietellingen uit. Het resultaat van de tellingen is onder meer een naar frequentie geordende lijst van in het corpus voorkomende woorden. Als frequent definieren we de 998 meest frequente woorden. Het zijn woorden die twintig keer of vaker in het corpus voorkomen. Ze maken bijna 77 procent van de tokens in het corpus ($n = 202526$) uit. Het totaal aantal types = 18210; de gemiddelde frequentie is ruim 11. Voor de deellijst van 998

Tabel 3*Gemiddelde en SD predictoren in de woordreeksen fictie, non-fictie en fictie en non-fictie samen*

	Fictie en non-fictie		Fictie		Non-fictie	
	Gemiddelde	SD	Gemiddelde	SD	Gemiddelde	SD
GWL	4.45	0.38	4.36	0.27	4.55	0.46
GWLG	1.48	0.13	1.45	0.10	1.50	0.16
GFREQ77	78.18	6.63	79.32	6.35	77.03	6.84
PTYPES	75.32	6.04	75.10	6.06	75.52	5.99

woorden geldt een gemiddelde frequentie van ongeveer 156. We noemen deze lijst in het vervolg de *FREQ77*-lijst. *FREQ77* is het percentage tokens in een tekst dat in de *FREQ77*-lijst voorkomt.

PTYPES staat voor het percentage verschillende woorden. Deze maat voor de type/token-ratio, dit wil zeggen de verhouding tussen het aantal verschillende woordvormen en het totaal aantal woorden, is onze maat voor de diversiteit van de woordenschat.

Meetresultaten

In Tabel 3 geven we per te onderscheiden steekproef (fictie, non-fictie en fictie en non-fictie samen) het gemiddelde en de SD van de waarden voor *GWL*, *GWLG*, *FREQ77* en *PTYPES*.

De waarden zijn, met het oog op de vereiste homogeniteit, gebaseerd op tellingen in de eerste 80 woorden van de reeksen, het aantal dat overeenkomt met het gemiddeld aantal in de reeksen gelezen woorden.

Tussen haakjes: uit de literatuur en uit de praktijk van het (aanvankelijk) leesonderwijs is bekend dat woorden met bepaalde medeklinkercombinaties (bijvoorbeeld: sch..., schr...) moeilijker te verklanken zijn dan woorden zonder die combinaties. Toch hebben we 'lastige medeklinkercombinaties' niet als predictor in ons onderzoek betrokken. 'Gewone' teksten verschillen namelijk niet of nauwelijks als het om dit kenmerk gaat. Daarvoor hebben we een logische verklaring: 'lastige medeklinkercombinaties' is geen variabele die door schrijvers van 'gewone' teksten gemanipuleerd wordt. Voor deze verklaring vinden we steun in de empirie. Als maat voor 'lastige medeklinkercombinaties' kan de gemiddelde lettergreep-lengte per letter worden gehanteerd. Immers:

lastige medeklinkercombinaties zijn relatief lang, dus hoe groter de lettergreep-lengte, hoe vaker lastig te verklanken medeklinkercombinaties in een tekst zullen voorkomen. In ons onderzoeksmateriaal blijkt de lettergreep-lengte gemiddeld ± 3.1 te zijn. De standaardafwijking van dit gemiddelde is ± 0.09 . Uit de correlatie met ons criterium wordt de betekenis van de verschillen tussen de teksten in ons corpus wat deze predictor betreft duidelijk: er is geen correlatie. Voor andere woordkenmerken die de verklanking bemoeilijken geldt hetzelfde als voor 'lastige medeklinkercombinaties'. Als er onder deze kenmerken al potentiële predictoren schuil gaan (bijvoorbeeld niet-klankzuivere leenwoorden), dan bestaat er een aanzienlijke kans dat we die 'gevangen' hebben met de predictor *FREQ77*.

4 Multipelle regressie-analyse

In paragraaf 2 gaven we aan hoe de criterium-scores in ons onderzoek bepaald zijn. In paragraaf 3 presenteerden we vervolgens een overzicht van de onafhankelijke variabelen of predictoren, waarvan we op grond van gegevens uit de literatuur konden verwachten, dat ze een substantieel gedeelte van de variantie van het criterium zouden verklaren. In deze paragraaf rapporteren we de resultaten van de uitgevoerde multipelle regressie-analyse. In deze analyse gaat het er om de optimale gewichtsverhouding te bepalen waarmee de predictoren het criterium voorspellen. In Tabel 4 rapporteren we per te onderscheiden steekproef de correlatiecoëfficiënten tussen de predictoren onderling en ook de resultaten van de multipelle regressie-analyse. In de multipelle regressie-analyse bleek dat *PTYPES*' bijdrage aan de voorspelling van het criterium niet significant

Tabel 4

Resultaten regressie-analyses in de steekproeven fictie, non-fictie en in fictie en non-fictie samen

	Fictie en non-fictie				Fictie				Non-fictie			
	1	2	3	4	1	2	3	4	1	2	3	4
1 GWL	1.00				1.00				1.00			
2 GWLG	.95	1.00			.90	1.00			.96	1.00		
3 FREQ77	-.65	-.54	1.00		-.55	-.46	1.00		-.70	-.59	1.00	
4 PTYPES	.37	.27	-.44	1.00	.26	.06	-.48	1.00	.46	.40	-.41	1.00
5 CRIT	-.88	-.85	.73	-.39	-.86	-.84	.66	-.27	-.90	-.85	.78	-.49
INTERCEPT	114.49				126.84				110.35			
FREQ77	0.28				0.24				0.32			
GWL	-12.33				-14.59				-12.01			
R (R ²)	.90 (.82)				.89 (.79)				.92 (.85)			

was. Vandaar dat deze predictor geen deel uit maakt van de multipele regressievergelijkingen. Zoals verwacht is de overlap tussen GWL en GWLG zeer groot. In de vergelijking is geen plaats voor beide predictoren. GWL doet het als predictor iets beter dan GWLG. Daarom is GWL in de definitieve multipele regressie-analyse als predictor ingevoerd en niet GWLG.

Kruisvalidering

Een multipele regressie-analyse schat regressiecoëfficiënten uit de correlatiematrix die afkomstig is uit een steekproef van observaties; de schattingen kunnen zeer wel beïnvloed zijn door de specifieke eigenaardigheden van die steekproef. Als we de schattingen willen beschouwen en gebruiken als benaderingen van de regressiecoëfficiënten die in de populatie (van teksten) gelden, dan moeten we aantonen dat de schattingen niet al te veel zijn 'aangetaast' door de eigenaardigheden van de steekproef waaruit ze geschat zijn.

Een techniek die men daar vaak voor gebruikt is de kruisvalidatie. Deze komt er op

neer dat men de regressiecoëfficiënten schat uit één steekproef (de screeningssteekproef; Lord & Novick, 1968) en deze schattingen gebruikt om het criterium in een onafhankelijke steekproef (de calibratiesteekproef) te voorspellen. In het algemeen vindt men dat de multipele correlatie in de calibratiesteekproef lager uitvalt dan in de screeningssteekproef; een eventueel verschil wordt 'krimp' genoemd. Een kleine krimp is een teken van stabiliteit van de schatters van de regressiecoëfficiënten.

In ons onderzoek hebben we de teksten behorend tot de subcorpora 'Fictie' en 'Non-fictie' beurtelings de rol van screeningssteekproef en calibratiesteekproef toebedeeld, en aldus twee kruisvalidaties uitgevoerd. De geschatte regressiecoëfficiënten zijn weergegeven in Tabel 4. In Tabel 5 geven we een overzicht van de resultaten.

Uit Tabel 5 blijkt dat de krimp van de R van 'Non-fictie' naar 'Fictie' nihil is en dat er van 'Fictie' naar 'Non-fictie' sprake is van geringe krimp. We gingen ook na of de regressievergelijkingen voor het corpus 'Fictie' en 'Non-fic-

Tabel 5

Resultaten kruisvalidatie

calibratiesteekproef	screeningssteekproef		Non-fictie	
	Fictie	Non-fictie	Fictie	Non-fictie
R	0.89	0.88	0.92	0.92

tie' als aan elkaar gelijk kunnen worden beschouwd: de verschillen tussen beide vergelijkingen blijken niet significant te zijn ($\alpha = .05$). Dat wijst op de validiteit van de vergelijkingen in de afzonderlijke steekproeven.

Vanwege een grotere stabiliteit komt de regressievergelijking op basis van de data in de steekproeven fictie en non-fictie samen in aanmerking voor gebruik in de praktijk (zie Kerlinger & Pedhazur, 1973, blz. 284). We noemen deze vergelijking Cito Index Lees-techniek, afgekort CILT.

$$\text{CILT} = 114.49 + 0.28 \times \text{FREQ77} - 12.33 \times \text{GWL} \quad (1)$$

5 Van een relatieve naar een domeingerichte index

De CLIB, we spraken over deze index in de inleiding, verwijst als index van de leesbaarheid van teksten naar een subpopulatie of domein van leerlingen met een leesvaardigheid die minimaal vereist is om een tekst te kunnen lezen en, omgekeerd, als een index voor de leesvaardigheid, gemeten met de Clib-toetsen, verwijst de CLIB naar een domein van teksten. Als een leerling op de Clib-toetsen een CLIB van 42 heeft behaald, dan kan deze teksten begrijpen met een CLIB van 42 en lager. Kan ook de CILT van een domeingerichte betekenis worden voorzien? In ieder geval niet zonder een aanvulling op het onderzoek waarvan we in de vorige paragrafen verslag deden en waarin het uitsluitend te doen was om de 'ontwikkeling' van een relatieve index voor de leestech-niek. Een domeingerichte CILT vereist een ander (en 'duurder') criterium voor de leestech-niek. Nadat we ons CILT-onderzoek hadden afgerond, ontstond in het project dat een-voudig klassikaal afneembare toetsen voor de leessnelheid voor het Cito-Leerlingvolgsys-tem ontwikkelt (Toetsen Leestempo, Krom, 1996), de behoefte aan een domeingerichte index en de vraag is nu of we, gebruik makend van de resultaten van het CILT-onderzoek, de CILT alsnog van een domeingerichte betekenis kunnen voorzien. Dat is de vraag waar het in deze paragraaf om gaat. Op het praktisch nut van een domeingerichte CILT komen we in de bespreking terug.

Als we de CILT willen voorzien van een domeingerichte betekenis dan moeten we in staat zijn vast te stellen hoe groot de technische leesvaardigheid van een lezer moet zijn om met succes een tekst met een gegeven CILT te kunnen decoderen. Minder cruciaal voor een absolute interpretatie, maar wel handig is het als behalve de door teksten veronderstelde leestech-niek ook de leestech-niek van lezers in de CILT kan worden uitgedrukt. Ter bepaling van de gedachten geven we een mogelijke opzet van een onderzoek naar een domeingerichte CILT met als uitgangspunt de relatieve CILT. In zo'n onderzoek zouden we eerst kunnen vaststellen wat voor elke leerling in een steek-proef van leerlingen de moeilijkste tekst is in een steekproef van teksten die hij of zij met succes technisch kan lezen. Op deze manier kunnen we de leestech-niek van de proefperso-nen uitdrukken in een CILT, namelijk de CILT van de moeilijkste tekst die hij of zij kan deco-deren. Daarnaast zouden we met een valide en betrouwbare toets (bijvoorbeeld met de boven-genoemde toetsen Leestempo) kunnen vaststel-len wat de decodeervaardigheid van de proef-persoonen is. We kennen dan van elke leerling de CILT en de score op de decodeertoets. Als de steekproeven van teksten en leerlingen aan het doel van het onderzoek beantwoorden, kan worden nagegaan of vanuit de CILT de scores op de gebruikte toets kunnen worden voorspeld (of andersom).

Wij hebben een variant van het onderzoek met hierboven geschetste opzet uitgevoerd. Wij hebben gebruik gemaakt van gegevens uit een onderzoek dat Staphorsius (1994) uitvoerde naar de bruikbaarheid van een tekstenschaal voor de beoordeling van de leesvaardigheid van leerlingen en naar de validiteit van de CLIB als 'absolute' leesindex. Staphorsius vroeg leerkrachten (aantal = 87; aantal ver-schillende scholen: 21; totaal aantal leerlingen in groep 5 tot en met 8: 1844) de vaardigheid 'begrijpend lezen' en de 'leestech-niek' van hun leerlingen te bepalen met behulp van een tek-stenschaal. Deze schaal bestond uit 12 teksten met ongeveer dezelfde gemiddelde CILT (en een iets grotere spreiding) als in de 48 woord-reeksen in het 'CILT-onderzoek'. Aan de leer-krachten werd gevraagd de vaardigheid 'begrij-pend lezen' van elke leerling in hun klas te schatten door de moeilijkste tekst op de

tekstenschaal aan te geven die de leerling naar zijn of haar oordeel met begrip zou kunnen lezen. Bovendien vroeg Staphorsius de leerkrachten van elke individuele leerling aan te geven wat naar schatting de moeilijkste tekst was die deze technisch 'voldoende goed' zou kunnen lezen. In het onderzoek waren representatieve steekproeven leerlingen uit groep 5 tot en met 8 betrokken (zie Staphorsius, 1994). Alle leerlingen hadden ook de Clib-toetsen gemaakt zodat Staphorsius behalve over het leerkrachtenoordeel ook over een 'directe' maat voor de leesvaardigheid van elke leerling (uitgedrukt op de CLIB-schaal) beschikte. Van deze dataset maken we nu gebruik bij onze pogingen om de relatieve CILT om te zetten in een absolute CILT. We beschikken over een maat voor de leestechiek van de leerlingen via de tekstenschaal. Van alle leerlingen weten we zo wat volgens hun leerkracht de moeilijkste tekst is die zij 'technisch' kunnen lezen. Van die tekst kennen we ook de CILT. We weten via de direct bepaalde CLIB voor welke 'gemiddelde leerling' deze CILT geldt: de gemiddelde leerling midden groep 4, eind groep 4, midden groep 5 enzovoort. Via het onderzoek dat is uitgevoerd ter normering van onder meer de toetsen Leestempo (Krom, 1996) en de DMT3 kennen we de waarden voor het met deze toetsen bepaalde leestempo van deze 'gemiddelde leerlingen'. We weten nu bijvoorbeeld van de derde tekst op de tekstenschaal wat, uitgedrukt op de Leestempo-schaal, het gemiddelde leestempo is van leerlingen die deze tekst wel, maar de volgende op de tekstenschaal nog niet voldoende goed kunnen lezen. We merkten net al op dat de derde tekst net als de elf andere teksten ook een CILT heeft. We kunnen nu een analyse uitvoeren waarin we deze CILT-waarden regresseren op bijvoorbeeld de Leestempo-waarden van de 12 teksten. We zoeken dus de coëfficiënten A en B, zo dat

$$CILT_i = A + B \times LEESTEMPO_i + \varepsilon_i \quad (2)$$

waarbij de variantie van de residuen zo klein mogelijk is. Deze coëfficiënten bepalen de regressievergelijking waarmee we de waarde van een tekst op de Leestempo-schaal kunnen omzetten in een waarde op de CILT-schaal. Deze analyse hebben we inderdaad uitgevoerd

($r = .94$). We hadden uiteraard ook de gemiddelde Leestempo-scores als (domeingericht) criterium kunnen hanteren en kunnen nagaan welke combinatie van predictoren dit criterium het best voorspelde. De stabiliteit van de CILT op basis van 48 woordreeksen echter is groter dan de stabiliteit van een index gebaseerd op 12 teksten.

Van een leerling met een bepaalde score op de toetsen Leestempo (Krom, 1996) kunnen we nu vaststellen voor welke teksten deze vermoedelijk al wel een voldoende leestechiek heeft en voor welke teksten dat naar alle waarschijnlijkheid nog niet het geval is. Immers de scores op de toetsen Leestempo kunnen worden omgezet naar een CILT, terwijl met formule 1 ook de CILT van teksten kan worden bepaald. Dat maakt het mogelijk teksten (boeken, artikelen, brochures enzovoort) te vinden met een CILT die ongeveer overeenkomt met de via de toetsen Leestempo bepaalde CILT van de leerling. Zo is de CILT van een domeingerichte interpretatie voorzien. We hadden uiteraard dezelfde oefening kunnen uitvoeren met de LVS-EMT of een andere voor ons doel valide toets. Ons onderzoek heeft met dit resultaat een voorlopig einde. Voorlopig, want we zullen onze verkenningen van domeingerichte criteria voor de leestechiek voortzetten.

6 Bespreking

Lezers verschillen in leestechiek en teksten verschillen in vereiste leestechiek. In dit verslag hebben we een voorspeller van relatieve verschillen tussen teksten in vereiste leestechiek, geïntroduceerd. Ons onderzoeksmateriaal bestond uit 48 reeksen van 150 woorden. Deze zijn met behulp van de EMT-tekst-procedure op basis van teksten in een steekproef uit de populatie jeugdlectuur samengesteld. Als criterium voor de vereiste leestechiek hanteerden we het aantal woorden dat door de proefpersonen gemiddeld correct gelezen werd. De proefpersonen waren leerlingen uit het basisonderwijs met een leestechiek die overeenkomt met de gemiddelde leestechiek van de leerlingen aan het einde van groep 5. Als predictoren hanteerden we het percentage frequente woorden, het percentage verschillende woorden, de gemiddelde woordlengte in letters en in letter-

grepen. De multiële regressie-analyse resulteerde in de CILT, een regressievergelijking met het percentage frequente woorden en de gemiddelde woordlengte in letters als predictoren ($R = .90$; $R^2 = .82$). De bijdrage van de gemiddelde woordlengte in lettergrepen en van het percentage verschillende woorden aan de voorspelling van het criterium naast de andere predictoren was niet significant. Van willekeurige teksten kan, nadat het percentage frequente woorden en de gemiddelde woordlengte in letters in die teksten bepaald is, de CILT berekend worden. We zijn er in geslaagd onder meer de scores op de toetsen Leestempo (Krom, 1996) om te zetten in een CILT die aangeeft wat de moeilijkste tekst is die een leerling kan verklanken. Daarmee heeft de CILT, naast een relatieve, een domeingerichte betekenis gekregen. We komen nu toe aan de vraag naar de praktische waarde van de CILT. De onderwijspraktijk kan al beschikken over de CLIB. Kan het onderwijs met indexen als de CILT iets dat het met de CLIB niet kan?

Voor de meeste leerlingen in de fase van het aanvankelijk lezen zijn de teksten die ze te lezen krijgen conceptueel gezien triviaal terwijl de decodeerproblemen in verhouding nog groot zijn. Voor een grote meerderheid van de leerlingen smaakt de overwinning van die problemen dan vermoedelijk nog zó goed, dat de inhoudelijke eenvoud van het aangereikte leesgoed ze niet al te zeer stoort. De decodeerbaarheid neemt al tijdens het aanvankelijk leesonderwijs snel zo toe dat zelfs conceptueel

gezien nog te moeilijke teksten technisch binnen het bereik van die vaardigheid komen. Het lijkt er dan ook op dat de leestehnik (aspect: tempo) zich bij de meeste leerlingen ruwweg vanaf eind groep 4 kan ontwikkelen door het lezen van conceptueel gezien passende teksten. Voor zwakke decodeerders ligt dit vermoedelijk enigszins anders. Voor een aantal van deze lezers blijft de discrepantie bestaan, waarvan aan het begin van het leesonderwijs voor vrijwel alle leerlingen sprake is. Door hun matige leestehnik kan het gebeuren dat deze leerlingen niet of nauwelijks toekomen aan het begrijpen van teksten die conceptueel gezien geen enkel probleem vormen (Perfetti, 1985). Voor leerlingen met een grote discrepantie tussen leestehnik en conceptuele mogelijkheden, zo vermoedde men ook in het project 'LVS-Leestempo' (Krom, 1996), zouden indexen zoals de CILT van didactische betekenis kunnen zijn, vooral als zou blijken dat teksten met een ongeveer gelijke conceptuele lading verschillen in de leestehnik die ze veronderstellen. Tabel 6 geeft van deze verschillen een indruk, maar vereist eerst een nadere toelichting.

In Tabel 6 vergelijken we op basis van Staphorsius (1996) het Clib-niveau van teksten (totale aantal = 9991) met de index van het Katholiek Pedagogisch Centrum (Visser, Van Laarhoven en Ter Beek, 1994) voor de **leestehnik**, het zogenaamde AVI-niveau (in het onderwijs ook wel 'de AVI' genoemd) van dezelfde teksten. De teksten zijn afkomstig uit bijna 1500 verschillende bronnen (fictie en

Tabel 6
Gemiddelde, SD en bereik van AVI-niveau per Clib-niveau

Clib-niveau		AVI-niveau**			
		Gem.	SD	min.	max.
4	(266)*	5.56(B5)	0.99	4(E4)	10(>E6)
5	(1342)	7.14(E5)	0.97	5(E4)	10(>E6)
6	(2967)	8.56(E6)	0.88	5(B5)	10(>E6)
7	(3045)	9.63(>E6)	0.61	4(M4)	10(>E6)
8	(1477)	9.97(>E6)	0.18	8(B6)	10(>E6)

* Aantal teksten in het aangegeven Clib-niveau

** Bijvoorbeeld (B4) betekent: komt ongeveer overeen met het gemiddelde begin groep 4; (M4) = ongeveer gemiddelde midden groep 4; (E4) = ongeveer gemiddelde eind groep 4; >E6 = niveau dat na het moment E6 behaald wordt.

non-fictie) en hebben een gemiddelde lengte van ruim 420 woorden. De Clib-niveaus zijn niveau-aanduidingen voor het *begrijpen* van teksten. Per Clib-niveau kan het gemiddelde en de standaardafwijking van het AVI-niveau van de teksten worden afgelezen. Bijna 900 teksten behoren tot hogere of lagere niveaus dan de in de tabel opgenomen Clib-niveaus. Bijvoorbeeld Clib-niveau 4 verwijst naar teksten die geschikt zijn voor de gemiddelde leerling aan het einde van groep 4 in de basisschool. Voor de berekening van de gemiddelde AVI-index per Clibniveau hebben we alle teksten die boven AVI-niveau 9 uitkomen met een 10 gescoord en alle teksten lager dan niveau 5 met een 4. Op basis van de normgegevens van Visser, Van Laarhoven en Ter Beek (1994) hebben we achter de AVI-niveaus aangegeven welke 'gemiddelde leerling' op welk moment over een bepaald niveau beschikt. De momenten 'begin leerjaar' (september/oktober) en 'eind leerjaar' (maart/april) zijn gebaseerd op de gerapporteerde empirische gegevens. De waarden voor het moment 'midden leerjaar' zijn door ons op grond van deze empirische waarden geschat.

Uit Tabel 6 blijkt dat teksten met een vergelijkbare conceptuele moeilijkheid (uitgedrukt in Clib-niveau) onderling wat betreft de volgens het AVI-systeem vereiste leestehnik behoorlijk verschillen, soms zelfs 'jaren'. Een index voor de leestehnik maakt het mogelijk te midden van teksten met ongeveer dezelfde conceptuele moeilijkheid de teksten te identificeren met de laagste decodeermoeilijkheid. Met een domeingerichte index kan bovendien worden vastgesteld of zelfs de tekst met de laagste decodeerlading in vergelijking met de aanwezige leestehnik wel leesbaar is. Indexen voor de leestehnik kunnen zo bezien in het onderwijs aan leerlingen die specifiek op de leestehnik uitvallen een rol spelen. Uit onderzoek in de onderwijspraktijk echter moet blijken of dat ook een substantiële rol is. De uitkomst van dat onderzoek is in de eerste plaats afhankelijk van de vraag in hoeverre het domeingerichte criterium staat voor de minimaal nodige leestehnik en in de tweede plaats van de vraag hoe vaak de betreffende leerlingen inderdaad aan een bij hun leesdoel passende en conceptueel en technisch gezien leesbare tekst geholpen kunnen worden. Tus-

sen haakjes: het is opvallend dat teksten die door gemiddelde leerling 'eind groep 4' met begrip gelezen kunnen worden, volgens de AVI-normering voor deze leerlingen zelfs gemiddeld genomen nog te grote decodeerproblemen bevatten.

We sluiten af met enkele concluderende opmerkingen over de CILT. Deze index lijkt in ieder geval bruikbaar voor een domeingerichte beschrijving van het niveau en het van de ontwikkeling van de leestehnik in het primair onderwijs. Didactische betekenis heeft de CILT naast de CLIB naar het zich laat aanzien vooral voor het onderwijs aan zeer zwakke decodeerders. Vermoedelijk kan de differentiatie van het leesonderwijs aan de andere leerlingen gebaseerd worden op de vaardigheid 'begrijpend lezen'. Een strikt uitgevoerde differentiatie op basis van de leestehnik van deze leerlingen zou er, gelet op de gegevens in Tabel 6, zelfs toe kunnen leiden dat ze teksten 'moeten' lezen die ze niet kunnen begrijpen en dat ze teksten niet 'mogen' lezen, terwijl ze die wel zouden kunnen begrijpen.

Literatuur

- Ballard, P.B. (1920). *Mental Tests*. Londen: University of London Press.
- Brus, B.T., & Voeten, M.J.M. (1973). *Een-Minuu-Test, vorm A en B*. Nijmegen: Berkhout.
- Hudson, P.T.W., Bergman, M.W., Houtmans, M.J.M., & Nas, G.L.J. (1984). De bestudering van woordherkenning als basis voor het lezen. In A.J.W.M. Thomassen, L.G.M. Noordman & P.A.T.M. Eling (Red.), *Het leesproces* (pp. 27 - 52). Lisse: Swets & Zeitlinger.
- Just, M. A., & Carpenter, P.A. (1987). *The psychology of reading and language comprehension*. New-ton: Allyn and Bacon.
- Kerlinger, F.N., & Pedhazur, E.J. (1973). *Multiple regression in behavioural research*. New York: Holt, Rinehart and Winston.
- Krom, R.S.H. (1996). *Toetsen Leestempo*. Arnhem: CITO.
- Lord, F.M., & Novick M.R. (1968). *Statistical theories of mental test scores*. New York: Addison Wesley.
- Perfetti, C.A. (1985). *Reading ability*. New York: Oxford University Press.

- Rayner, K., & Pollatsek, A. (1989). *Psychology of reading*. Englewood Cliffs: Prentice Hall.
- Scarborough, D.L., Cortese, C., & Scarborough, H.S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 1-17.
- Scarborough, D.L., Gerard, L., & Cortese, C. (1979). Accessing lexical memory: The transfer of word repetition effects across task and modality. *Memory & Cognition*, 7, 3-12.
- Staphorsius, G. (1992). *Clib-toetsen*. Arnhem: CITO.
- Staphorsius, G. (1994). *Leesbaarheid en leesvaardigheid. De ontwikkeling van een domeingericht meetinstrument*. (Proefschrift Universiteit Twente). Arnhem: Cito.
- Staphorsius, G. (1996). *Overzicht van tekstkenmerken van bijna 10.000 teksten voor de jeugd uit bijna 1500 verschillende bronnen*. (P335-doc 601). Arnhem: CITO.
- Staphorsius, G., & Krom, R.S.H. (1985). *Cito leesbaarheidsindex voor het basisonderwijs*. Arnhem: CITO.
- Staphorsius, G., Krom, R.S.H., & De Geus, K. (1988). *Frequenties van woordvormen en letterposities in jeugdlektuur*. Arnhem: CITO.
- Taylor, L. (1953). 'Cloze procedure'. A new tool for measuring readability. *Journalism Quarterly*, 30, 415-433.
- Thomassen, A.J.W.M., Noordman, L.G.M., & Eling, P.A.T.M. (1984). *Het leesproces*. Lisse: Swets & Zeitlinger.
- Thomassen, A.J.W.M., Noordman, L.G.M., & Eling, P.A.T.M. (1991). *Lezen en begrijpen*. Lisse: Swets & Zeitlinger.
- Verhoeven, L. (1995). *Drie Minuten Toets*. Arnhem: CITO.
- Visser, J., Laarhoven, A. van, & Beek, A. ter (1994). *AVI-Toetspakket*. Den Bosch: KPC.
- Vogel, M., & Washburne, C.W. (1928). An objective method of determining grade placement of childrens reading material. *Elementary School Journal*, 28, 373-381.
- Zwarts, M. (1990). *Balans van het taalonderwijs aan het einde van de basisschool*. Arnhem: CITO

Manuscript aanvaard 24 - 10 - 1996

164

PEDAGOGISCHE
STUDIËN

Auteurs

G. Staphorsius (wetenschappelijk medewerker, Cito Instituut voor Toetsontwikkeling)

N.D. Verhelst (wetenschappelijk medewerker, Cito Instituut voor Toetsontwikkeling; Hoogleraar Psychometrie, faculteit der Toegepaste Onderwijskunde, Universiteit Twente)

Adres: Cito Instituut voor Toetsontwikkeling
Postbus 1034
6801 MG Arnhem

Abstract

The Development of a Domain Referenced Index of the Decoding Load of Texts

G. Staphorsius & N.D. Verhelst. *Pedagogische Studiën*, 1997, 74, 154-164.

The CRIE (CITO Readability and Reading ability Index for Elementary education) is an index of both the reading ability of students in Dutch primary education and the readability of non-fiction prose. P-CRIE, a computer program, can indicate - in CRIE-values, ranging from 0 -100, - reading ability required by texts. The CRIE tests are an example of domain-referenced measurement of reading comprehension. The CRIE of a given reader refers to a set or domain of texts which can be comprehended by that individual with a given probability of success. The CRIE tests and P-CRIE can contribute to the individualization of reading instruction. The CRIE is primarily an index of the comprehensibility of texts and reading comprehension.

In addition to an index of readability as comprehensibility, an index of the decoding load of texts might be useful, especially for readers having a poor performance on decoding (aspect: speed), but whose comprehension of oral language is quite normal. Joining an index of the decoding skill required by texts to an index of the comprehensibility of texts, the matching of reading proficiency and readability can be optimized.

To develop such an index (of both the decoding skill and the 'decode-ability' of texts, we carried out a study, which will be reported on here. Our investigation has resulted in the CID (Cito Index of Decoding) a multiple regression equation with 'word length' and 'word frequency' as predictors ($R = .90$). The mean number of words decoded in one minute was selected as a measure of the decoding load of texts.

Korte-termijneffecten van coping-gedrag van docenten-in-opleiding

W.F. Admiraal, F.A.J. Korthagen en Th. Wubbels

Samenvatting

In dit artikel wordt het coping-gedrag van aanstaande docenten tijdens hun stage beschreven. Het betreft gedrag in de klas naar aanleiding van klassesituaties die problematisch zijn ervaren. Hierbij staan twee korte-termijneffecten van coping-gedrag centraal: de mate waarin taakgericht gedrag van leerlingen is toegenomen en de ervaren tevredenheid met de eigen reactie. Van 27 docenten-in-opleiding zijn ongeveer 300 situaties en reacties in kaart gebracht. De steekproef is representatief voor de hele groep van ongeveer 150 aanstaande docenten die in dezelfde periode de opleiding aan de Universiteit Utrecht hebben gevolgd.

Uit de analyses blijkt dat reacties van de docenten-in-opleiding uit ons onderzoek over het algemeen de taakgerichtheid van de klas niet beïnvloeden. Er zijn wel verschillen in de effecten van de reacties van de respondenten. Deze verschillen worden voor een deel verklaard door het coping-gedrag. In situaties waarin (de afwezigheid van) de taakgerichtheid van de klas als problematisch is ervaren, bestaat effectief coping-gedrag uit een (monotone) nadruk op het oproepen van spanning in de relatie met leerlingen en het variëren in intensiteit van de activiteiten. Voorbeelden van docentgedrag waarin spanning wordt opgeroepen, zijn ordemaatregelen, zoals leerlingen waarschuwen of de klas uit sturen, en ontevreden wachten op minder rumoer in de klas. Een voorbeeld van docentgedrag waarin de intensiteit van de activiteit wordt gevarieerd, is leerlingen bij de les betrekken door afwisselend gerichte vragen te stellen, antwoorden zwijsend af te wachten en antwoorden te corrigeren of te belonen. Eerder onderzoek maakt aannemelijk dat de gevonden effecten op korte termijn op gespannen voet kunnen staan met effecten op lange termijn: het beantwoorden van niet-taakgericht gedrag van leerlingen met ordemaatregelen is op korte ter-

mijn weliswaar effectief, maar kan contraproductief zijn op langere termijn. Wat betreft de ervaren tevredenheid zijn er geen significante verbanden aangetoond met het coping-gedrag van aanstaande docenten.

1 Inleiding

Veel beginnende docenten ervaren een discontinuïteit tussen de betrekkelijk veilige opleidingssituatie en de realiteit van het zelfstandig lesgeven in een eerste baan. In een overzichtsartikel geeft Veenman (1984) aan dat beginnende docenten gevraagd naar hun problemen, vooral problemen noemen die te maken hebben met het functioneren in de klas en het houden van orde. Uit onderzoek naar de ontwikkeling in 'concerns' van docenten (bijvoorbeeld Fuller & Bown, 1975) komt naar voren dat beginnende docenten aanvankelijk meer gericht zijn op hun functioneren in de klas dan op meso- of macro-concerns. Recent onderzoek van MacDonald (1993) wijst uit dat iets dergelijks ook geldt voor docenten-in-opleiding. In de lerarenopleiding wordt veel aandacht besteed aan de wijze waarop aanstaande docenten het beste om kunnen gaan met problemen in de onderwijspraktijk, of meer specifiek, in de klas.

In dit artikel zijn we gericht op het vergroten van het inzicht in effecten op korte termijn van de wijze waarop docenten-in-opleiding omgaan met problematische gebeurtenissen in de klas. Een dergelijk inzicht kan op termijn leiden tot aanwijzingen voor de opleiding van docenten. De wijze waarop docenten omgaan met problematische gebeurtenissen zal worden aangeduid als het *coping-gedrag* van docenten. In de volgende paragraaf zal het (theoretische) kader van ons onderzoek worden toegelicht. Tevens zal voor twee kernbegrippen uit het onderzoekskader (respectievelijk 'stress' en 'coping') een deel van het onderzoek worden

besproken dat is gedaan naar problematische gebeurtenissen van docenten en de wijze waarop zij hiermee omgaan.

2 Stress en coping

In ons onderzoek richten we ons op problematische klassesituaties van docenten-in-opleiding. Onderzoek van Bergen en Van Opdorp (1989) maakt aannemelijk dat dergelijke situaties als incidenten worden ervaren. In de transactionele theorie over stress en coping wordt veel belang gehecht aan dergelijke dagelijkse beslommingen. Lazarus en Folkman (1987), die veel over onderzoek naar stress en coping vanuit een transactioneel perspectief hebben gerapporteerd, geven aan dat dagelijkse beslommingen over het algemeen veranderingen in gezondheid beter voorspellen dan meer ingrijpende gebeurtenissen.

Lazarus en Folkman (1984) vatten factoren in onderzoek naar stress en coping samen in een model. De voor stress en coping belangrijke aspecten zijn ingedeeld in drie onderzoeksniveaus en vier groepen factoren. Wij hebben dit model gebruikt om stress bij en coping door docenten weer te geven (zie Figuur 1). De drie onderzoeksniveaus zijn het maatschappelijke, het psychologische en het fysiologische niveau. De vier groepen van factoren zijn de voorafgaande condities, de mediërende processen, de korte-termijneffecten en de lange-termijneffecten. Een uitgangspunt van het model is dat elke stressvolle situatie verloopt van voorafgaande condities via mediërende processen naar effecten op korte en lange termijn. Mede omdat we in ons onderzoek informatie willen verzamelen over de onderwijspraktijk van docenten-in-opleiding, zijn wij vooral geïnteresseerd in factoren op het psychologische onderzoeksniveau. In dit artikel concentreren wij ons op problematische klassesituaties (als één van de voorafgaande condities uit Figuur 1), coping-gedrag (als één van de mediërende processen) en effecten op korte termijn. Hierna zullen deze drie aspecten worden toegelicht.

2.1 Problematische gebeurtenissen

Stress wordt in de literatuur beschreven in termen van zowel een stimulus als een respons. In

de responsgeoriënteerde benadering (bijvoorbeeld Selye, 1979) wordt een reactie alleen een stressrespons genoemd als deze volgt op een aanwijsbare stressor, terwijl in de stimulusgeoriënteerde benadering (zie bijvoorbeeld Thoits, 1983) een stimulus slechts als stressor wordt beschouwd als het een identificeerbare stressrespons oproept. Bijvoorbeeld, een fysiologische respons als verhoging van de bloeddruk is alleen een stressrespons als de verhoging plaatsvindt naar aanleiding van een stressor (bijvoorbeeld een ingrijpende gebeurtenis), terwijl het afleggen van een rij-examen alleen een stressor is als de examenkandidaat een stressrespons vertoont (bijvoorbeeld een verhoogde bloeddruk).

Beide benaderingen zijn ons inziens onvolledig. In de transactionele benadering van stress en coping worden de stimulus- en de responsgeoriënteerde benadering gecombineerd en nader uitgewerkt (zie Lazarus & Folkman, 1984). Stress wordt opgevat als een relatie tussen kenmerken van de omgeving en van een persoon. De discrepantie tussen de ervaren eisen vanuit de omgeving en de waargenomen mogelijkheden om hiermee om te gaan, bepaalt of een persoon stress ervaart. Deze discrepantie leidt tot stress als de eisen van de omgeving als te hoog worden ervaren, gezien de mogelijkheden die de betreffende persoon bij zichzelf waarneemt (Lazarus & Folkman, 1987). Door docenten als problematisch ervaren gebeurtenissen in de klas kunnen worden opgevat als een vorm van een dergelijke discrepantie.

Uit het eerder genoemde onderzoek van Veenman (1984) komt, zoals gezegd, naar voren dat veel problematische situaties die beginnende docenten ervaren, te maken hebben met de interactie met leerlingen, zoals ordeproblemen, het motiveren van leerlingen en het omgaan met verschillen tussen individuele leerlingen. Ook Peters (1985) heeft empirisch onderzoek gedaan naar problematische situaties van docenten. In zijn onderzoek is aan docenten in het voortgezet onderwijs gevraagd situaties te formuleren die zij in hun beroepsuitoefening als problematisch ervaren. Door drie codeurs zijn bijna 2.700 omschrijvingen van situaties van ruim 250 docenten ingedeeld in 13 categorieën. Uit de omschrijvingen is vervolgens een representatieve steekproef getrokken van 37 situaties. Daarna is aan ongeveer

	voortgaande condities	mediërende processen	effecten op korte termijn	effecten op lange termijn
maatschappelijk niveau	<ul style="list-style-type: none"> * onderwijsbeleid school & overheid * rol, status en salaris beroep leraar * onderwijskansen leerlingen * reorganisaties 	<ul style="list-style-type: none"> * politieke steun * vakbondsteun * steun van beroepsgroep * lerarenopleiding 	<ul style="list-style-type: none"> * financiële tegemoetkoming * bijscholing * andere beleidsveranderingen 	<ul style="list-style-type: none"> * statusverandering * beroepsrolverandering
	Docent:			
	<ul style="list-style-type: none"> * gedragsstijl (coping-stijl, interactie-stijl, onderwijsstijl) * zelfbeeld * (onderwijs)opvattingen * achtergrondkenmerken 	<ul style="list-style-type: none"> * inschatting omgeving * inschatting eigen mogelijkheden * herinschatting omgeving * gedrag in klas, school en daarbuiten * kiezen voor andere werkgever partner, therapeut 	<ul style="list-style-type: none"> * gevoelens docent * veranderingen in interacties * veranderingen in schoolomgeving * kiezen voor bijscholing * kiezen voor andere werkgever 	<ul style="list-style-type: none"> * beroepstevredenheid * functioneren als docent * uittreden uit het beroep
psychologisch niveau	<ul style="list-style-type: none"> Klas- en schoolomgeving * patronen in interactie met leerlingen, vragers van steun aan collega's, gen, collega's en ouders * schoolomgeving en -voorzieningen * taakhoud docent * onderwijsprogramma 			
fysiologisch niveau	<ul style="list-style-type: none"> * fysieke toestand docent * risicogedrag docent (roken, alcohol, drugs) * taakbelasting * fysieke kenmerken werkomgeving (bijv. temperatuur) 	<ul style="list-style-type: none"> * verhoging bloeddruk * verhoging cortisolgehalte * blozen * zenuwtrekjes 	<ul style="list-style-type: none"> * psychosomatische klachten (vermoedheid, hoofdpijn) * ziekte * absentie 	<ul style="list-style-type: none"> * burnout-verschijnselen * chronische ziekte

Figuur 1. Voorbeelden van factoren die van belang kunnen zijn voor onderzoek bij docenten naar stress en coping. De factoren zijn gegroepeerd per onderzoeksniveau en per groep van factoren. In de meest linkse kolom is het onderzoeksniveau aangegeven, in de bovenste regel zijn de groepen van factoren aangeduid

180 docenten gevraagd de formuleringen van de 37 situaties te ordenen. Situaties waarvan de docenten vonden dat ze op elkaar leken, moesten zij bij elkaar plaatsen. Na hiërarchische clusteranalyse zijn vijf groepen als problematisch ervaren situaties onderscheiden: functioneren van de docent tijdens de les, ongewenst gedrag van leerlingen tijdens de les, hulpmiddelen en materialen, invloed van de organisatie op het functioneren, en functioneren van en samenwerking met collega's. De eerste twee van de bovengenoemde typen situaties hebben betrekking op interacties in de klas tussen de docent en leerlingen. Voorbeelden van problematische situaties, die zijn gegroepeerd onder 'functioneren van de docent tijdens de les' zijn het geven van extra uitleg en weinig aandacht kunnen geven aan individuele leerlingen. 'Ongewenst gedrag van leerlingen tijdens de les' houdt bijvoorbeeld in dat leerlingen hun spullen zijn vergeten, huiswerk niet hebben gemaakt, naar het toilet gaan, door de uitleg van een docent heen kletsen en het meubilair bekladden. De andere drie typen situaties verwijzen volgens Peters naar de organisatie van het onderwijs. Onder het type 'hulpmiddelen en materialen' worden situaties verstaan als het ontbreken van een borstel, storing via de intercom en een diaprojector die niet werkt. Situaties die verwijzen naar 'invloed van de overheid en de schoolorganisatie op het functioneren' zijn bijvoorbeeld maatregelen van de overheid, roosterwijzigingen en onverwachte sectiebijeenkomsten. De vijfde groep situaties, 'functioneren van en samenwerking met collega's', verwijst naar situaties waarin collega's geen interesse tonen voor het vak of problemen van de docent, een collega lessen heeft overgenomen op een manier die de docent niet aanspreekt of er lang wordt vergaderd zonder resultaat. Omdat wij onderzoek doen bij docenten-in-opleiding en uit ander onderzoek (zie bijvoorbeeld MacDonald, 1993) bekend is dat hun 'concerns' voornamelijk betrekking hebben op het microniveau, zullen wij ons concentreren op de eerste drie groepen situaties die Peters noemt.

2.2 Coping-gedrag van aanstaande docenten

Coping-gedrag wordt in de transactionele benadering van stress en coping omschreven als het constant veranderende geheel van acties

en cognities waarmee wordt geprobeerd specifieke eisen van de omgeving en/of van de persoon die de waargenomen mogelijkheden van een persoon te boven gaan, te reduceren, tolereren, accepteren of weg te nemen (Lazarus & Folkman, 1984). Hierbij kan coping-gedrag vooral gericht zijn op de beleving van een situatie of een gebeurtenis (emotieregulerend coping-gedrag) of op het veranderen van een gebeurtenis (probleemgericht coping-gedrag). Het eerste type gedrag verwijst bijvoorbeeld naar het zoeken van minder belastende verklaringen voor het ontstaan van een situatie, het denken aan andere dingen dan aan de betreffende gebeurtenis of zich uitleven in hobby's. Probleemgericht coping-gedrag bestaat onder meer uit het nadenken over oplossingen en het ingrijpen in de situatie. Over het algemeen zou emotieregulerend coping-gedrag vooral optreden in situaties waarin het gevoel is ontstaan niets te kunnen veranderen, terwijl probleemgericht coping-gedrag met name zou worden ingezet wanneer de indruk bestaat dat een situatie gewijzigd kan worden. Lazarus en Folkman merken ook op dat veel coping-gedrag zowel probleemgericht als emotieregulerend kan zijn. Tevens maken zij aannemelijk dat probleemgericht coping-gedrag meer afhankelijk is van de context waarin het gedrag plaatsvindt en dat emotieregulerend coping-gedrag vooral wordt beïnvloed door relatief stabiele persoonlijkheidskenmerken.

Reacties van docenten op problematische situaties zijn onderzocht door het coping-gedrag van docenten binnen en buiten het klaslokaal in kaart te brengen (zie bijvoorbeeld Dewe, 1985; Freeman, 1987; Rudow, 1990; Wahl, 1991; Wood & Dorsey, 1989; Wulk, 1988). In het kader van het onderhavige onderzoek is met name het onderzoek van belang waarin coping-gedrag *in de klas* is bestudeerd. In onderzoek van Den Hertog (1990) bijvoorbeeld, zijn onder meer gedragsintenties ontlokt aan docenten en docenten-in-opleiding middels beschrijvingen van situaties die over het algemeen door docenten als belastend worden ervaren. Aan ruim 800 docenten in het voortgezet onderwijs en ongeveer 140 docenten-in-opleiding zijn beschrijvingen van zes situaties voorgelegd. Vier van de voorgelegde situatiebeschrijvingen hebben betrekking op gebeurtenissen in de klas en twee verwijzen naar inter-

acties met collega's. De docenten is door middel van een vragenlijst gevraagd aan te geven in hoeverre een aantal mogelijke reacties voor hen van toepassing zijn, als zij de docent uit de beschrijving zouden zijn. Er worden vijf typen gedragsintenties onderscheiden die respectievelijk verwijzen naar dominant-onvriendelijk, dominant-vriendelijk, meegaand, organiserend, en vermijgend gedrag. Deze indeling in typen coping-gedrag is gebaseerd op literatuur over stress en coping (Folkman & Lazarus, 1985) en over interpersoonlijk gedrag van docenten in de klas (Créton & Wubbels, 1984). Voorbeelden van reactie-intenties naar aanleiding van de situatie waarin een leerling niet oplet en met haar buurvrouw praat, zijn 'ik verwijder haar uit de les' (dominant-onvriendelijk), 'ik waarschuw haar vriendelijk, maar beslist, dat ze nu verder stil moet zijn' (dominant-vriendelijk), 'ik laat haar blijken dat ik wel begrip heb voor haar gedrag' (meegaand), 'ik neem me voor een gesprek te regelen met het meisje en de klasseleraar' (organiserend) en 'ik reageer er niet op en ga gewoon door met de les' (vermijgend). Ook in het onderhavige onderzoek is het coping-gedrag van docenten bestudeerd vanuit meerdere perspectieven op docentgedrag.

2.3 Effecten op korte-termijn

Coping-gedrag is niet gedefinieerd in termen van effecten, ofschoon verwacht kan worden dat bepaald coping-gedrag bepaalde effecten sorteert. Zo kan bepaald coping-gedrag (bijvoorbeeld ingrijpen in een wanordelijke situatie in de klas) gericht zijn op het oplossen van een probleem, maar het behoeft hierin niet te resulteren. Lazarus en Folkman (1984) maken bij effecten van coping-gedrag onderscheid in drie componenten: het functioneren op het werk en in het sociale leven (de gedragscomponent), de ervaren tevredenheid met het bestaan (de affectieve component) en de lichamelijke gezondheid (de fysiologische component). Hierbij maken zij een onderscheid in effecten op korte termijn en op lange termijn. Effecten op korte termijn kunnen een verandering van klas, school of baan (gedragscomponent), een slecht humeur (affectieve component) of een verhoging van de hartslag (fysiologische component) inhouden. Effecten op lange termijn zijn bijvoorbeeld uittreding uit het beroep

(gedragscomponent), depressieve stemming (affectieve component) of psychosomatische klachten (fysiologische component). Korte-termijneffecten zijn dikwijls gerelateerd aan een specifieke context waarin de problematische situatie optreedt en het coping-gedrag wordt ondernomen. Lange-termijneffecten daarentegen hebben nauwelijks nog betrekking op de specifieke situatie waarop is gereageerd, en zijn dikwijls ook moeilijker te veranderen.

In onderzoek naar docentgedrag en prestaties van leerlingen (zie voor een overzicht Brophy & Good, 1986; Doyle, 1986) wordt de tijd die leerlingen aan de leertaak besteden (de mate van taakgerichtheid van leerlingen) als indicatie beschouwd van effectief lesgeven. Brophy en Good maken in hun overzicht aanneemelijk dat docenten prestaties van leerlingen positief beïnvloeden, wanneer zij erin slagen leerlingen zoveel mogelijk aan te zetten tot (academische) leeractiviteiten en de tijd die aan andere taken wordt besteed, te minimaliseren. Dergelijk gedrag van docenten komt bijvoorbeeld naar voren in de wijze waarop zij de lesstof presenteren, de klas organiseren, orde-maatregelen nemen en pauzes houden. Derhalve is taakgerichtheid van de klas in ons onderzoek gekozen als één van de vast te stellen korte-termijneffecten.

3 Opzet van het onderzoek

3.1 Probleemstelling en onderzoeksvragen

Stress en coping moeten als onderdeel van een proces worden opgevat. Niet alleen de factoren die zijn weergegeven in Figuur 1 maken onderdeel uit van dit proces, ook binnen de factoren, bijvoorbeeld coping-gedrag, kunnen verschillende fasen worden onderscheiden. In ons onderzoek zou bepaald coping-gedrag van een docent-in-opleiding in een wanordelijke situatie een reactie kunnen zijn in een lange reeks van reacties op wanorde (in die betreffende klas), maar ook een reactie op een voor de docent nieuwe situatie. Toch hebben wij in dit onderzoek effecten van coping-gedrag onderzocht door telkens één combinatie van situatie, gedrag en effecten uit een wellicht lange reeks van combinaties te bestuderen. Wij zijn van mening dat dit een noodzakelijke stap is in de richting van onderzoek naar het meer volledig

Tabel 1
Variabelen in de analyses ter beantwoording van de twee onderzoeksvragen

onderzoeksvraag	onafhankelijke variabelen	covariaten	afhankelijke variabelen
1	- coping-gedrag - aard van de situatie - interactie tussen beiden	n.v.t.	- taakgerichtheid klas
2	- coping-gedrag	- eerste inschatting - tweede inschatting	- ervaren tevredenheid

in kaart brengen van coping-processen bij docenten-in-opleiding.

Met het onderhavige onderzoek richten we ons op effecten op korte termijn van de wijze waarop docenten-in-opleiding omgaan met problematische gebeurtenissen in de klas. We hebben gekozen voor twee korte-termijneffecten. Het eerste effect betreft de mate waarin docenten-in-opleiding de taakgerichtheid van de klas veranderen richting meer taakgericht gedrag. Zoals eerder aangegeven, is in evaluatie van docentgedrag in de klas de mate waarin leerlingen taakgericht aan het werk zijn een veel gekozen criteriumvariabele (zie Brophy & Good, 1986; Doyle, 1986; Evertson, 1989). Met deze criteriumvariabele sluiten we aan bij wat Lazarus en Folkman (1984) de gedragscomponent in effecten noemen. Het tweede effect betreft de mate waarin een docent-in-opleiding na een reactie tevreden of ontevreden is met het eigen functioneren. Hiermee richten we ons op wat Lazarus en Folkman de affectieve component van effecten noemen. De onderzoeksvragen zijn als volgt geformuleerd:

- 1 *Wat is de relatie van coping-gedrag van aanstaande docenten met de verandering in taakgerichtheid van de klas?*
- 2 *Wat is de relatie van coping-gedrag van aanstaande docenten met de ervaren tevredenheid van docenten met hun reactie?*

Zoals eerder aangegeven maken Lazarus en Folkman (1984) aannemelijk dat probleemgerichte coping en de effecten ervan afhankelijk zijn van de context waarin het gedrag optreedt, en dat emotieregulerend coping-gedrag meer afhankelijk is van relatief stabiele persoonlijkheidskenmerken. Wat betreft de taakgerichtheid van de klas is het aannemelijk dat coping-gedrag in situaties waarin de taakgerichtheid als problematisch is ervaren andere effecten sorteert dan in andere problematische klas-situaties. Derhalve zullen we in de analyses ter

beantwoording van de eerste onderzoeksvraag ook informatie over de aard van de problematische situatie opnemen. We verwachten dan significante verbanden tussen coping-gedrag en de taakgerichtheid van de klas in situaties waarin de taakgerichtheid als problematisch is ervaren, maar geen significante verbanden in andersoortige situaties. In de analyses ter beantwoording van de tweede onderzoeksvraag zullen we gegevens over de inschatting van de betekenis van een situatie opnemen. De reden hiervan is dat het aannemelijk is dat er een relatief stabiele relatie bestaat tussen de inschatting van een situatie en de ervaren tevredenheid met de eigen reactie op een situatie. Bijvoorbeeld, in een als zeer problematisch ervaren situatie lijkt het onwaarschijnlijk dat een docent tevreden is, nog ongeacht het gehanteerde coping-gedrag. In Tabel 1 zijn de variabelen in de analyses samengevat.

3.2 Steekproef

Het onderzoek is uitgevoerd bij 27 docenten-in-opleiding (DIO's) van ongeveer 150 DIO's die in 1990 en 1991 de universitaire lerarenopleiding van de Universiteit Utrecht volgden. Onze steekproef bleek niet te verschillen (bij $\alpha=0,05$) van de andere docenten-in-opleiding wat betreft coping-stijl¹, sekse, leeftijd en vakgebied.

De Utrechtse universitaire lerarenopleiding is een eenjarige tweede-fase-opleiding, waarin naast andere curriculumonderdelen twee stages worden doorlopen: de triostage en de individuele eindstage. Zeven studenten zijn tijdens hun triostage in het onderzoek betrokken. De triostage is een stage in de eerste helft van de opleiding, waarbij een DIO samen met twee andere DIO's een aantal lessen van een ervaren docent observeert en deze lessen vervolgens gedeeltelijk overneemt. Twintig andere respondenten hebben aan het onderzoek deelgenomen

tijdens hun individuele eindstage. Deze stage, in de tweede helft van de opleiding, houdt in dat één DIO gedurende drie maanden alle lessen van een paar klassen overneemt van een ervaren docent. In deze stage is de DIO verantwoordelijk voor de lessen, inclusief bijvoorbeeld rapportcijfers en contacten met ouders en de schoolleiding. Koetsier (1991) geeft een uitgebreide beschrijving van de Utrechtse universitaire lerarenopleiding, in het bijzonder van de individuele eindstage.

3.3 Operationalisaties van de variabelen

Hieronder worden de operationalisaties van de variabelen beschreven. Ook zijn aspecten van betrouwbaarheid, begripsvaliditeit, gelijktijdige validiteit en interne validiteit onderzocht. Voor informatie over de betrouwbaarheid en de validiteit verwijzen we naar Admiraal (1994). Daar is uiteengezet dat de resultaten van deze analyses bevredigend zijn. De betrouwbaarheid van de instrumenten (i.c. de interbeoordelaars-overeenstemming in de vorm van Cohens κ of de 'pooled within groups' correlatiecoëfficiënt), ligt in alle gevallen rond of ruim boven de door ons gehanteerde norm van 0,80 (zie Popping, 1988).

In ons onderzoek zijn op twee manieren gegevens verzameld: door middel van een video-opname van een les van de respondent en door middel van een interview volgens de 'stimulated recall'-methode na de les. Van elke respondent is één les op video opgenomen. De video-opnamen zijn gebruikt om informatie te verzamelen over het coping-gedrag van aanstaande docenten en de taakgerichtheid van leerlingen (en als stimulus in de interviews na de les); gegevens uit de interviews verwijzen naar informatie over de aard van problematische situaties, de inschatting van de betekenis van een situatie en de ervaren tevredenheid met de eigen reactie. De precieze gegevensverzameling en de wijze van coderen zullen later worden besproken.

Over het algemeen hanteert de respondent in de geobserveerde les als didactische werkvorm klassikale instructie of onderwijsleergesprek. Practicumlessen zijn voor dit onderzoek niet geobserveerd.

Aard van de problematische situatie

We hebben de aard van de classesituatie gede-

finieerd in termen van de door de respondent waargenomen kenmerken van de situatie in de klas. De situaties zijn gecodeerd aan de hand van negen categorieën. Deze indeling is gebaseerd op het reeds beschreven onderzoek van Peters (1985) naar problematische situaties van docenten en situatiebeschrijvingen van docenten uit onderzoek van Wulk (1988) en onze voorstudie waarin wij een aantal lessen van aanstaande docenten hebben geobserveerd en hen hebben geïnterviewd. Voor het onderhavige onderzoek is het van belang onderscheid te maken tussen situaties waarin de taakgerichtheid van de klas als problematisch is ervaren (43 procent van de 306 situaties) en in andersoortige situaties, zoals problemen met de lesstofoverdracht, de organisatie van de les of vragen en opmerkingen van leerlingen over de lesstof.

Inschatting van de betekenis van een situatie

In ons onderzoek is in overeenstemming met de transactionele theorie van stress en coping (zie Lazarus & Folkman, 1984) de inschatting van de betekenis van een situatie onderverdeeld in twee variabelen: de 'eerste inschatting' en de 'tweede inschatting'. De eerste inschatting verwijst naar een beoordeling door de respondent van de gevolgen die een situatie in de klas heeft voor het verloop van de les of voor het eigen welbevinden. De tweede inschatting betreft de inschatting van de mogelijkheden om met de situatie om te gaan. De onderzoeksinstrumenten met betrekking tot deze variabelen bestaan uit 11 beoordelingschalen (zes voor de eerste inschatting en vijf voor de tweede inschatting), die zijn gebaseerd op onderzoek over inschatting van levensgebeurtenissen van Smith en Ellsworth (1985) en Mauro, Sato en Tucker (1992). Voorbeelden van de eerste inschatting zijn de mate waarin een situatie als negatief dan wel positief wordt ervaren, de mate waarin een situatie opwinding oproept, en de mate waarin verantwoordelijkheid wordt genomen voor het ontstaan van een situatie. Voorbeelden van de tweede inschatting zijn de mate waarin actief wordt ingegrepen, de mate waarin verantwoordelijkheid wordt genomen voor de eigen reactie, en de mate waarin wordt ingeschat dat een reactie effectief is. De in totaal 11 beoordelingschalen voor de inschatting van de betekenis van

een situatie zijn vijfpuntsschalen van het Likert-type, waarbij een schaal meer van toepassing is naarmate de score hoger is. Omdat in dit artikel de inschatting van de betekenis van een problematische situatie als covariaat in de analyse van de relatie tussen coping-gedrag en de ervaren tevredenheid wordt gezien, volstaan we voor een beschrijving van de beoordelingsschalen met een verwijzing naar Admiraal (1994).

Coping-gedrag

In ons onderzoek beperken we coping-gedrag van aanstaande docenten tot de waargenomen activiteiten van docenten tijdens of direct volgend op een problematische situatie in de klas. We karakteriseren coping-gedrag van aanstaande docenten vanuit drie perspectieven. Deze perspectieven zijn gebaseerd op onderzoek naar (opties van) coping-gedrag van werknemers in de commerciële dienstverlening door Farell (1983), methodisch handelen van docenten (Westerhof, 1987) en interpersoonlijk gedrag van docenten (Créton & Wubbels, 1984). De drie perspectieven zijn weergegeven in beoordelingsschalen, die uitgebreid zijn beschreven in Admiraal, Wubbels en Korthagen (1996a; 1996b). Het coping perspectief van gedrag verwijst naar de betekenis die, gezien de aard van de situatie, aan gedrag kan worden toegekend. Het methodisch handelen van docenten heeft betrekking op docentgedrag in termen van lesorganisatie en didactisch handelen. Het interpersoonlijke perspectief van docentgedrag verwijst naar de relationele boodschap die door een docent in gedrag ten opzichte van leerlingen tot uiting wordt gebracht.

De drie perspectieven van coping-gedrag leveren theoretisch onderscheidbare aspecten op. Aangezien wij het coping-gedrag als één concept beschouwen, hebben op de scores van de vier beoordelingsschalen een principale componenten-analyse uitgevoerd voor elk type problematische situaties (PRINCALS; Gifi, 1983) om onderliggende dimensies van de verschillende aspecten coping-gedrag te extraheeren. Twee dimensies van coping-gedrag (waarbij de scores variëren van ongeveer -3 tot +3) blijken te kunnen worden onderscheiden, die in elk type situatie te zamen tenminste 50 procent van de variantie in scores verklaren.

De eerste dimensie van docentgedrag die is geëxtraheerd, hebben we de 'activiteitsdimensie' (actief versus passief) genoemd: de intensiteit waarmee een docent in klassesituaties de inhoudelijke informatie van het docentgedrag tracht over te brengen door middel van een actieve interactie met leerlingen. *Lage scores* op de activiteitsdimensie worden in alle typen problematische situaties gevonden, wanneer aanstaande docenten zwigend afwachten, naar leerlingen luisteren, iets opzoeken in een boek of het bord uitvegen. *Hoge scores* op de activiteitsdimensie verwijzen naar gedrag van aanstaande docenten dat is gericht op het betrekken van leerlingen bij de les (vragen stellen, 'beurten' geven of antwoorden belonen of juist afwijzen) en op het handhaven van de orde in de klas.

De tweede dimensie van docentgedrag die is geëxtraheerd, hebben we de 'spanningsdimensie' (spanning versus ontspanning) genoemd: de mate waarin er door het gedrag van docenten spanning in de relatie met leerlingen wordt opgeroepen. *Lage scores* op de spanningsdimensie worden gevonden voor instructieactiviteiten, zoals lesstofoverdracht en helpen van leerlingen, en organisatorische handelingen. *Hoge scores* op de spanningsdimensie worden gevonden bij ordemaatregelen en (ontevreden) zwijgen door de aanstaande docenten. Dergelijk gedrag toont een afwijzende houding ten aanzien van de leerlingen en interpreteren wij als veel spanning oproepend.

Effecten op korte termijn

In ons onderzoek zijn twee korte-termijneffecten opgenomen: het verschil tussen de taakgerichtheid van de klas voor en na het optreden van een problematische situatie (de verandering in taakgerichtheid), en de ervaren tevredenheid met het eigen functioneren. De operationalisatie van taakgerichtheid van de klas is gebaseerd op de 'Classroom Activity Record' van Evertson en Burry (1988) en bestaat uit een vijfpuntsbeoordelingsschaal waarmee wordt gescoord in welke de mate de klas taakgericht gedrag vertoont. De schaalpunten zijn samengesteld uit informatie over de intensiteit en over de mate van variatie tussen leerlingen in de taakgerichtheid van het gedrag, het aantal leerlingen dat taakgericht gedrag vertoont en het aantal seconden dat leerlingen taakgericht

gedrag vertonen. Wij hebben deze informatie gecombineerd in vijf schaalpunten, die variëren van 'helemaal niet-taakgericht' (1) tot 'helemaal taakgericht' (5). Willett (1989) heeft aannemelijk gemaakt dat bij het meten van verandering met twee meetpunten het gebruik van verschillcores tussen voor- en nameting de voorkeur heeft boven het als covariaat in de analyses opnemen van de scores op de voormeting. We hebben er derhalve voor gekozen om als effect op korte termijn *het verschil* tussen de taakgerichtheid van de klas voor en na een reactie van een docent in de analyses op te nemen. Een negatieve score op deze dimensie betekent dat de klas na een reactie is veranderd in de richting van meer taakgericht gedrag; een positieve score betekent dat de klas is veranderd in de richting van minder taakgericht gedrag. Een score rond 0 wil zeggen dat er geen verandering in de taakgerichtheid heeft plaatsgevonden.

De variabele 'ervaren tevredenheid' heeft betrekking op een al dan niet positieve waardering van de respondent van zijn eigen gedrag. Uitspraken waarin de ervaren tevredenheid naar voren komt, kunnen verwijzen naar een waardering achteraf van de reacties (bijvoorbeeld 'dit was niet zo handig gedaan van mij'), waardering van gevolgen van de eigen reactie (bijvoorbeeld 'hierdoor wilde opeens de hele klas hun proefwerkcijfer weten') of alternatieve reacties (bijvoorbeeld 'ik had beter direct kunnen beginnen met de les'). Elke relevante uitspraak van een respondent is door een codeur gescoord op een vijfpuntsbeoordelingschaal van 1 (zeer ontevreden) tot 5 (zeer tevreden). Er is een 'ongeldige' score toegekend als er geen uitspraken konden worden geselecteerd of als ervaringen onduidelijk zijn geformuleerd.

3.4 Gegevensverzameling en codeereenheden

In ons onderzoek verwijst coping-gedrag van een aanstaande docent in een problematische situatie naar waarneembare activiteiten tijdens een fragment van de opgenomen les. Het gedrag in elke situatie vormt een reactie die een lengte heeft tussen ongeveer tien seconden en vijf minuten. Het begin en het eind van elke reactie op video zijn door de respondent tijdens het interview zelf bepaald. Om het gedrag in de reacties in kaart te brengen is gebruik gemaakt

van 'event sampling', dat wil zeggen dat elke verandering in gedrag van een aanstaande docent (beschreven in termen van de instrumenten uit §3.3) een nieuwe codeereenheid bepaalt. Aan elke eenheid zijn meerdere scores toegekend op basis van de drie perspectieven op docentgedrag. Na de in de vorige paragraaf beschreven principale componenten-analyses wordt het coping-gedrag in elke reactie gekenmerkt door één score of een reeks van scores op de activiteits- en de spanningsdimensie. Voor iedere dimensie afzonderlijk kan coping-gedrag worden weergegeven aan de hand van een profiel van scores op de betreffende dimensies. Voor elke reactie hebben wij dit profiel van scores op een dimensie samengevat aan de hand van de gemiddelde score en de spreiding in scores op de activiteits- en de spanningsdimensie².

De taakgerichtheid van de klas is voor elke reactie op twee momenten in kaart gebracht: gedurende een fragment van een minuut (zes eenheden van 10 seconden) direct voorafgaand aan het begin van een reactie en gedurende een fragment van een minuut (zes eenheden van 10 seconden) direct na afloop van een reactie. Het resultaat is twee scores met betrekking tot taakgerichtheid van de leerlingen (vooraf en achteraf) die staan voor het gemiddelde van toegekende scores direct voor en direct na een reactie. Zoals eerder beschreven, zijn deze twee scores samengevat in één verschillscore.

In de interviews hebben we de respondenten gevraagd ongeveer tien situaties uit de voorgaande les te noemen die hun aandacht hebben geëist en/of energie hebben gevegd³. Ook hebben we hun gevraagd aan te geven wat er tijdens de gebeurtenis door hen heen ging (hun inschatting van de betekenis van een situatie), en wat zij achteraf vonden van hun reactie (hun ervaren tevredenheid met hun reactie). Aan de omschrijving van een respondent van een problematische situatie is één score toegekend. Alle uitspraken die betrekking hebben op de eerste inschatting, de tweede inschatting en de ervaren tevredenheid met de eigen reactie zijn geselecteerd en gecodeerd door middel van de eerder genoemde beoordelingsschalen. Voor elke situatie (en reactie) zijn de scores van alle uitspraken op de verschillende beoordelingsschalen samengevat in een gemiddelde score. In totaal zijn 306 klassesituaties en hetzelfde

Tabel 2

Variabelen in de analyse van de relatie tussen coping-gedrag en taakgerichtheid van leerlingen

variabele		beschrijving
coping-gedrag	GEM-ACT	gemiddelde score per reactie op de activiteitsdimensie (gemiddelde = 0,06; standaardafwijking = 0,59)
	GEM-SPA	gemiddelde score per reactie op de spanningsdimensie (gemiddelde = -0,04; standaardafwijking = 0,64)
	SD-ACT	standaardafwijking in scores per reactie op de activiteitsdimensie (gemiddelde 0,72; standaardafwijking = 0,50)
	SD-SPA	standaardafwijking in scores per reactie op de spanningsdimensie (gemiddelde 0,79; standaardafwijking = 0,55)
aard problematische situatie	SIT	reacties op situaties waarbij de taakgerichtheid van de klas als problematisch is ervaren (1; N=79) of op andere situaties (0; N=108)
verandering in taakgerichtheid	VERSCHIL	verschil tussen taakgerichtheid vooraf en na een reactie van een docent (gemiddelde = -0,14; standaardafwijking = 1,06)

aantal reacties geanalyseerd. Voor informatie over de wijze waarop de kwaliteit van het bepalen van de codeereenheden is nagegaan, verwijzen we naar Admiraal (1994). De betrouwbaarheid, in de vorm van Cohens κ of percentage overeenkomst, ligt boven de in ons onderzoek gehanteerde norm van 0,80.

3.5 Analyses

We hebben gegevens verzameld over docenten en bij elke docent over reacties op een aantal situaties. Dit betekent dat onze gegevens een hiërarchische structuur hebben. We hebben daarom bij het beantwoorden van de onderzoeksvragen gebruik gemaakt van meerniveau-analyses (Mln; Woodhouse, Rasbash, Goldstein, Yang & Plewis, 1996). Teneinde na te gaan of meerniveau-analyses noodzakelijk zijn, is eerst nagegaan hoeveel variantie er binnen (op reactieniveau) en tussen docenten (op docent- of lesniveau) bestaat in de scores op verandering in taakgerichtheid en ervaren tevredenheid. Deze gegevens vormen de basis voor de intraklasrelatie (ρ), de proportie variantie tussen docenten in scores op de kortetermijneffecten. Wanneer de variantie tussen docenten niet significant afwijkt van 0, achten wij meerniveau-analyses niet noodzakelijk. Vervolgens worden (meerniveau)regressie-analyses uitgevoerd om beide onderzoeksvragen te beantwoorden⁴.

4 Resultaten

4.1 Coping-gedrag en verandering in taakgerichtheid van de klas

In de analyses ter beantwoording van de eerste onderzoeksvraag vormen de gemiddelde score en de standaardafwijking in scores op de activiteits- en de spanningsdimensie van coping-gedrag, de aard van de situatie (gecombineerd in situaties waarin de taakgerichtheid van de klas als problematisch is ervaren en andere problematische situaties) en de interactie tussen de aard van de situatie en de coping-dimensies de predictoren. Het verschil tussen de taakgerichtheid van de klas voor en na een reactie is als criteriumvariabele opgenomen. De variabelen in de analyses en beschrijvende statistiek zijn opgenomen in Tabel 2. Omdat het voor een deel van de reacties het aantal op de video waarneembare leerlingen te klein was, beschikken we voor 187 reacties van 24 respondenten over informatie over de verandering in taakgerichtheid van de klas.

Voor de scores op de verandering in taakgerichtheid van de klas is de variantie opgedeeld in twee porties: variantie op situatie/reactieniveau en variantie op docent/les-niveau. De variantiecomponent op docent/les-niveau is vrijwel gelijk aan 0 (en de intraklasrelatie dus ook). Dit betekent dat we ervan uitgaan dat er geen variantie tussen docenten in scores op verandering in taakgerichtheid bestaat. Dit houdt ook in dat we voor de beantwoording van

Tabel 3

Resultaten van de regressie-analyses voor verandering in taakgerichtheid van de klas. Weergegeven zijn de significante (bij $\alpha=0,05$) regressiecoëfficiënten β , de standaardfout (tussen haakjes), de proportie verklaarde variantie in totaal (R^2) en voor het verschil tussen de stappen (R^2_{verschil}), en de F-waarde behorend bij het verschil

	stap 1	stap 2	stap 3
aard problematische situatie SIT			
coping-gedrag			
GEM-ACT			
GEM-SPA			
SD-ACT			
SD-SPA			
interacties			
SIT*GEM-ACT			-0,15 (0,07)
SIT*GEM-SPA			-0,33 (0,10)
SIT*SD-ACT	-0,20 ((0,07)	-0,34 (0,10)	0,21 (0,10)
SIT*SD-SPA		0,20 (0,10)	
R^2_{verschil}	0,04	0,03	0,02
F_{verschil}	7,9	4,5	4,4
R^2	0,04	0,06	0,09

deze onderzoeksvraag gebruik maken van 'klassieke' regressie-analyses, waarvan de resultaten zijn samengevat in Tabel 3⁵.

Opvallend is dat de constante niet significant afwijkt van 0. Dit betekent dat over het algemeen de aanstaande docenten in ons onderzoek niet in staat zijn in een les de taakgerichtheid van de klas door hun reactie te veranderen. In Tabel 3 is aangegeven dat in het uiteindelijke model drie interactie-effecten zijn geselecteerd. Zij verklaren negen procent van de variantie in verandering van taakgerichtheid van de klas. Er zijn geen significante hoofdeffecten gevonden. Dit bevestigt onze verwachting dat de relatie tussen coping-gedrag en de verandering in taakgerichtheid van de klas niet voor alle typen situaties hetzelfde is, en zelfs afwezig is, wanneer we informatie over de aard van de problematische situatie niet in de analyses opnemen. De interactie-effecten kunnen als volgt worden beschreven.

Wat betreft de gemiddelde score op de spanningsdimensie (SIT*GEM-SPA) blijkt dat in situaties waarin de taakgerichtheid van de klas als problematisch is ervaren, het verband met de verandering in taakgerichtheid sterker negatief is dan in andersoortige problematische klassesituaties. Zoals eerder aangegeven, is de score op verandering in taakgerichtheid lager naarmate de klas verandert richting meer taak-

gericht gedrag (na een reactie). Dit betekent dat het genoemde interactie-effect inhoudt dat naarmate aanstaande docenten in situaties waarin de taakgerichtheid van de klas als problematisch is ervaren, meer spanning vertonen in de relatie met leerlingen, zij de klas meer veranderen in de richting van taakgericht gedrag. Bijvoorbeeld, naarmate aanstaande docenten meer of sterkere disciplinemaatregelen nemen in dit type situaties, wordt de klas meer taakgericht.

Wat betreft de spreiding in scores op de activiteitsdimensie (SIT*SD-ACT) komt naar voren dat in situaties waarin de taakgerichtheid als problematisch is ervaren, het verband met de verandering in taakgerichtheid negatief is, terwijl er in andersoortige situaties geen verband bestaat. Dit betekent dat naarmate in eerstgenoemde situatietypen aanstaande docenten in een reactie meer variatie vertonen in hun gedrag wat betreft de activiteitsdimensie, zij de klas veranderen in de richting van taakgerichtheid na een reactie. Bijvoorbeeld, naarmate aanstaande docenten afwisselend ingrijpen in de klasomgeving door bijvoorbeeld disciplinemaatregelen, 'beurten' geven, of antwoorden belonen of afwijzen, en zich afzijdig houden, zijn zij meer in staat de klas te veranderen richting taakgerichtheid.

Wat betreft de spreiding in scores op de

spanningsdimensie (SIT*SD-SPA) blijkt dat in situaties waarin de taakgerichtheid van de leerlingen als problematisch is ervaren, het verband met de verandering in taakgerichtheid positief is, terwijl in de andere situatietypen geen verband bestaat. Dit betekent dat naarmate in de eerstgenoemde situaties aanstaande docenten *minder* variatie vertonen in hun gedrag wat betreft de spanningsdimensie, zij de klas meer veranderen in de richting van taakgericht gedrag na een reactie. Bijvoorbeeld, naarmate aanstaande docenten zich beperken tot disciplinemaatregelen (en deze niet afwisselen met bijvoorbeeld presenteren van de lesstof of

inhoudelijke vragen aan leerlingen stellen), zijn zij meer in staat de klas te veranderen richting taakgerichtheid.

Wanneer we effecten op korte termijn in termen van verandering naar meer taakgerichtheid van de klas interpreteren, blijkt, kort samengevat, effectief coping-gedrag van aanstaande docenten te bestaan uit een (monotone) nadruk op het oproepen van spanning en het variëren in intensiteit van activiteiten. Dit geldt uitsluitend voor situaties waarin de taakgerichtheid van de klas als problematisch is ervaren. Bijvoorbeeld, dit betekent dat docenten in een situatie waarin de klas aan het begin van les

Tabel 4

Variabelen in de analyse van de relatie tussen coping-gedrag en de ervaren tevredenheid met de eigen reactie

variabele	beschrijving
coping-gedrag	GEM-ACT gemiddelde score per reactie op de activiteitsdimensie (gemiddelde = 0,06; standaardafwijking = 0,59)
	GEM-SPA gemiddelde score per reactie op de spanningsdimensie (gemiddelde = -0,04; standaardafwijking = 0,64)
	SD-ACT standaardafwijking in scores per reactie op de activiteitsdimensie (gemiddelde 0,72; standaardafwijking = 0,50)
	SD-SPA standaardafwijking in scores per reactie op de spanningsdimensie (gemiddelde 0,79; standaardafwijking = 0,55)
Eerste inschatting	GEM-E1 gemiddelde score per reactie op arousal-dimensie (gemiddelde = 3,42; standaardafwijking = 0,40)
	GEM-E2 gemiddelde score per reactie op waarderingsdimensie (gemiddelde = 2,55; standaardafwijking = 0,52)
	GEM-E3 gemiddelde score per reactie op belemmeringsdimensie (gemiddelde = 2,75; standaardafwijking = 0,33)
	GEM-E4 gemiddelde score per reactie op zekerheidsdimensie (gemiddelde = 3,21; standaardafwijking = 0,50)
	GEM-E5 gemiddelde score per reactie op verantwoordelijkheidsdimensie (gemiddelde = 2,49; standaardafwijking = 0,73)
	GEM-E6 gemiddelde score per reactie op intensionaliteitsdimensie (gemiddelde = 3,00; standaardafwijking = 0,35)
Tweede inschatting	GEM-T1 gemiddelde score per reactie op veranderingsdimensie (gemiddelde = 3,74; standaardafwijking = 0,90)
	GEM-T2 gemiddelde score per reactie op activiteitsdimensie (gemiddelde = 3,26; standaardafwijking = 0,66)
	GEM-T3 gemiddelde score per reactie op zekerheidsdimensie (gemiddelde = 3,18; standaardafwijking = 0,75)
	GEM-T4 gemiddelde score per reactie op verantwoordelijkheidsdimensie (gemiddelde = 4,03; standaardafwijking = 0,66)
	GEM-T5 gemiddelde score per reactie op effectiviteitsdimensie (gemiddelde = 2,95; standaardafwijking = 0,34)
ervaren tevredenheid	TEVREDEN gemiddelde score per reactie op de ervaren tevredenheid (gemiddelde = 2,76; standaardafwijking = 1,02)

Tabel 5

Resultaten van de meerniveau-analyses voor de ervaren tevredenheid met de eigen reactie. Weergegeven zijn de schatting, de standaardfout in de schatting (tussen haakjes), de totale proportie verklaarde variantie op situatie/reactieniveau, de passing van het model ten opzichte van het lege model ($2*\log(lh)$) en het daarbij behorende aantal vrijheidsgraden ($df_{verschil}$)

	model 1 eerste inschatting	model 2 tweede inschatting	model 3 covariaten en coping- gedrag
coping-gedrag			
GEM-ACT			
GEM-SPA			
SD-ACT			
SD-SPA			
eerste inschatting			
GEM-E1	-		
GEM-E2	0,16 (0,07)		0,13 (0,07)
GEM-E3	-		
GEM-E4	-		
GEM-E5	-		
GEM-E6	-		
tweede inschatting			
GEM-T1		-	
GEM-T2		-	
GEM-T3		-	
GEM-T4		-0,15 (0,08)	-0,15 (0,08)
GEM-T5		0,34 (0,08)	0,32 (0,08)
$2*\log(lh)_{verschil}$	5,1	19,0	22,6
$df_{verschil}$	1	2	3
verklaarde variantie reactieniveau	0,02	0,07	0,08

zeer onrustig is, taakgericht gedrag bevoornden, wanneer zij ordemaatregelen afwisselen met zwijgend de effecten hiervan afwachten. Minder effectief is het om in deze situatie ordemaatregelen af te wisselen met (pogingen om te) beginnen met het presenteren van de lesstof of het overhoren van het huiswerk.

4.2 Coping-gedrag en ervaren tevredenheid met de reactie

In de analyses ter beantwoording van de tweede onderzoeksvraag vormen de gemiddelde score en de standaardafwijking in scores op de activiteits- en de spanningsdimensie van coping-gedrag de predictoren. De 11 beoordelingsschalen met betrekking tot eerste en tweede inschatting zijn als covariaten in de analyses opgenomen. De ervaren tevredenheid met de eigen reactie is als criteriumvariabele opgenomen. De variabelen in de analyses en beschrijvende statistiek zijn weergegeven in Tabel 4. Omdat niet alle respondenten over alle reacties uitspraken hebben gedaan over de inschatting van een situatie en/of de ervaren tevredenheid

met de reacties, beschikken we voor 184 reacties van alle 27 respondenten over de benodigde informatie.

Voor de scores op ervaren tevredenheid is de variantie opgedeeld in twee porties: variantie op situatie/reactie-niveau en variantie op docent/les-niveau. De variantiecomponent op docent/les-niveau is 0,21 en op situatie/reactie-niveau 0,84. Dit betekent dat de intraklas correlatie gelijk is aan 0,20; ofwel 20 procent van de variantie in scores op ervaren tevredenheid wordt verklaard door verschillen tussen docenten. Dit houdt in dat we voor de beantwoording van de tweede onderzoeksvraag gebruik maken van meerniveau-analyses⁶. De resultaten zijn samengevat in Tabel 5.

In Tabel 5 is aangegeven dat drie covariaten in de uiteindelijke analyses opgenomen zijn. Dit zijn de gemiddelde score op de waardering van een situatie (GEM-E2), de mate waarin zij zichzelf verantwoordelijk zien voor hun reactie (GEM-T4), en de mate waarin wordt ingeschat dat hun reactie effectief is (GEM-T5). Deze effecten houden in dat naarmate aanstaande

docenten een situatie als minder negatief beoordelen, zich minder verantwoordelijk achten voor hun reactie, en hun reactie als meer effectief inschatten, zij meer tevreden zijn met hun reactie. Deze covariaten verklaren acht procent van de variantie in ervaren tevredenheid (op situatie/reactie-niveau). We zijn niet in staat effecten van de vier maten van coping-gedrag op de ervaren tevredenheid aan te tonen. Dit betekent dat verschillen in coping-gedrag tussen de diverse reacties niets (significant) verklaren van de verschillen in ervaren tevredenheid met de reacties.

5 Conclusie en discussie

In dit artikel zijn twee effecten op korte termijn onderzocht van coping-gedrag van aanstaande docenten. Van 27 docenten hebben we ongeveer 300 reacties op problematische klas-situaties in kaart gebracht en in verband gebracht met de verandering in taakgerichtheid van leerlingen na een reactie en de ervaren tevredenheid met de eigen reactie. Wat betreft de verandering in taakgerichtheid van de klas zijn drie (interactie-)effecten aangetoond. In situaties waarin de (afwezigheid van) taakgerichtheid van leerlingen als problematisch is ervaren, bestaat effectief coping-gedrag van aanstaande docenten uit een (monotone) nadruk op het oproepen van spanning en het variëren in intensiteit van activiteiten. Bijvoorbeeld, dit betekent dat docenten in een situatie waarin de klas aan het begin van de les zeer onrustig is, taakgericht gedrag bevorderen, wanneer zij ordemaatregelen afwisselen met zwijgend de effecten hiervan afwachten. Minder effectief is het om in deze situatie ordemaatregelen af te wisselen met (pogingen om te) beginnen met het presenteren van de lesstof of het overhoren van het huiswerk. Wat betreft de ervaren tevredenheid zijn er geen significante verbanden aangetoond met het coping-gedrag van aanstaande docenten.

Lazarus en Folkman (1984) hebben aangegeven dat de probleemgerichte functie van coping afhankelijk is van de context waarin het gedrag zich voordoet, en dat emotieregulerende coping meer bepaald wordt door persoonskenmerken. In ons onderzoek hebben we hiervoor twee, zij het bescheiden, bevestigingen

gevonden. Ten eerste zijn er nauwelijks verschillen tussen docenten wat betreft de verandering van de taakgerichtheid van de klas die in ons onderzoek is gebruikt om de probleemgerichte functie van coping-gedrag in kaart te brengen. Daarentegen verschillen docenten wel wat betreft de ervaren tevredenheid met de eigen reactie. Deze tevredenheid is in ons onderzoek gebruikt om de emotieregulerende functie van coping-gedrag te onderzoeken. Een tweede aanwijzing voor het door Lazarus en Folkman veronderstelde verschil tussen probleemgerichte en emotieregulerende coping, zijn de gevonden effecten op korte termijn. Wat betreft de verandering in taakgerichtheid van de klas hebben we alleen interactie-effecten tussen de aard van de problematische situatie en coping-gedrag, maar geen hoofdeffecten gevonden. Dit is een aanwijzing dat informatie over de aard van de problematische situatie van belang is om te kunnen voorspellen welk coping-gedrag het meest effectief is om taakgericht gedrag van de klas te veroorzaken. Wat betreft de ervaren tevredenheid zijn geen effecten van coping-gedrag aangetoond. Daarvoor hebben wij de volgende verklaring. Uit Lazarus en Folkman (1984) kan worden opge-maakt dat de emotieregulerende functie van coping in cognities van personen tot uiting komt. Uit onderzoek naar interactieve cognities van docenten (zie bijvoorbeeld Yinger, 1986) komt naar voren dat docenten tijdens het lesgeven nauwelijks tijd hebben voor dergelijke interactieve cognities. Dit maakt aannemelijk dat emotieregulerend coping-gedrag niet tijdens het lesgeven, maar vooral na afloop van de les plaatsvindt. Dat zou een verklaring kunnen zijn voor de afwezigheid in ons onderzoek van significante verbanden tussen coping-gedrag en ervaren tevredenheid met de eigen reactie.

Eerder in dit artikel (zie Figuur 1) is een schema gepresenteerd van kenmerken van stress en coping waarop onderzoek zich zou moeten richten. In het onderhavige onderzoek is een beperkt deel van het schema in het onderzoek opgenomen. In verband met interpretatie van de gevonden resultaten en mogelijkheden voor vervolgonderzoek zouden wij drie aspecten uit dit schema willen toelichten.

Het eerste aspect betreft de gevonden resultaten met betrekking tot de korte-termijneffecten.

ten in relatie tot effecten op lange termijn. De gevonden aanwijzingen voor positieve kortetermijneffecten op taakgerichtheid van de klas kunnen op gespannen voet staan met consequenties op langere termijn. Wubbels, Créton en Holvast (1988) hebben aannemelijk gemaakt dat ordemaatregelen, zoals het waarschuwen van leerlingen, en leerlingen de klas uit sturen, kunnen resulteren in een problematische interactie tussen docent en leerlingen. Zij maken duidelijk dat in dergelijke interacties het ordeverstorende gedrag van leerlingen niet alleen wordt gevolgd door corrigerend gedrag van de docent, maar dat leerlingen de correcties van de docent ook gebruiken om de orde te verstoren. Agressief gedrag van de één wordt door de ander aangegrepen als een prikkel voor nog agressiever gedrag. Docent en leerlingen kunnen verstrikt raken in een vicieuze cirkel waarin beide partijen elkaars gedrag symmetrisch kunnen versterken. Resultaten uit het onderhavige onderzoek met betrekking tot de kortetermijneffecten geven aan dat de kans op een dergelijke, op lange termijn, symmetrische escalatie bij aanstaande docenten reëel is. In de begeleiding van aanstaande docenten zou dus benadrukt moeten worden dat het bovengenoemde onderzoeksresultaat, dat lijkt te wijzen op het belang van ordemaatregelen voor de taakgerichtheid van de klas op korte termijn, tot andere effecten op lange termijn kan leiden. Deze bewering is echter niet gebaseerd op empirische resultaten van het onderhavige onderzoek.

Ten tweede willen wij wijzen op het mogelijke belang van het in kaart brengen van (persoons)kenmerken van de docent. In ons onderzoek is van elke aanstaande docent slechts één les in het onderzoek betrokken. Hierdoor blijft het onduidelijk in welke mate de gevonden resultaten kunnen worden toegeschreven aan verschillen in stabiele (persoons)kenmerken van docenten, of aan verschillen in (over lessen wisselend) coping-gedrag. Wellicht dat variabelen op docentniveau, zoals de coping-stijl (bijvoorbeeld gemeten met behulp van de Utrechts Coping Lijst (UCL); Schreurs, Van de Willige, Tellegen & Brosschot, 1988), de interpersoonlijke stijl (bijvoorbeeld gemeten met de Vragenlijst Interpersoonlijk Leraarsgedrag (VIL); Créton & Wubbels, 1984) of persoonlijkheidstrekken, zoals bijvoorbeeld gemeten

met de Myers-Briggs Type Indicator (MBTI; Myers & McCaulley, 1985) in toekomstig onderzoek naar stress en coping moeten worden opgenomen.

Ten derde wijzen wij op de rol van de opleiding bij het voorkomen van stress bij en het bevorderen van adequaat coping-gedrag van aanstaande docenten (in Figuur 1 aangeduid op het maatschappelijk niveau). Opmerkelijk in ons onderzoek is dat de aanstaande docenten in een les over het algemeen geen verandering in de taakgerichtheid van leerlingen na een reactie realiseren. De resultaten in het onderhavige onderzoek kunnen daarom gezien worden als een uitdaging voor de opleiding van docenten. Daarbij moet echter gewezen worden op het gegeven dat de communicatie tussen docent en leerlingen een systeemkarakter vertoont (Créton, Wubbels & Hooymayers, 1993). Het systeemkarakter van de communicatie tussen docent en leerlingen wordt onder meer gekenmerkt door een mate van stabiliteit en weerstand tegen verandering. Wanneer er een bepaalde mate van stabiliteit is bereikt, aarzelen zowel leerlingen als de docent om te veranderen. De eerste dagen op school zetten een trend voor de rest van het jaar, waarbij een eenmaal gestabiliseerd patroon moeilijk te veranderen is. Dat is wellicht een tweede verklaring voor de afwezigheid in ons onderzoek van een verband tussen coping-gedrag en de ervaren tevredenheid. De resultaten van dit onderzoek leiden dan ook tot de hypothese dat de opleiding zich dan ook kan richten op het uitrusten van aanstaande docenten met kennis, inzichten en vaardigheden die voorkómen dat er problematische situaties ontstaan.

Noten

1. Coping-stijl is gemeten door middel van de Utrechtse Coping-Lijst (UCL, Schreurs, Van de Willige, Tellegen & Brosschot, 1988).
2. Cronbach en Gleser (1953) onderscheiden drie kenmerken die bepalend zijn voor de overeenkomst tussen profielen van scores, namelijk niveau ('elevation'), spreiding ('scatter') en vorm ('shape'). Omdat door het gebruik van 'event sampling' in ons onderzoek acties niet noodzakelijk even lang duren, is het naar onze mening niet zinvol het scoreprofiel te karakteriseren aan

de hand van de vorm van het scoreprofiel.

3. Achteraf is gebleken dat de respondenten vrijwel uitsluitend situaties noemen die zij als belastend of problematisch hebben ervaren. Daarom spreken we in dit artikel van 'problematische situaties'.
4. Uit controle-analyses is gebleken dat respondenten die tijdens de triostage in het onderzoek zijn betrokken, niet verschillen van de andere respondenten wat betreft de scores op beide criteriumvariabelen en op coping-gedrag. Een uitzondering vormt de standaardafwijking in scores op de activiteitsdimensie: studenten uit de individuele eindstage variëren meer in de intensiteit van de waargenomen activiteiten (bij $\alpha=0,05$) dan studenten uit de triostage. In de regressie-analyses zal, gezien de geringe verschillen, geen onderscheid gemaakt worden in beide typen stage.
5. Bij het uitvoeren van de regressie-analyses hebben we de volgende stapsgewijze procedure gevolgd. In elke stap is de variabele in het model toegevoegd die de meeste variantie verklaart. Deze stappen zijn herhaald totdat geen enkele variabele het model significant verbetert.
6. We hebben daarbij de volgende procedure gevolgd. In de eerste analyses is bepaald welke covariaten in de uiteindelijke analyses moeten worden opgenomen. Dit is afzonderlijk gedaan voor de eerste en de tweede inschatting (respectievelijk model 1 en 2). Vervolgens zijn de predictoren en de eventuele covariaten in de analyses opgenomen (model 3). In alle analyses is net zoals in de eerder beschreven 'klassieke' regressie-analyses een stapsgewijze procedure gevolgd. In elke stap is die variabele in het model opgenomen die de meeste verklaarde variantie aan het (in de eerste stap nog lege) model toevoegt. Deze stappen zijn herhaald totdat geen enkele variabele het model significant verbetert. De scores op de 11 beoordelingsschalen zijn getransformeerd tot standaardcores met een gemiddelde van 0 en een standaardafwijking van 1.

Literatuur

- Admiraal, W.F. (1994). *Reacties van docenten op aandachtseisende situaties in de klas*. Utrecht: W.C.C.
- Admiraal, W., Wubbels, Th., & Korthagen, F. (1996a). Gedrag van aanstaande docenten in de aandachtseisende situaties in de klas. *Tijdschrift voor Onderwijsresearch*, 21(2), 162-181.
- Admiraal, W., Wubbels, Th., & Korthagen, F. (1996b). Student teacher behaviour in response to daily hassles in the classroom. *Social Psychology of Education*, 1(1), 25-46.
- Bergen, Th.C.M., & Opdorp, C.A.W. van (1989). Probleemsituaties in het onderwijs. *Pedagogische Studiën*, 66(9), 327-340.
- Brophy, J.E., & Good, T.L. (1986). Teacher behavior and student achievement. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd edition), (pp. 328-375). New York: MacMillan P.C.
- Créton, H.A., & Wubbels, Th. (1984). *Ordeproblemen bij beginnende leraren*. Utrecht: W.C.C..
- Créton, H., Wubbels, Th., & Hooymayers, H. (1993). A systems perspective on classroom communication. In T. Wubbels & J. Levy (Eds.), *Do you know what you look like?*, (pp. 1-12). London: The Falmer Press.
- Cronbach, L.J., & Gleser, G.C. (1953). Assessing similarity between profiles. *Psychological Bulletin*, 50(6), 456-473.
- Dewe, P.J. (1985). Coping with work stress: an investigation of teachers' actions. *Research in Education*, 33, 27-40.
- Doyle, W. (1986). Classroom organization and management. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd edition), (pp. 392-431). New York: MacMillan P.C.
- Evertson, C.M. (1989). Improving elementary classroom management: a school-based training program for beginning the year. *Journal of Educational Research*, 83(2), 82-90.
- Evertson, C.M., & Burry, J.A. (1988, April). *Capturing classroom context: the observation instrument as lens for assessment*. Paper gepresenteerd op de jaarlijkse bijeenkomst van de American Educational Research Association, New Orleans.
- Farell, D. (1983). Exit, voice, loyalty, and neglect as responses to job satisfaction: a multidimensional scaling study. *Academy of Management Journal*, 26(4), 596-607.
- Folkman, S., & Lazarus, R.S. (1985). If it changes it must be a process: study of emotion and coping

- during three stages of a college examination. *Journal of Personality and Social Psychology*, 48(1), 150-170.
- Freeman, A. (1987). The coping teacher. *Research in Education*, 38, 1-38.
- Fuller, F.F., & Bown, O. (1975). Becoming a teacher. In K. Ryan (Ed.), *The 74th yearbook of the National Society for the Study of Education*, (pp. 25-52). Chicago: The University of Chicago Press.
- Gifi, A. (1983). *Principals user's guide*. Leiden: Department of Datatheory.
- Hertog, P.C. den (1990). *Persoonlijkheidskenmerken en causale attributies*. Amsterdam: Thesis Publishers.
- Koetsier, C.P. (1991). *Een brug tussen opleiding en praktijk*. Utrecht: W.C.C..
- Lazarus, R.S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer P.C..
- Lazarus, R.S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality*, 1(2), 141-169.
- MacDonald, C.J. (1993). Coping with stress during the teaching practicum: the student teacher's perspective. *The Alberta Journal of Educational Research*, 39(4), 407-418.
- Mauro, R., Sato, K., & Tucker, J. (1992). The role of appraisal in human emotions: a cross-cultural study. *Journal of Personality and Social Psychology*, 62(2), 301-317.
- Myers, I.B., & McCaulley, M.H. (1985). *Manual: a guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.
- Nunnally, J.C. (1978). *Psychometric Theory* (2nd edition). New York: McGraw-Hill.
- Peters, V.A.M. (1985). *Docenten en hun probleem-situaties*. Nijmegen: Academisch Proefschrift.
- Popping, R. (1988). On agreement indices for nominal data. In W.E. Saris & I.N. Gallhofer (Eds.), *Sociometric Research, volume 1*. Basingstoke, (pp. 90-105). Hampshire: MacMillan.
- Rudow, B. (1990). Empirische Untersuchungen zur Belastung und Beanspruchung in Lehrerberuf. *Zeitschrift für Pädagogische Psychologie*, 4(2), 75-85.
- Schreurs, P.J.G., Willige, G. van de, Tellegen, B., & Brosschot, J.F. (1988). *De Utrechtse Coping Lijst: UCL* (Handleiding). Lisse: Swets & Zeitlinger.
- Selye, H. (1979). *The stress of life* (rev. edition). New York: Van Nostrand Reinhold.
- Smith, C.A., & Ellsworth, Ph.C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, 48(4), 813-838.
- Thoits, P.A. (1983). Dimensions of life events as influences upon the genesis of psychological distress and associated conditions: an evaluation and synthesis of the literature. In H.B. Kaplan (Ed.), *Psychosocial stress: trends in theory and research*, (pp. 33-103). New York: Academic Press.
- Veenman, S.A.M. (1984). Perceived problems of beginning teachers. *Review of Educational Research*, 54(2), 143-178.
- Wahl, D. (1991). *Handeln unter Druck*. Weinheim: Deutscher Studien Verlag.
- Westerhof, K.J. (1989). *Effectiviteit van leerkrachtgedrag*. Groningen: Academisch Proefschrift.
- Willett, J.B. (1989). Questions and answers in the measurement of change. *Review of Research in Education*, 15, 345-422.
- Wood, F.H., & Dorsey, B. (1989). Aversiveness of teacher-chosen interventions and student problem characteristics: is there a relationship? *Psychology in the Schools*, 26, 389-397.
- Woodhouse, G. Rasbash, J., Goldstein, H., Yang, M., & Plewis, I. (1996). *Multilevel modelling applications: a guide for users of Mln*. London: University of London.
- Wubbels, T., Créton, H.A., & Holvast, A. (1988). Undesirable classroom situations: a systems communication perspective. *Interchange*, 19(2), 25-40.
- Wulk, J. (1988). *Lehrerbelastung*. Frankfurt: Peter Lang.
- Yinger, R.J. (1986). Examining thought in action: a theoretical and methodological critique of research on interactive teaching. *Teaching & Teacher Education*, 2(3), 263-282.

Manuscript aanvaard 3-1-1997

Auteurs

W.F. Admiraal – onderzoeker,

F.A.J. Korthagen – senior docent-onderzoeker,

Th. Wubbels – hoogleraar

zijn werkzaam bij het Interfacultair Instituut voor de Lerarenopleiding, Onderwijsontwikkeling en Studietoelichting (IVLOS) van de Universiteit Utrecht.

Adres: IVLOS – Universiteit Utrecht, Heidelberglaan 8, 3584 CS Utrecht.

Abstract

Immediate outcomes of student teachers' coping

W.F. Admiraal, F.A.J. Korthagen & Th. Wubbels.
Pedagogische Studiën, 1997, 74, 165-182.

In teaching, stressful encounters frequently occur. Ineffective ways of coping with these encounters may result in feelings of distress and a decrease of learning possibilities for pupils. The present research examines coping during teaching practices by evaluating the relation between classroom behaviour in response to daily hassles in the classroom and immediate outcomes in a sample of 27 student teachers, enrolled in our teacher education programme. Subjects' experienced daily hassles in the classroom, the ways they coped with the demands of these hassles, and two immediate outcomes (pupils' time-on-task and the subjects' self-evaluation of their coping responses) were assessed. About 300 responses and hassling events were characterised. The relation between coping and the two outcomes was evaluated by a series of (multi level) hierarchical regression analyses.

In general, student teachers do not change the pupils' time-on-task in a lesson. However, differences were found between student teachers, that were to some extent explained by student teachers' coping behaviour. In situations in which the pupils' time-on-task had been appraised as hassling, student teachers demonstrated effective coping behaviour as they monotonously aroused tension in the relationship with pupils, and varied the intensity of activities. For example, in a disorderly classroom at the beginning of a lesson, a student teacher accomplished more time-on-task, when disciplinary actions were alternated with silently waiting for the effects of this actions. Less effective student teacher behaviour consisted of alternating disciplinary action with efforts to start with the lesson by instructing or testing homework.

Leerpotentieelonderzoek bij allochtone leerlingen: een verkenning met de Leerpotentieeltest voor Inductief Redeneren

A.P.A.C. Slenders en W.C.M. Resing

Samenvatting

In dit onderzoek staat de vraag centraal of de Leerpotentieeltest voor Inductief Redeneren (LIR) bruikbare informatie oplevert over het leerpotentieel van allochtone leerlingen uit het basisonderwijs en het onderwijs aan moeilijk lerende kinderen (MLK). De onderzoeksgroep bestaat uit 7 à 8-jarige basisschoolleerlingen en 10 à 13-jarige MLK-leerlingen, met gelijke mentale leeftijd en aanvangsniveau. Geconcludeerd wordt dat de allochtone MLK-leerlingen en de drie jaar jongere basisschoolleerlingen tijdens de leerpotentieeltrainingen evenveel hulp nodig hebben en dat de training een positieve invloed heeft op de prestaties van beide groepen leerlingen op diverse inductieve redeneertaken. Deze resultaten komen overeen met resultaten uit onderzoek dat eerder werd uitgevoerd bij 7 à 8-jarige autochtone leerlingen. De LIR blijkt een diagnostisch waardevol instrument te zijn voor allochtone leerlingen.

Inleiding

Voor zwakke schoolvorderingen van allochtone leerlingen zijn diverse verklaringen mogelijk. Een matige kennis van de Nederlandse taal kan hen parten spelen, evenals ontoereikende cognitieve mogelijkheden. Het in kaart brengen van de leermogelijkheden van allochtone kinderen is niet eenvoudig. De idee is dat traditionele intelligentietests hierin tekort schieten. Het construeren van nieuwe diagnostische instrumenten, waaronder leerpotentieeltests, is één van de mogelijke benaderingswijzen waarmee de problemen, die bij het testen van allochtone kinderen naar voren komen, kunnen worden aangepakt. Leerpotentieeltests geven namelijk niet alleen een beeld van het momentane niveau van de ontwikkeling van (een be-

paald aspect van) intelligentie, maar ook van het vermogen om van leerervaringen te profiteren (Hessels, 1993; Resing, 1990; 1993b; 1997a). Daardoor wordt duidelijk of kinderen door middel van gerichte instructie en hulp al dan niet in staat zijn hun schoolvorderingen op een hoger niveau te brengen. In het onderhavige onderzoek staat het gebruik van de LIR, de Leerpotentieeltest voor Inductief Redeneren, bij Turkse en Marokkaanse leerlingen afkomstig uit het basis- en het MLK-onderwijs centraal.

Het gebruik van traditionele intelligentietests ter bepaling van het intelligentieniveau wordt nogal eens bekritiseerd (zie onder andere Budoff, 1987a; Feuerstein, Rand & Hoffman, 1979; Hamers en Ruijsenaars, 1984). Zo zouden intelligentietests slechts meten wat leerlingen voorafgaand aan de testafname hebben geleerd. Bovendien bieden intelligentietests nauwelijks aangrijpingspunten voor interventie en leveren zij een onvoldoende bijdrage aan het vormgeven van handelingsvoorschriften voor leerkrachten. Naast deze algemene punten van kritiek gelden voor allochtone leerlingen enkele specifieke kritiekpunten. Zelfs indien gecorrigeerd wordt voor sociaal-economisch niveau, behalen zij over het algemeen lagere scores dan autochtone leerlingen op intelligentietests die oorspronkelijk voor autochtone leerlingen zijn ontwikkeld, zoals de RAKIT, SON-R en WISC-R (Bleichrodt, Resing, Drenth & Zaal, 1987; Resing, Bleichrodt & Drenth, 1986; Snijders, Tellegen & Laros, 1988; Wechsler, 1974). Bij het gebruik van intelligentietests wordt er van uitgegaan dat alle leerlingen in gelijke mate de kans hebben gehad zich de, voor de test benodigde, kennis en vaardigheden eigen te maken. Vooral bij allochtone leerlingen blijkt deze vooronderstelling moeilijk vol te houden. Daarnaast hebben allochtone leerlingen over het algemeen weinig

ervaring met 'test-situaties' en worden zij benadeeld door de veelal talige instructies bij de taken (Babad & Budoff, 1974; Hessels, 1993; Testscreeningscommissie, 1990).

Door afname van leerpotentieeltests wordt onder andere getracht de problemen die ontstaan bij het testen van allochtone leerlingen, waar mogelijk, te ondervangen. Voor een algemeen overzicht van onderzoek op het terrein van leerpotentieel wordt verwezen naar de overzichtswerken van Lidz (1987) in de Verenigde Staten en Hamers, Sijtsma en Ruijsse-naars (1993) in Europa.

De LIR is opgezet vanuit het pretest-training-posttest model en onderscheidt zich hiermee van een traditionele intelligentietest doordat het kind niet één maar meerdere malen (meestal twee keer) wordt getest, waarbij tussen de meetmomenten een trainingsfase is ingelast. De tester is actief bij het hele testgebeuren betrokken en ziet het kind frequenter dan het geval is bij standaard testafname. Er ontstaat als het ware een miniatuur-simulatie van het onderwijsleerproces.

De mate waarin het kind hulp nodig heeft in de trainingsfase en de prestaties op de posttest vormen volgens diverse onderzoekers een indicatie voor het leervermogen (Hamers, Castellijns & Pennings, 1989; Resing, 1993a; Wijnstra, 1987). Tijdens de trainingsfase kan immers nagegaan worden in welke mate het kind baat heeft bij de geboden hulp en van wat voor type hulp het kind het meest profiteert. Daarnaast is de trainingsfase vooral voor (allochtone) kinderen, die met de testvereisten weinig ervaring hebben opgedaan, belangrijk. De trainingsfase verlaagt de kunstmatigheid van de erop volgende testsituatie. Tevens reduceert de training de invloed van niet relevante verschillen tussen leerlingen in voorwaardelijke kennis en vaardigheden, zoals het begrijpen van instructies en dergelijke (Budoff, 1987a; Hegarty & Lucas, 1978; De Leeuw & Resing, 1985). Kortom, leersucces in de trainingsfase weerspiegelt het reële prestatievermogen van het kind ten aanzien van cognitieve en schoolse taken. Uit onderzoek met de LIR is gebleken dat testcores na training het leren op school en prestaties op schooltaken beter voorspellen dan pretestcores, c.q. de standaard intelligentietest (vgl. Resing, 1990; 1993a).

Met name in de groep zwak presterende

leerlingen leidt afname van een leerpotentieeltest tot differentiatie. Uit onderzoek van onder andere Babad en Budoff (1974), Budoff (1987a; 1987b), Hamers en Ruijsse-naars (1982), Hegarty en Lucas (1978), Hessels (1993) en Resing (1990; 1997b) blijkt dat er kinderen zijn die na gerichte training beter presteren dan op grond van traditioneel intelligentie-onderzoek het geval leek te zijn. Zij werden op grond van hun lage intelligentie ten onrechte als 'moeilijk lerend' geïnclassificeerd. Met name bij allochtone kinderen is dit gevaar aanwezig.

Onderzoek met de LIR (Resing, 1990) heeft geleid tot een instrument om het leerpotentieel van 7 à 8-jarige autochtone leerlingen uit basis- en Speciaal Onderwijs in kaart te brengen. In het onderhavige onderzoek wordt nagegaan of bij allochtone leerlingen de prestaties op de posttests, als gevolg van training, verbeteren en of er transfer naar verwante taken optreedt. Transfer wordt hier opgevat als het "effect van het zelfstandig kunnen generaliseren van het in de training geleerde naar nieuwe situaties of taken" (vgl. Resing, 1990, p. 43). Tevens staat de vraag centraal of de LIR praktisch bruikbare informatie oplevert over het leerpotentieel van allochtone leerlingen uit het basis- en MLK-onderwijs. De gerapporteerde resultaten worden in de conclusie vergeleken met gegevens verkregen uit eerder onderzoek bij autochtone leerlingen.

1 Methode

1.1 Onderzoeksgroep

De onderzoeksgroep bestond uit 40 allochtone leerlingen: 20 basisschoolleerlingen (11 jongens en 9 meisjes) en 20 kinderen uit het Speciaal Onderwijs (4 jongens en 16 meisjes). Gekozen werd voor basisschoolleerlingen en moeilijk lerende kinderen vanwege de relatief sterke toename van allochtone leerlingen in het MLK-onderwijs (Teunissen & Golhof, 1987; Wijnstra, 1987). Besloten werd het onderzoek te beperken tot Turkse en Marokkaanse leerlingen vanwege hun relatief grote vertegenwoordiging in het Speciaal Onderwijs (Teunissen & Golhof, 1987; Wijnstra, 1987) en vanwege het feit dat zij nog steeds de grootste achterstanden in leerprestaties vertonen (Hessels, 1993).

Zestig procent van de deelnemende leerlingen had de Marokkaanse nationaliteit, veertig procent de Turkse. Voor deelname aan het onderzoek werd de eis gesteld dat de leerling eenvoudige instructies in het Nederlands kon begrijpen.

Om het probleem ten aanzien van een ongelijke beginsituatie bij chronologisch equivalente leeftijdsgroepen te voorkomen (de basisschoolleerlingen behalen de hoogste scores op de pretest, de MLK-leerlingen de laagste) werden basisschool- en MLK-leerlingen van vergelijkbare mentale leeftijd in het onderzoek betrokken. Beide groepen leerlingen hadden derhalve een gelijk aanvangsniveau, uitgedrukt in termen van ruwe scores op twee taken uit de pretest: de subtests Analogieën en Exclusie van de RAKIT (Revisie Amsterdamse Kinder Intelligentie Test) (Bleichrodt, Drenth, Zaal & Resing, 1984). Op basis van testgegevens uit dossieronderzoek werd besloten MLK-leerlingen in het onderzoek te betrekken die twee tot drie jaar ouder waren dan de deelnemende basisschoolleerlingen. De gemiddelde leeftijd van de basisschoolleerlingen bedroeg zeven jaar en negen maanden (7;3-8;4), die van de MLK-leerlingen tien jaar en zes maanden (9;3-11;6). Het gemiddelde RAKIT-IQ bedroeg voor de allochtone basisschoolleerlingen 87,6 (range 74-104), voor de MLK-leerlingen 70,2 (range 57-88).

1.2 Onderzoeksofzet

Het onderzoek kent een 'randomized pretest-training-posttest control group design'. Op basis van de subtests Analogieën en Exclusie van de RAKIT werden 'randomized' blokken leerlingen gevormd om zoveel mogelijk uit te sluiten dat de groepen voorafgaand aan de training al noemenswaardig van elkaar zouden verschillen. Per schooltype werden de leerlingparen vervolgens a-select aan één van beide condities (experimenteel, controle) toegewezen. De twee experimentele groepen (tien basisschoolleerlingen en tien MLK-leerlingen) kregen de LIR aangeboden. De controlegroepen volgden in de tussentijd het onderwijs in de klas.

De gemiddelde RAKIT-IQ-scores van de experimentele en controle basisschoolleerlingen (respectievelijk 88,3 en 86,8) verschilden niet significant van elkaar. De IQ-scores van de

experimentele en controle MLK-leerlingen (respectievelijk 70,7 en 69,7) verschilden eveneens niet significant van elkaar.

Ter controle op het gebruik van 'randomized-blocks' zijn variantie-analyses uitgevoerd. De verschillen in gemiddelde pretest-scores tussen basisschool- en MLK-leerlingen bleken, zoals verwacht, niet significant te zijn. Dit gold zowel voor de subtest Exclusie ($F=.701$; $p=.407$) als voor de subtest Analogieën ($F=1.26$; $p=.268$). Ook wanneer de controlegroepen en de experimentele groepen onderling vergeleken werden, waren de verschillen niet significant.

1.3 Leerpotentieelmaten

Wat betreft het gebruik van leerpotentieelmaten zijn vanuit de literatuur twee benaderingen te onderscheiden. Budoff (Babad & Budoff, 1974; Budoff, 1987a; Budoff, 1987b) en Feuerstein (Feuerstein, Rand, Jensen, Kaniel & Tzuriel, 1987) bijvoorbeeld richten zich op de maximale hoeveelheid vooruitgang van de leerling na een vastgestelde hoeveelheid instructie. In ander onderzoek (vgl. Resing, 1990) werd leerpotentieel gedefinieerd in termen van leerefficiëntie: de minimale hoeveelheid hulp die nodig is om een gespecificeerde hoeveelheid leren en/of transfer te bewerkstelligen. Het onderhavige onderzoek gaat van deze laatste operationalisatie van leerpotentieel uit.

In het onderzoek werden de volgende leerpotentieelmaten gebruikt: de hoeveelheid benodigde hulpstappen om een leer criterium te bereiken (vier trainingsopgaven zonder hulp correct beantwoorden), het aantal trainingsopgaven waarbij hulp gegeven is en de scores na training op verschillende posttests.

1.4 Onderzoeksmateriaal

Hieronder volgt een beschrijving van de Leerpotentieeltest voor Inductief Redeneren en van de gebruikte pre- en posttests.

1.4.1 Leerpotentieeltest voor Inductief Redeneren

De LIR is ontwikkeld bij twee subtests van de RAKIT: Exclusie en Verbale Analogieën. Beide subtests bestaan uit inductieve redeneertaken. Inductief redeneren, een van de belangrijkste aspecten van intelligentie, is het vermo-

gen om vanuit specifieke ervaringen te generaliseren en nieuwe, meer abstracte concepten te vormen. De hieraan ten grondslag liggende activiteit wordt met inductie aangeduid: het genereren van algemene regels, ideeën of concepten vanuit een aantal specifieke gevallen (Nickerson, Perkins & Smith, 1985). Inductie leidt niet tot absolute geldigheid, maar tot een bepaalde graad van waarschijnlijkheid. Redenen voor het gebruik van inductieve redeneertaken als uitgangspunt voor de ontwikkeling van een leerpotentieeltest zijn hun centrale positie in vele intelligentietheorieën en theorieën omtrent cognitieve ontwikkeling en de mogelijkheid tot het maken van een taakanalyse.

Bij de subtest Exclusie gaat het erom dat het kind uit vier abstracte figuren één figuur kiest

die niet aan een, door het kind zelf ontdekte, regel voldoet waaraan de drie andere figuren wel voldoen. Bij de subtest (verbale) Analogieën (bijvoorbeeld mens : neus = olifant : ?) moet het kind een relatie leggen tussen de eerste twee termen van de analogie. Deze relatie moet vervolgens worden toegepast op de derde term, hetgeen leidt tot de vierde term die eenzelfde verband moet hebben met de derde als de tweede met de eerste.

De LIR blijkt een betrouwbaar instrument te zijn (Cronbachs Alpha van .89 voor Exclusie en .90 voor Analogieën). De soortgenoot-validiteit (i.c. de relatie met de Leertest voor Etnische Minderheden (Hamers, Hessels & Van Luit, 1991)) is hoog te noemen (zie Resing & Van Wijk, 1996).

Leerpotentieel werd in dit onderzoek gede-

ANALOGIEËN DEEL-GEHEEL A : B = C : ?	
STAP 0	Welk woord moet er in dit lege vakje komen? Je mag niet raden. +: Ja, goed. Leg uit. (naar volgend item, reden noemen) -: We gaan nog even verder, ik ga je een beetje helpen. (NAAR 1)
STAP 1	Wat moet je ook weer precies doen? Je kijkt eerst naar (A) en (B), die horen bij elkaar. Op dezelfde manier moet je een woord zoeken bij (C), dat daar net zo bij past. +: Ja, goed. Leg uit. (naar volgend item, reden noemen) -: We gaan weer even verder. (NAAR 2)
STAP 2	Hoe heb je het net ook weer gedaan? Helpt dat je om het goede woord te vinden? +: Ja, goed. Leg uit. (naar volgend item, reden noemen) -: We gaan weer even verder. (NAAR 3)
STAP 3	Denk goed aan (A) en (B). Wat komt er dan in het lege vakje? +: Ja, goed. Leg uit. (naar volgend item, reden noemen) -: We gaan weer even verder. (NAAR 4)
STAP 4	Waarom horen (A) en (B) bij elkaar? Wat hoort er dan bij (C)? +: Ja, goed. Leg uit. (naar volgend item, reden noemen) -: We gaan weer even verder. (NAAR 5)
STAP 5	(A) en (B) horen bij elkaar want (A) is een stukje van (B) . Wat hoort er dan net zo bij (C)? +: Ja, goed. Leg uit. (naar volgend item, reden noemen) -: We gaan weer even verder. (NAAR 5)
STAP 6	We gaan er zinnetjes van maken. Jij moet zeggen of het klopt. Is ... een stukje van (D)? (5x) +: Ja, goed. Leg uit. (naar volgend item, reden noemen) -: We gaan weer even verder. (NAAR 7)
STAP 7	Pl. geeft het antwoord en noemt de reden.

Figuur 1. Schematische weergave hulpstappen Analogietraining (ontleend aan Resing, 1990)

finieerd in termen van leerefficiëntie. Getracht werd te bereiken dat, als gevolg van training, de prestaties van de leerlingen verbeteren op zowel taken die erg lijken op de trainingstaken als op taken die qua inhoud of vorm daarvan iets afwijken (vgl. Resing, 1990). Om dit te bewerkstelligen zijn de trainingsprocedures gericht op zowel het aanleren van specifieke oplossingsstrategieën als op het aanleren van metacognitieve vaardigheden, waaronder controleren en plannen. Het doel van de LIR-afname was het bepalen van de minimaal benodigde hoeveelheid hulp per trainingsopgave. Essentieel in de training is dan ook het stapsgewijs ondersteunen van het uit te voeren oplossingsproces, tot op het niveau waarop de opgave voor het kind oplosbaar is. De hulpstappen zijn gestructureerd van algemene metacognitieve hints, gericht op een herdefinitie van de taak, naar meer specifieke hulpstappen met als doel het leggen van relaties tussen de kenmerken van de opgaven (zie ook Campione & Brown, 1987; Campione, Brown, Ferrara, Jones & Steinberg, 1985). Figuur 1 bevat een overzicht van de hulpstappen voor een analogie-opgave. De hulp die bij de Exclusie-opgaven wordt gegeven, heeft een vergelijkbare structuur.

1.4.2 Pre- en posttests

Voor en na de leerpotentieeltraining kregen de leerlingen diverse taken aangeboden. In het onderzoek werd onderscheid gemaakt tussen taken die het directe effect van training meten en taken die het transfereffect van training in kaart brengen. Om het directe effect te kunnen meten, zijn taken gebruikt waarvan zowel de inhoud als de oplossingsprocedure overeenkomt met of sterk lijkt op de taken uit de leerpotentieeltraining. Ten aanzien van het transfereffect werd een onderscheid gemaakt tussen twee typen transfer. Specifieke horizontale transfer doet zich voor in geval de inhoud (het materiaal) verschilt van de inhoud van de trainingstaken, maar de oplossingsprocedure dezelfde is. In geval van specifieke verticale transfer wijkt de inhoud van de transfertaken af van de inhoud van de trainingstaken en is slechts een gedeelte van de vereiste oplossingsprocedure tijdens de training aangeleerd. De in de training aangeleerde oplossingsprocedures zijn dus niet toereikend om op grond daarvan

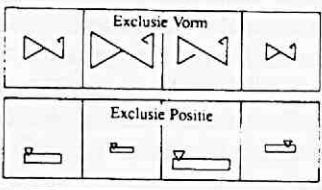
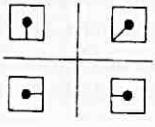
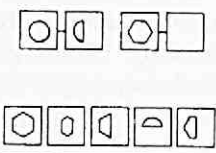
zonder meer in staat te zijn verticale transfertaken op te lossen.

1.5 Onderzoeksprocedure

De pretest werd in twee delen afgenomen. Allereerst werd van de RAKIT de verkorte vorm (waaronder de subtest Exclusie-R) en de subtest Analogieën-R afgenomen. Deze subtests werden gebruikt om specifieke horizontale transfer te kunnen meten. Het tweede deel van de pretest bestond uit taken die het directe effect (Exclusie-T en Analogieën-T) en het specifieke verticale transfereffect meten (Visuele Analogieën en Verbale Exclusie). In Figuur 2 zijn voorbeelden van opgaven uit de pre- en posttests weergegeven. De directe effectmaten Exclusie-T en Analogieën-T bestaan deels uit trainingsopgaven, deels uit vergelijkbare, nieuwe opgaven.

Na de pretest kregen de kinderen uit de experimentele groepen de Exclusie- en Analogietrainingen uit de LIR aangeboden. In de Exclusietraining wordt onderscheid gemaakt tussen trainingsopgaven waarbij het exclusieprincipe gebaseerd is op vormverandering (één plaatje hoort er niet bij, want dat heeft een andere vorm) dan wel positieverandering (één plaatje hoort er niet bij, want dat staat op een andere plaats). In de Analogietraining komen twee relaties voor, deel-geheel (...is een stukje van...) en functie (...gebruik je om te...). In Figuur 2 zijn voorbeelden van opgaven uit de trainingen opgenomen. Zowel de Exclusie- als de Analogietraining omvat drie deeltrainingen. Tijdens de eerste twee deeltrainingen wordt op één transformatie of relatie getraind. In de derde deeltraining krijgen de kinderen, ter stimulering van flexibel strategiegebruik, de verschillende soorten opgaven door elkaar aangeboden. Trainingssessies worden afgebroken wanneer een leercriterium – vier achtereenvolgende trainingsopgaven zonder hulp correct beantwoorden en een juiste reden geven voor het gekozen antwoord – wordt bereikt. De overige trainingsopgaven worden na het bereiken van dit criterium zonder hulp of feedback aan de kinderen voorgelegd.

De leerlingen volgden twee deeltrainingen per week. De totale training nam een periode van drie weken in beslag. Twee weken na de laatste deeltraining kregen alle leerlingen de posttest. In het eerste deel van de posttest wer-

TYPE TRANSFER	
Direct effect (tevens voorbeelden trainingsopgaven)	
Exclusie-T 	Analogieën-T (functie en deel-geheel) potlood : schrijven = mond : ... (lippen huilen tong praten) jas : kraag = schoen : ... (laars hak sok voet)
Specifieke, horizontale transfer	
Exclusie-R 	Analogieën-R jongen : huis = vos : ... (hol wolf hok schuur)
Specifieke, verticale transfer	
Verbale Exclusie raam deur dak huis	Visuele Analogieën 

Figuur 2. Voorbeelden transfer- en trainingsopgaven

den Exclusie-R en Analogieën-R van de RAKIT afgenomen. Enkele dagen daarna kregen de kinderen Exclusie-T, Analogieën-T, Verbale Exclusie en Visuele Analogieën opnieuw voorgelegd.

2 Resultaten

In deze paragraaf worden enkele gegevens over het directe en transfereffect van de LIR gepresenteerd. Daarnaast komen resultaten met betrekking tot de leerpotentieelmaten en moeilijkheidsgraad van de training aan de orde. Afgesloten wordt met informatie over individuele leerpotentieelprofielen.

2.1 Effect van training

Nagegaan werd of afname van de LIR heeft geleid tot verbeterde prestaties op zowel de directe effectmaten als de verschillende transfertaken. In Tabel 1 zijn, voor zowel de basis-

school- als de MLK-leerlingen, per conditie de gemiddelde scores op de pre- en posttests weergegeven.

Om het directe effect van de training te kunnen bepalen zijn op de pre- en posttestscores van de experimentele en de controlegroepen op de tests Exclusie-T en Analogieën-T variantieanalyses met herhaalde meting uitgevoerd met Conditie en Tijdstip van Afname als onafhankelijke variabelen en de resultaten op de pre- en posttests als afhankelijke variabelen. De resultaten zijn geanalyseerd voor de hele groep en per schooltype.

Uit de resultaten voor de gehele onderzoeksgroep (basisschool- en MLK-leerlingen) bleek voor Exclusie-T sprake van significante effecten voor Conditie ($F=7.92$; $p=.008$) en Tijdstip van Afname ($F=30.86$; $p<.001$). Ook de interactie tussen Conditie en Tijdstip was significant ($F=7.59$; $p=.009$). Voor Analogieën-T was er voor de variabelen Tijdstip ($F=24.46$; $p<.001$) en de interactie tussen Con-

Tabel 1

Gemiddelde scores op pre- en posttest voor Exclusie-T, Analogieën-T (directe effect), Exclusie-R, Analogieën-R (horizontale transfer), Verbale Exclusie en Visuele Analogieën (verticale transfer) naar schooltype en experimentele conditie

School	Conditie	Tijd	Ex-T	An-T	Ex-R	An-R	VeEx	ViAn
Basis	Exp.	voor	15.0	14.2	34.5	7.4	15.2	6.4
		na	20.0	21.7	40.3	13.9	17.2	9.6
	Contr.	voor	14.1	16.7	32.7	8.2	16.9	7.5
		na	15.4	19.5	38.5	15.0	20.8	9.6
MLK	Exp.	voor	15.6	18.0	31.9	11.0	15.7	7.6
		na	20.1	23.6	39.2	17.0	18.0	8.0
	Contr.	voor	11.9	15.5	31.8	8.7	16.4	6.2
		na	13.8	15.9	37.3	12.3	17.7	7.1

ditie en Tijdstip van Afname sprake van significante effecten ($F=9.02$; $p=.005$). Op grond van de significante interactie-effecten tussen Conditie en Tijdstip van Afname kan voor beide directe effectmaten worden geconcludeerd dat leerlingen uit de experimentele condities als gevolg van de leerpotentieeltraining een significant grotere vooruitgang boekten dan leerlingen uit de controlecondities.

De resultaten van variantie-analyses met herhaalde meting per schooltype voor Exclusie-T en Analogieën-T staan vermeld in Tabel 2. De variabele Schooltype gaf geen significant hoofdeffect te zien. Het interactie-effect tussen Conditie en Tijdstip van Afname was voor de basisschoolleerlingen significant voor zowel Exclusie-T als Analogieën-T. Voor de MLK-leerlingen was er alleen een significant interactie-effect voor Analogieën-T. Indien de posttestscores van de experimentele leerlingen vergeleken worden met de scores van de controleleerlingen, blijken de basisschool- en MLK-leerlingen in eenzelfde mate van de LIR te profiteren. Er treedt, gezien het feit dat

schooltype geen significant effect oplevert, geen differentiatie op in trainingseffect tussen leerlingen uit het basisonderwijs en leerlingen uit het MLK-onderwijs.

Om een transfereffect van training te kunnen bepalen, werd onderscheid gemaakt tussen specifieke horizontale en verticale transfer. Allereerst komen de resultaten met betrekking tot horizontale transfer aan de orde. Om na te gaan of training een extra effect op de testcores heeft gehad in vergelijking tot alleen testherhaling, zijn wederom variantie-analyses met herhaalde meting uitgevoerd. Uit de analyses bleek voor de totale onderzoeksgroep ten aanzien van Tijdstip van Afname voor zowel Exclusie-R ($F=36.60$; $p<.001$) als Analogieën-R ($F=53.92$; $p<.001$) sprake van significante effecten. Voor de variabele Conditie en de interactie tussen Conditie en Tijdstip van Afname daarentegen was geen sprake van significante verschillen. Deze resultaten kwamen overeen met de resultaten van de analyses per schooltype (zie Tabel 3).

Variantie-analyses met herhaalde meting

Tabel 2

Resultaten van variantie-analyses met herhaalde meting met Experimentele Conditie, Tijdstip van Afname en Interactie tussen Experimentele Conditie en Tijdstip van Afname voor directe effecttaken (Exclusie-T en Analogieën-T) naar schooltype

Factor	Basisschoolleerlingen			MLK-leerlingen			
	F	df	Signif. F	F	df	Signif. F	
Ex-T	Exp. Cond.	1.63	1	ns	7.88	1	.012*
	Voor/Na	13.95	1	.002*	15.53	1	.001*
	Interactie EC x vn	4.81	1	.042*	2.56	1	ns
An-T	Exp. Cond.	.00	1	ns	4.43	1	.050*
	Voor/Na	20.49	1	.000*	6.25	1	.022*
	Interactie EC x vn	55.22	1	.054*	4.69	1	.044*

* significantieniveau $p=.05$

Tabel 3

Resultaten van variantie-analyses met herhaalde meting met Experimentele Conditie, Tijdstip van Afname en Interactie tussen Experimentele Conditie en Tijdstip van Afname naar schooltype voor horizontale transfertaken (Ex-R, An-R) en verticale transfertaken (VeEx, ViAn)

	Factor	Basisschoolleerlingen			MLK-leerlingen		
		F	df	Signif. F	F	df	Signif. F
Ex-R	Exp. Cond.	.63	1	ns	.14	1	ns
	Voor/Na	18.48	1	.000*	16.78	1	.001*
	Interactie Ec x vn	.00	1	ns	.33	1	ns
An-R	Exp. Cond.	.15	1	ns	1.55	1	ns
	Voor/Na	37.54	1	.000*	18.09	1	.000*
	Interactie Ec x vn	.02	1	ns	1.13	1	ns
VeEx	Exp. Cond.	1.28	1	ns	.01	1	ns
	Voor/Na	4.74	1	.043*	1.69	1	ns
	Interactie Ec x vn	.49	1	ns	.13	1	ns
ViAn	Exp. Cond.	.08	1	ns	1.02	1	ns
	Voor/Na	14.83	1	.001*	.76	1	ns
	Interactie Ec x vn	.64	1	ns	.11	1	ns

* significantieniveau $p=.05$

met Schooltype als variabele brachten geen effect voor schooltype aan het licht. Kortom, de posttestscores op de horizontale transfertaken hangen niet onvoorwaardelijk samen met het al dan niet ontvangen hebben van leerpotentieeltraining. Deze resultaten kunnen verklaard worden vanuit het feit dat zowel de experimentele als de controlegroepen voor Analogieën-R en Exclusie-R op de posttest significant hogere scores (geanalyseerd met 'paired' t-toetsen) behaalden dan op de pretest. Dit gold zowel voor de basisschool- als voor de MLK-leerlingen. Herhaalde testafname lijkt een (vrijwel) even groot effect te hebben op horizontale transfertaken als de metacognitieve leerpotentieeltraining.

Pre- en posttestscores op de taken Visuele Analogieën (vorm) en Verbale Exclusie (deelgeheel) geven informatie over verticale transfer. Voor de totale onderzoeksgroep lieten de resultaten van variantie-analyses met herhaalde meting met Conditie en Tijdstip van Afname als variabelen alleen een significant effect zien ten aanzien van Tijdstip van Afname. Dit gold voor zowel Verbale Exclusie ($F=6.20$; $p=.017$) als Visuele Analogieën ($F=9.94$; $p=.003$). De analyses per schooltype resulteerden voor de basisschoolleerlingen in eenzelfde beeld (zie Tabel 3). Voor de MLK-leerlingen kon echter geen enkel significant effect worden vastgesteld.

Een variantie-analyse met herhaalde meting met Schooltype liet een significant interactie-effect zien voor Schooltype en Tijdstip van Afname voor Visuele Analogieën ($F=4.03$; $p=.05$). 'Paired' t-toetsen onderschreven dit resultaat: alleen de basisschoolleerlingen ($p=.008$ voor de controlegroep en $p=.028$ voor de experimentele groep) vertoonden op de test Visuele Analogieën een significante verbetering van pre- naar posttest. De MLK-leerlingen behaalden op de posttest vergelijkbare scores als op de pretest.

2.2 Leerpotentieelmaten en moeilijkheidsgraad trainingen

In deze subparagraaf wordt nader ingegaan op de vraag of de oudere MLK-leerlingen minder, evenveel, dan wel meer hulp en trainingstijd nodig hadden dan de basisschoolleerlingen. Tevens komt de moeilijkheidsgraad van de LIR voor allochtone leerlingen aan de orde.

Tabel 4 bevat de hoeveelheid hulpstappen die de experimentele leerlingen gemiddeld genomen per deeltraining nodig hadden. Met behulp van variantie-analyses zijn de verschillen in gemiddeld aantal benodigde hulpstappen tussen beide schooltypen getoetst. Er werd geen significant verschil gevonden. De oudere, qua mentale leeftijd vergelijkbare, MLK-leerlingen hadden evenveel hulp nodig als de basisschoolleerlingen om het leercriterium te

Tabel 4

Hoeveelheid benodigde hulpstappen per deeltraining: gemiddelden (M) en standaarddeviaties (SD) per schooltype

Type training	Basisschoolleerlingen		MLK-leerlingen	
	M	SD	M	SD
Analogieën				
Deel Geheel	16.6	7.96	9.9	7.02
Functie	13.5	8.89	6.9	6.42
Gemengd	11.7	9.18	9.7	6.63
Exclusie				
Vorm	4.2	5.65	6.8	5.63
Positie	12.5	14.32	7.6	9.07
Gemengd	7.7	6.62	7.5	4.81

bereiken. De correlatie tussen de totale hoeveelheid hulpstappen voor de Analogie- en de Exclusietraining bedroeg .77 voor de totale onderzoeksgroep. De correlatie tussen beide hulpscores was hoger voor de basisschoolleerlingen (.89) dan voor de MLK-leerlingen (.60).

De Exclusietraining bedroeg in het totaal gemiddeld 75 minuten voor de basisschoolleerlingen en 72 minuten voor de MLK-leerlingen. Dit verschil bleek niet significant. De hoeveelheid benodigde trainingstijd voor de Analogietraining vertoonde eenzelfde beeld (77,5 minuten voor de basisschoolleerlingen en 76 minuten voor de MLK-leerlingen).

Het aantal trainingsopgaven, waarbij de leerlingen hulp nodig hadden, zegt iets over de relatieve moeilijkheidsgraad van de training. Tijdens de Analogietraining kregen de basisschoolleerlingen gemiddeld per deeltraining bij 6,3 en de MLK-leerlingen bij gemiddeld 4,6 (van de 15) opgaven hulp aangeboden. Dit verschil bleek niet significant te zijn ($p=.215$). Bij de Exclusietraining lagen deze gemiddelden nog dichter bij elkaar: 3,8 en 3,7 (van de 18) opgaven hulp per deeltraining voor respectievelijk de basisschool- en MLK-leerlingen ($p=.892$). De LIR blijkt voor de onderzochte MLK-leerlingen niet moeilijker te zijn dan voor de jongere basisschoolleerlingen.

Nagegaan is ook in hoeverre er per schooltype verschillen in moeilijkheidsgraad bestaan tussen de (verbale) Analogietraining en de (niet-verbale) Exclusietraining. Voor de allochtone basisschoolleerlingen bleek de Analogietraining moeilijker dan de Exclusietraining: zij hadden per deeltraining bij gemiddeld 6,3 Analogie-opgaven en 3,8 Exclusie-opgaven hulp nodig ($p=.005$). Voor de allochtone MLK-

leerlingen was er geen significant verschil tussen de twee typen trainingen in het aantal trainingsopgaven waarbij zij hulp nodig hadden (per deeltraining 3,7 voor Exclusie-opgaven en 4,6 voor Analogie-opgaven). Ook op basis van de hoeveelheid benodigde hulpstappen bleek de Analogietraining moeilijker dan de Exclusietraining voor allochtone basisschoolleerlingen ($p<.001$), terwijl er bij de MLK-leerlingen geen verschil in moeilijkheidsgraad waarneembaar was.

2.3 Individuele verschillen in leerpotentieel

Bovenstaande resultaten hebben betrekking op groepsgewijze analyses en bieden geen inzicht in de resultaten van individuele leerlingen. In het kader van het gebruik van de leerpotentieeltest als diagnostisch instrument zijn tevens leerpotentieelgegevens van individuele leerlingen van belang. Inzake de hulpbehoefte per leerling bleken er binnen beide schooltypen grote individuele verschillen te bestaan (zie ook Tabel 5). Het aantal benodigde hulpstappen van de basisschoolleerlingen varieerde voor de Analogietraining van 8 tot 82 en voor de Exclusietraining van 1 tot 79. De MLK-leerlingen hadden voor de Analogietraining tussen de 8 en de 50 en voor de Exclusietraining tussen de 0 en de 45 hulpstappen nodig. Tabel 5 presenteert informatie over de individuele leerpotentieelprofielen van de leerlingen uit de experimentele conditie. Per leerling zijn de pre- en posttestscores op de directe effectmaten (Exclusie-T en Analogieën-T) gecombineerd met de hulpbehoefte.

Uit eerdere analyses is naar voren gekomen dat de basisschoolleerlingen als groep wat betreft prestaties op de posttests en de hoeveel-

Tabel 5

Individuele leerpotentieelprofielen: pre- en posttestscores directe effectmaten (Exclusie-T en Analogieën-T) gecombineerd met hulpbehoefte per leerling per schooltype

Leerling	Ex-T score pretest	Ex-T score posttest	hulp Excl.	An-T score pretest	An-T score posttest	hulp Anal.
BASISSCHOOLLEERLINGEN						
1	6	7	79	11	13	82
2	10	14	29	9	20	52
3	13	20	42	11	24	48
4	15	23	29	8	23	50
5	15	25	30	18	25	39
6	16	20	16	10	11	51
7	17	23	3	24	24	12
8	18	24	1	12	27	27
9	18	20	22	13	23	39
10	22	24	3	26	27	8
MLK-LEERLINGEN						
11	10	17	45	18	25	23
12	12	15	39	9	17	50
13	12	19	45	10	23	43
14	12	20	18	10	23	32
15	16	22	24	14	26	19
16	17	18	9	26	24	8
17	17	24	8	26	26	8
18	18	19	12	28	23	18
19	19	22	0	26	26	24
20	23	25	19	13	23	40

heid benodigde hulp, weinig verschillen vertoonden met de MLK-leerlingen. Uit Tabel 5 echter wordt duidelijk dat binnen beide schooltypen grote verschillen in leerpotentieelprofielen bestaan.

3 Conclusie

Levert de LIR bruikbare informatie op over het leerpotentieel van allochtone kinderen uit het basis- en MLK-onderwijs? Ter beantwoording van deze vraag worden de resultaten uit het onderhavige onderzoek vergeleken met de onderzoeksresultaten van Resing (1990).

Uit het onderhavige onderzoek kan geconcludeerd worden dat de prestaties van de leerlingen, die een metacognitieve leerpotentieeltraining kregen aangeboden, significant meer toegenomen zijn dan de prestaties van de controlegroepen op de taken die het directe effect van training meten. De resultaten met betrekking tot de horizontale transfertaken vertonen een ander beeld. Herhaalde testafname blijkt

een even groot effect op de testcores op Exclusie-R en Analogieën-R (horizontale transfertaken) te hebben als de metacognitieve training. Bij zowel de directe effectmaten als de horizontale transfertaken laten de MLK-leerlingen wat betreft hun vooruitgang in prestaties van pre- naar posttest een vrijwel identiek beeld zien als de basisschoolleerlingen. Wanneer de tijdens training geleerde strategieën toegepast moeten worden op taken met een afwijkende inhoud en een deels andere oplossingsprocedure (verticale transfer), scoren de basisschoolleerlingen op de test Visuele Analogieën hoger dan de MLK-leerlingen met een vergelijkbaar beginniveau. Dit effect van schooltype is onafhankelijk van het wel of geen training hebben gehad. Met betrekking tot Verbale Exclusie, de tweede verticale transfertaak, wordt geen effect van schooltype gevonden.

Bovenstaande resultaten komen gedeeltelijk overeen met de resultaten van Resing (1990). In haar onderzoek bleken de prestaties van de leerlingen die een metacognitieve training ontvingen, significant meer toegenomen te zijn

dan de prestaties van de controlegroepen op zowel de directe effecttaken als de specifieke horizontale transfertaken. Er was uitsluitend sprake van specifieke horizontale transfer en in heel beperkte mate van verticale transfer. In het onderzoek van Resing bleken deze effecten zowel op korte als op lange termijn aanwezig. Het ontbreken van een horizontaal trainingseffect in het onderhavige onderzoek kan het gevolg zijn van de beperkte groepsgrootte.

Het bekende probleem ten aanzien van het effect van transfer komt ook in beide onderzoeken naar voren: de leerlingen lijken veel van de training op te steken, maar zij passen het geleerde niet uit zichzelf toe als hun een andere, afwijkende taak wordt aangeboden. Opgemerkt dient te worden dat in dit onderzoek, evenals in dat van Resing, transfer op een statische wijze werd gemeten. De kinderen kregen zonder verdere verwijzingen naar de training de transfertaak voor zich en tijdens afname van de transfertaak werd hun geen hulp geboden. Onderzoekers die transfer op een dynamische wijze meten, rapporteren wel effecten van transfer bij kinderen (zie bijvoorbeeld Campione, Brown, Ferrara en Bryant, 1984).

In het onderzoek van Resing (1990) participeerden autochtone leerlingen uit het basisonderwijs, het onderwijs aan leerlingen met leren opvoedingsmoeilijkheden (LOM) en moeilijk lerende kinderen (MLK), tussen 7 en 8 jaar oud, in chronologisch equivalente leeftijdsgroepen¹. Uit dit onderzoek bleek dat de autochtone MLK-leerlingen drie keer zoveel hulp nodig hadden als de autochtone basisschoolleerlingen en de autochtone LOM-leerlingen twee keer zoveel hulp als de autochtone basisschoolleerlingen teneinde het leer criterium te bereiken.

In het onderhavige onderzoek zijn wat betreft de duur van de training en de hoeveelheid trainingsopgaven waarbij de kinderen hulp nodig hebben gehad, geen verschillen gevonden tussen allochtone basisschoolleerlingen en allochtone MLK-leerlingen met een vergelijkbaar beginniveau en mentale leeftijd. Daarnaast blijken de oudere allochtone MLK-leerlingen evenveel hulp nodig te hebben als de jongere allochtone basisschoolleerlingen om het leer criterium te bereiken. De allochtone leerlingen uit het onderhavige onderzoek blijken evenveel hulp nodig te hebben als de

autochtone basisschoolleerlingen uit het onderzoek van Resing. De totale trainingstijd die de allochtone leerlingen nodig hebben, ligt iets hoger dan die van de autochtone leerlingen. Dit komt mede doordat het (een aantal) allochtone leerlingen meer moeite en tijd kost hun antwoorden te verbaliseren dan autochtone leerlingen. Ook de instructie verloopt iets langzamer.

Geconcludeerd kan worden dat de allochtone basisschoolleerlingen eenzelfde leerpotentieelpatroon laten zien als de allochtone leerlingen uit het MLK-onderwijs die qua mentale leeftijd op een vergelijkbaar cognitief niveau functioneren maar twee tot drie jaar ouder zijn. Tevens kan worden geconcludeerd dat de onderzochte 7- tot 8-jarige allochtone basisschool- en de 9- tot 11-jarige allochtone MLK-leerlingen eenzelfde leerpotentieelpatroon laten zien als de 7- tot 8-jarige autochtone basisschoolleerlingen uit het onderzoek van Resing. Zij hebben evenveel hulp nodig en ook hun vooruitgang op de directe effecttaken is vergelijkbaar (zie ook Slenders & Resing, 1994).

Wie de resultaten ten aanzien van het directe en transfereffect nogmaals beschouwt, kan zich afvragen of een uitgebreid trainingsonderzoek in alle gevallen zinvol is of dat volstaan kan worden met herhaalde testafname of eventueel oefening met het testmateriaal. Een mogelijke verklaring voor het optredende test-herst effect (gemeten bij de controlegroepen) is dat in de testafname tegemoet is gekomen aan een aantal van de kritiekpunten ten aanzien van het gebruik van tests bij allochtone kinderen. Het is waarschijnlijk dat de allochtone leerlingen ten tijde van de posttest testvaardiger ('test-wiser') zijn dan bij de pretest. De gehele testsituatie en de gebruikte testmaterialen en -opgaven zijn hun ten tijde van de posttest bekend. De posttestscores reflecteren vermoedelijk de werkelijke vaardigheden van de leerlingen beter dan de pretestscores. Op de directe effectmaten is bovendien nog een sterk, individueel verschillend, effect van training waarneembaar.

De LIR levert daarnaast ook bruikbare, gestandaardiseerde informatie over het leervermogen van individuele leerlingen; deze is zowel kwantitatief als kwalitatief van aard. Informatie over individuele leerpotentieelprofielen, kwalitatieve informatie over bijvoorbeeld het

oplossingsproces (naar aanleiding van de verbalisaties) en de mogelijkheid tot uitgebreide observaties van werkhouding, motivatie en dergelijke op verschillende tijdstippen, maken de LIR tot een diagnostisch waardevol instrument. Het is niet alleen geschikt gebleken voor autochtone leerlingen uit het basis- en Speciaal Onderwijs in de leeftijd van 7 tot 8 jaar, maar ook voor allochtone leerlingen en voor MLK-leerlingen met een hogere chronologische, maar met een gelijke mentale, leeftijd. Om deze redenen is het extra investeren in de afname van een leerpotentieeltest gerechtvaardigd.

Noot

1. Als gevolg van de chronologisch equivalente onderzoeksgroepen ontstond er een ongelijke beginsituatie voor de drie groepen: de basisschoolleerlingen behaalden de hoogste scores op de pretest, de MLK-leerlingen de laagste. De spreiding in het (aanvangs)niveau van de leerlingen was zodanig groot dat het niet mogelijk was voor alle leerlingen om hun prestaties, ten gevolge van de training, optimaal te verbeteren. In het onderhavige onderzoek is dan ook gekozen voor basisschool- en MLK-leerlingen van gelijke mentale leeftijd.

Literatuur

- Babad, E.Y., & Budoff, M. (1974). Sensitivity and validity of learning potential measurement in three levels of ability. *Journal of Educational Psychology*, 66, 439-447.
- Bleichrodt, N., Resing, W.C.M., Drenth, P.J.D., & Zaal, J.N. (1987). *Revisie Amsterdamse Kinder Intelligentie Test*. Lisse: Swets & Zeitlinger.
- Bleichrodt, N., Drenth, P.J.D., Zaal, J.N., & Resing, W.C.M. (1984). *Revisie Amsterdamse Kinder Intelligentie Test. Handleiding*. Lisse: Swets & Zeitlinger.
- Budoff, M. (1987a). The validity of learning potential assessment. In C.S. Lidz (Ed.), *Dynamic assessment. An interactional approach to evaluating learning potential* (pp. 52-81). New York: The Guilford Press.
- Budoff, M. (1987b). Measures from assessing learning potential. In C.S. Lidz (Ed.), *Dynamic assessment. An interactional approach to evaluating learning potential* (pp. 173-195). New York: The Guilford Press.
- Campione, J.C., & Brown, A.L. (1987). Linking dynamic assessment with school achievement. In C.S. Lidz (Ed.), *Dynamic assessment. An interactional approach to evaluating learning potential* (pp. 82-115). New York: The Guilford Press.
- Campione, J.C., Brown, A.L., Ferrara, R.A., & Bryant, N.R. (1984). The zone of proximal development: Implications for individual differences and learning. In B. Rogoff & J.V. Wertsch (Eds.), *Children's learning in the zone of proximal development. New directions for Child development. No. 23* (pp. 77-91). San Francisco: Jossey-Bass, Inc.
- Campione, J.C., Brown, A.L., Ferrara, R.A., Jones, R.S., & Steinberg, E. (1985). Breakdowns in flexible use of information: Intelligence-related differences in transfer following equivalent learning performance. *Intelligence*, 9, 297-315.
- Feuerstein, R., Rand, Y., Jensen, M.R., Kaniell, S., & Tzuriel, D. (1987). Prerequisites for assessment of learning potential: The LPAD model. In C.S. Lidz (Ed.), *Dynamic assessment. An interactional approach to evaluating learning potential* (pp. 35-51). New York: The Guilford Press.
- Feuerstein, R., Rand, Y., & Hoffman, M.B. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments and techniques*. Baltimore: University Park Press.
- Hamers, J.H.M., Castelijns, J.H.M., & Pennings, A.H. (1989). De ontwikkeling van leertests. Ontstaan, onderzoek en toekomstbeeld. *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs*, 5, 66-75.
- Hamers, J.H.M., Hessels, M.G.P., & Luit, J.E.H. van (1991). *Leertest voor Etnische Minderheden: Test en handleiding*. Lisse: Swets & Zeitlinger.
- Hamers, J.H.M., & Ruijsenaars, A.J.J.M. (1982). Ontwikkelingen in de intelligentiediagnostiek. In D.B. Baarda & E.J. Zwaan (Eds.), *Alternatieven in de psychodiagnostiek* (pp. 115-150). Nijmegen: Dekker & van de Vegt.
- Hamers, J.H.M., & Ruijsenaars, A.J.J.M. (1984). *Leer-geschiktheid en leertests*. Nijmegen: Flevodruk Harlingen BV.
- Hamers, J.H.M., Sijtsma, K., & Ruijsenaars, A.J.J.M. (1993). *Learning potential assessment: theoretical, methodological and practical issues*. Amsterdam/Lisse: Swets & Zeitlinger.
- Hegarty, S., & Lucas, D. (1978). *Able to learn? The pursuit of culture-fair assessment*. Windsor: NFER Publishing Company.

- Hessels, M.G.P. (1993). *Leertest voor Etnische Minderheden: Theoretische en empirische verantwoording*. Rotterdam: RISBO.
- Leeuw, L. de, & Resing, W.C.M. (1985). Intelligentie-onderzoek op school. In H.P.M. Creemers, P.J.D. Drenth & D.B.P. Kallen (Eds.), *Losbladig onderwijskundig lexicon*, IO-2100-1/2100-22. Alphen aan den Rijn: Samsom.
- Lidz, C.S. (1987). *Dynamic assessment. An interactional approach to evaluating learning potential*. New York: The Guilford Press.
- Nickerson, R.S., Perkins, D.N., & Smith, E.E. (1985). *The teaching of thinking*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Resing, W.C.M. (1990). *Intelligentie en leerpotentieel: een onderzoek naar het leerpotentieel van jonge leerlingen uit het basis- en speciaal onderwijs*. Amsterdam/Lisse: Swets & Zeitlinger.
- Resing, W.C.M. (1993a). Measuring inductive reasoning skills: The construction of a learning potential test. In J.H.M. Hamers, K. Sijtsma & A.J.J.M. Ruijsenaars (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 219-242). Lisse: Swets & Zeitlinger.
- Resing, W.C.M. (1993b). Leerpotentieel en intelligentie: twee zijden van een medaille. In E.H. Rodigas & C.E.H.E. Veraart (Eds.), *Zorgbreedte* (pp. 30-38). Heerlen: OBD/CRO/PABO.
- Resing, W.C.M. (1997a). Leerpotentieel: wat is de meerwaarde? In T.J. Engelen-Snaterse & R. Kohnstamm (Eds.), *Trends in jeugd- en ontwikkelingspsychologie* (pp. 283-303). Lisse: Swets & Zeitlinger.
- Resing, W.C.M. (1997b). Intelligentie, leren en leerpotentieel. Meer van hetzelfde? In W. Tomic, H. van der Molen & W.K.B. Hofstee (Eds.), *Intelligentie en sociale competentie* (in druk). Lisse: Swets & Zeitlinger.
- Resing, W.C.M., Bleichrodt, N., & Drenth, P.J.D. (1986). Het gebruik van de RAKIT bij allochtoon etnische groepen. *Nederlands Tijdschrift voor de Psychologie*, 41, 179-188.
- Resing, W.C.M., & Wijk, A.M. van (1996). Leerpotentieel: onderzoek bij allochtone leerlingen uit het basisonderwijs. *Tijdschrift voor Orthopedagogiek*, 35, 432-444.
- Slenders, A.P.A.C., & Resing, W.C.M. (1994). *Het leerpotentieel van allochtone moeilijk lerende en basisschoolleerlingen*. Paper gepresenteerd op de Onderwijs Research Dagen, Utrecht.
- Snijders, J.Th., Tellegen, P.J., & Laros, J.A. (1988). *Snijders-Oomen niet-verbale intelligentietest: S.O.N.-R 5,5-17. Verantwoording en handleiding*. Groningen: Wolters-Noordhoff.
- Testscreeningscommissie (TSC) (1990). *Toepasbaarheid van psychologische tests bij allochtonen*. Utrecht: Landelijk Bureau Racismebestrijding (LBR-reeks, nr. 11).
- Teunissen, J.M.F.B.G., & Golhof, A. (1987). Etnische minderheden en speciaal onderwijs. In K. Doornbos & L.M. Stevens (Eds.), *De groei van het Speciaal Onderwijs. Deel A: Analyse van historie en onderzoek* (pp. 166-197). 's Gravenhage: Staatsuitgeverij.
- Wechsler, D. (1974). *Wechsler Intelligence Scale for Children - Revised*. New York: The Psychological Corporation.
- Wijnstra, J.M. (1987). *Op de grens van basis- en speciaal onderwijs: over verwijzing naar het speciaal onderwijs, in het bijzonder van leerlingen uit etnische en culturele minderheidsgroepen*. Amsterdam/Lisse: Swets & Zeitlinger.

Auteurs

A.P.A.C. Slenders is als Onderzoeker-in-Opleiding verbonden aan de Rijksuniversiteit Leiden bij de vakgroep Ontwikkelings- en Onderwijspsychologie.

Adres: Rijksuniversiteit Leiden, Postbus 9555, 2300 RB Leiden, e-mail: slenders@rulfsw.leidenuniv.nl

W.C.M. Resing is als universitair docent verbonden aan de Rijksuniversiteit Leiden bij de vakgroep Ontwikkelings- en Onderwijspsychologie.

Adres: Rijksuniversiteit Leiden, Postbus 9555, 2300 RB Leiden, e-mail: resing@rulfsw.leidenuniv.nl

Abstract

Research of learning potential with children of ethnic minority groups: the Learning Potential Test for Inductive Reasoning

A.P.A.C. Slenders & W.C.M. Resing. *Pedagogische Studiën*, 1997, 74, 183-196.

The central question in this article is whether the Learning Potential Test for Inductive Reasoning (LIR) provides useful information about the learning potential of children of various ethnic groups. Participants in this study are children from elementary schools and mentally retarded children. The children from primary education were 7 to 8 years old, the children from special education were 10 to 13 years old, but their mental level of performance before training was equal. The following has been concluded: both groups needed the same amount of hints during the learning potential training procedures. Their test scores after training increase substantial. These results are comparable with earlier research findings for Dutch children (Resing, 1990). The LIR provides diagnostic valuable information for children of ethnic minority groups.

Studiemotieven en leerstijlen van studenten in het hoger onderwijs voor ouderen*

E.C.M. De Ponti, J.G.L.C. Lodewijks en N. Rögels

Samenvatting

In deze studie is nagegaan, in welke opzichten oudere en jongere studenten in het Hoger Onderwijs van elkaar verschillen, wat betreft hun opvattingen over leren, hun motivationele oriëntaties, hun voorkeuren ten aanzien van de wijze van sturing van het leerproces en de studie-activiteiten die zij hierbij prefereren. Het studiegedrag van oudere studenten, die deelnemen aan het Hoger Onderwijs voor Ouderen (HOVO) werd onderzocht met behulp van onder meer de Inventaris Leerstijlen (ILS) voor het Hoger Onderwijs. Ter vergelijking dienden ILS-profielen van eerstejaars studenten van verschillende opleidingen aan de Katholieke Universiteit Brabant.

De gevonden verschillen tussen HOVO- en 'reguliere' W.O.-studenten lagen in sommige gevallen voor de hand, bijvoorbeeld dat HOVO-studenten zich in hun studie veel minder laten leiden door het streven naar het behalen van certificaten, dat zij hun studie veel minder positioneren in het perspectief van een beroep, en dat zij minder gericht zijn op het verwerven van kennis en vaardigheden met het oogmerk om daar later gebruik van te maken.

HOVO-cursisten richten zich daarnaast meer op een diepere verwerking van de leerstof, stellen een sterkere externe sturing op prijs, studeren vooral uitgaande van persoonlijke interesses en beschouwen hun leren vooral in het perspectief van het opbouwen van persoonlijke kennis (eigen kennisconstructie).

1 Levenslang leren

Als gevolg van de vele en snelle veranderingen die zich in de dagelijkse omgeving van mensen voordoen, worden aan het individu, als ouder, als burger en als werkende mens, bij voortdurende nieuwe eisen gesteld. Aanhoudende veranderingen in de sociale, de culturele, de technologische en de economische werkelijkheid vereisen van het moderne individu attente en flexibele aanpassingen aan snel veranderende maatschappelijke verwachtingen. Voor hedendaagse mensen is het welhaast noodzakelijk geworden om zich steeds weer op andere manieren te leren gedragen en zodoende greep te houden op het leven en op de zaken die zich in de samenleving voordoen. De startopleiding, die – eerder als een *passé-partout* – toegang gaf tot de wereld van werk en volwassenen, wordt – hoe hoog die ook eindigt – in het algemeen niet langer meer toereikend gevonden om bij leerlingen en studenten te voorzien in de kennis en vaardigheden die voor de rest van het leven nodig zijn (Lodewijks, 1993).

Het groeiende besef van de noodzaak tot levenslang leren is allerwegen doorgedrongen en maakt het begrijpelijk, dat ook binnen de wetenschap de laatste jaren de aandacht voor de ontwikkeling van het zogenaamde leervermogen (Shuell, 1988; Simons, 1990), voor het 'leren leren' (Maclure & Davies, 1991) en voor het bij voortdurende leren (Candy, 1991; Tuijnman & Van der Kamp, 1992) belangrijke aandachtspunten zijn geworden.

Scholingsbehoefte en scholingsbereidheid

Maar hoe is het gesteld met de behoefte van de mensen zelf voor wie de druk om je leven lang te leren steeds toeneemt? En hoe geven zij, mocht die behoefte aan verdergaande scholing al bestaan, hier gevolg aan? Een begin van een antwoord op deze vragen treffen we aan in een rapport van het Sociaal & Cultureel Planbureau

* Het hier besproken onderzoek was onderdeel van de doctoraalscriptie van de eerste auteur in het kader van haar studie in de Onderwijs- & Opleidingspsychologie aan de KU Brabant (De Ponti, 1995).

(SCP), waarin verslag wordt gedaan van een studie naar de opvattingen van Nederlanders over hun onderwijs en opleiding (Bronneman-Helmers, 1993). Uit dit rapport komt naar voren dat maar liefst een kwart van alle Nederlanders vindt dat zij te weinig hebben geleerd om hun dagelijkse werk goed te kunnen verrichten. Vooral mensen uit de lagere inkomensgroepen, werklozen en laag opgeleiden noemen het een voelbare handicap dat zij onvoldoende hebben geleerd.

De grote behoefte die er kennelijk bij veel Nederlanders bestaat om verder te leren of aanvullende opleidingen te volgen, blijkt echter merkwaardig genoeg niet te leiden tot een grote neiging tot zelf-geïnitieerde massale bijscholing, noch tot een manifeste opleidingsbereidheid. Integendeel, uit hetzelfde onderzoek van het SCP komt het beeld naar voren, dat mensen, zeker de al wat ouderen, en de lager opgeleiden bepaald niet staan te dringen om zich verder te doen opleiden. En dat er dus maatregelen zullen moeten worden bedacht om deze mensen weer te motiveren om voluit deel te nemen aan het levenslange leerproces.

Scholingsbehoefte bij ouderen

In tegenstelling tot de bevindingen in het rapport van het SCP, is de deelname van ouderen aan speciale vormen van Hoger Academisch Onderwijs de laatste jaren echter sterk stijgend. Het gaat hierbij om onderwijsvoorzieningen, waarvan men op latere leeftijd gebruik kan maken, meestal na afsluiting van deelname aan het arbeidsproces, of in aansluiting op een grotendeels voltooide maatschappelijke carrière.

Waarom bij juist deze ouderen de belangstelling voor studeren sterk toeneemt, is niet zonder meer duidelijk. Wel is er een aantal tentatieve verklaringen voorgesteld, die verband houden met de demografische ontwikkeling van de samenleving (vergrijzing).

Het procentuele aandeel van ouderen in de samenstelling van de bevolking is in de afgelopen jaren sterk gegroeid. Was in het jaar 1990 nog ongeveer één op de acht Nederlanders 65 jaar of ouder, in 2020 zal één op de vijf Nederlanders tot die leeftijdscategorie behoren. In datzelfde jaar zal een derde deel van de bevolking 55 jaar of ouder zijn (Herwijer, 1990).

Een tweede verklaring is dat mensen steeds eerder ophouden met werken, op een moment

dat zij nog vitaal zijn en de behoefte aan maatschappelijke zorg nog niet aanwezig is. Dit leidt ertoe dat er in het leven van mensen "een periode bij gekomen is" tussen 55 en 75 jaar, waardoor er een nieuwe groep vitale ouderen ontstaat die zelfstandig is en die goed voor zichzelf kan zorgen (Ter Steege, 1989).

Voorts is er sprake van een toename van het gemiddelde opleidingsniveau van de bevolking. Het percentage ouderen met een (semi-) hogere opleiding zal naar verwachting sterk toenemen. De behoefte aan verder onderwijs en verdere ontplooiing op latere leeftijd wordt hierdoor sterk beïnvloed.

Leerstijlen en leervermogen: kwaliteit en kwantiteit

Uit recent onderzoek blijkt, dat kwantitatieve verschillen in leervermogen en intelligentie tussen ouderen en jongeren veel kleiner zijn dan veelal wordt aangenomen (Herwijer, 1990; Cramer & Van der Kamp, 1990). De betekenis van het biologische, of deficiëntie-model (met het toenemen van de leeftijd verminderen allerlei vermogens, zo ook de leervermogens) dient op grond van deze inzichten te worden genuanceerd. Verschillen in leervermogen tussen ouderen en jongeren kunnen wellicht ook beter niet in kwantitatieve (meer-minder), maar veel eerder in kwalitatieve (hoe) termen worden beschouwd. Hierbij kan worden uitgegaan van het alternatieve competentie-model, waarin de specifieke kwaliteiten van ouderen worden onderkend en het begrip autonomie een centrale rol krijgt toebedeeld (Linschoten, Gerritsen, Schaap & Wolffensperger, 1989)

Over die *kwalitatieve* verschillen in het leren van oudere in vergelijking tot jongere studenten is overigens merkwaardig genoeg betrekkelijk weinig empirisch materiaal voorhanden. In dit artikel worden resultaten gepresenteerd van een onderzoek, waarin kwalitatieve verschillen in het leren/studeren van oudere en jongere studenten worden beschreven. Het onderzoek is met name opgezet om aanknopingspunten te vinden voor de verbetering van het Hoger Onderwijs aan senioren. Zo zou de inhoud en de structuur van dit onderwijs beter kunnen aansluiten bij de motivationele oriëntaties, de leerconcepties en leerstijlen van de oudere studenten. In deze studie zijn de leerstijlen van vijftig-plussers in kaart gebracht,

die een studie volgen aan een van de Senioren Academies in ons land. Hierbij is nagegaan in welke opzichten deze 'senior-studenten' verschillen van jongere studenten in het Hoger Onderwijs.

Leerstijlen van studenten in het hoger onderwijs

Onder een *leerstijl* wordt in het algemeen verstaan, de wijze waarop een lerende gewoonlijk (dat wil zeggen in uiteenlopende leersituaties) leert. In onze opvatting is een leerstijl een persoonsgebonden, relatief vaste combinatie van leer- en denkactiviteiten, die een lerende doorgaans toepast wanneer hij/zij aan het leren is.

De term leerstijl gebruiken wij overigens met enige aarzeling, omdat met het gebruik van dit begrip al te gemakkelijk de suggestie wordt opgeroepen, als zouden er verschillende, relatief onveranderbare, soorten of typen studenten zijn te onderscheiden. Naar ons inzicht is er naast de evidentie, dat er bij het leren in bepaalde (onderwijs)situaties vaste en persoonsgebonden patronen van leer- en denkactiviteiten voorkomen (bijvoorbeeld Vermunt, 1992), voldoende bewijs, dat lerenden hun gedragspatronen ook flexibel en uit zichzelf kunnen aanpassen, aan de eisen die verschillende omstandigheden naar hun oordeel stellen (bv.

Vermetten, Vermunt & Lodewijks, 1995).

In ons onderzoek sloten wij aan bij de 'leerstijl-theorie' zoals die door Vermunt (1992) is beschreven en die is gebaseerd op onderzoek bij studenten in het Wetenschappelijk Onderwijs. Volgens Vermunt dient een leerstijl breder te worden gedefinieerd, dan wij hierboven deden, en wel als een karakteristieke samenhang tussen iemands: (a) motivationele oriëntatie, (b) diens opvattingen over wat leren is, (c) zijn/haar voorkeuren voor de wijze van sturing van het leerproces (i.c. extern, vanuit het onderwijs, of intern, vanuit de student zelf) en (d) de kenmerkende verwerkingsactiviteiten die de persoon bij het leren toepast.

Op deze wijze gedefinieerd vond Vermunt, dat er onder W.O.-studenten vier verschillende soorten leerstijlen zijn te onderscheiden: (a) de betekenisgerichte, (b) de reproductiegerichte, (c) de toepassingsgerichte en (d) de ongerichte (of stuurloze) leerstijl. Deze vier leerstijlen zijn in Figuur 1 kort getypeerd.

Voor onze analyse is het op deze plaats voldoende om kort stil te staan bij de eerste twee leerstijlen. Studenten met een *betekenisgerichte leerstijl* geven zoveel mogelijk zelf vorm aan hun leren, zij leren vanuit een perspectief van persoonlijke interesse, en zijn erop gericht om

LEERSTIJLEN bij studenten in het Hoger Onderwijs

1 BETEKENISGERICHTE LEERSTIJL

Leer- en denkactiviteiten: relaterende en kritische verwerking
Regulatie: zelf-gestuurde regulatie
Leerconceptie: eigen kennisconstructie
Motivationele oriëntatie: persoonlijke interesse

2 REPRODUCTIEGERICHTE LEERSTIJL

Leer- en denkactiviteiten: memoriserende en analytische verwerking
Regulatie: extern-gestuurde regulatie
Leerconceptie: leren als het opnemen van aangeboden kennis
Motivationele oriëntatie: certificaat- en testgericht

3 TOEPASSINGSGERICHTE LEERSTIJL

Leer- en denkactiviteiten: concrete verwerking
Regulatie: zelf- én extern-gestuurde regulatie
Leerconceptie: leren om te gebruiken, toe te passen, om het nut
Motivationele oriëntatie: beroepsgericht

4 ONGERICHTE LEERSTIJL (STUURLOOS)

Leer- en denkactiviteiten: stuurloos leergedrag
Regulatie: ongereguleerde of ambivalente sturing
Leerconceptie: veel waarde aan sociale contacten en samenwerking met anderen en stimulerend onderwijs
Motivationele oriëntatie: ambivalent

Figuur 1. Leerstijlen bij studenten in het Hoger Onderwijs in termen van motivationele oriëntaties, leerconcepties, geprefereerde regulatie en geprefereerde verwerkingsactiviteiten (naar: Vermunt, 1992)

leer- en denkactiviteiten te ontplooiën, die leiden tot een dieper inzicht in de materie en tot persoonlijk geïntegreerde kennisconstructies. Hun opvatting van leren is, dat de informatie die geleerd moet worden, geïntegreerd moet raken binnen het persoonlijke kennisbestand.

Het leren van deze studenten wordt gekenmerkt door het toepassen van vooral relaterende en kritische verwerkingsactiviteiten. Dit wil zeggen dat zij er bij hun leren vooral op gericht zijn om (a) verbanden te zoeken tussen verschillende onderdelen van de leerinhoud, tussen de onderdelen en het geheel, de grote lijn van de stof en tussen nieuwe informatie en de eigen voorkennis of pre-concepties en (b) om mee te denken met auteurs, docenten en medeleerlingen, om een eigen mening te beredeneren en om op grond hiervan 'iets aan de informatie te veranderen', d.w.z. de stof niet letterlijk te nemen, maar deze te verbinden met het eigen denken en eigen inzichten.

Studenten met een *reproductiegerichte leerstijl* leren vooral met het oog op de toets die zij over hun leren moeten afleggen, of met het oog op het diploma, dat in het verschiet ligt. Hun leer- en denkactiviteiten zijn er vooral op gericht om de leerstof zó te structureren en te bewerken, dat deze gemakkelijker kan worden opgenomen. Hun idee van leren bestaat uit het in gecondenseerde en gememoriseerde vorm opnemen van kennis, zoveel mogelijk overeenkomstig de vorm, waarin die wordt aangeboden. Het leren van dit 'type' student wordt gekenmerkt door het toepassen van memoriserende en analytische verwerkingsactiviteiten. Het leren wordt gedomineerd door activiteiten als (a) het inprenten van afzonderlijke stukken informatie, door deze bijvoorbeeld een aantal malen voor zichzelf te repeteren, en door het herhalen van de studiestof en verder door activiteiten als (b) het opsplitsen van de studiestof in de onderdelen, waaruit deze is samengesteld en door het stap voor stap (meestal lineair) uitzoeken, hoeveel en welke verschillende aspecten aan een probleem, gedachtengang of theorie te onderkennen zijn.

Het zal duidelijk zijn, dat het leren van de op betekenis gerichte student, zoals Vermunt die heeft beschreven, naar aard en niveau sterk afwijkt van de reproductiegerichte student. Ook nadere analyses laten zien, dat de beide 'typen' niet alleen sterk verschillende leeracti-

viteiten ontwikkelen, maar ook, dat de uitkomsten van hun leren er kwalitatief anders uitzien. Noij en Lodewijks (1993) en Lodewijks en Noij (1994) vonden bijvoorbeeld, dat studenten met een betekenisgerichte leerstijl de leerstof 'dieper' verwerken en hogere vormen van begrip van de stof opbouwen (zij 'doorgronden' de stof en integreren de informatie beter in hun kennisbestand), dan studenten met een reproductiegerichte leerstijl (wier kennis in het algemeen beperkt blijft tot het 'weten', of het 'snappen' waar de stof over gaat). Het meest opvallende verschil is wel, dat de studenten die 'betekenisgericht' studeren, aanzienlijk meer *integrerende* en *constructieve leeractiviteiten* verrichten, dan reproductiegerichte studenten. Het gaat bij dit laatste om al die leeractiviteiten, waarbij de student de betekenis van het geleerde pogt te construeren op basis van het leer materiaal, door min of meer voortdurend met het materiaal te *interacteren*, door de inhoud ervan stelselmatig te *verbinden* met wat men al weet, wat men er zich persoonlijk bij kan voorstellen, wat men, in welke gevallen met die kennis zou kunnen doen, wat men zelf vindt over die kennis etc. Voorts blijkt uit onderzoek (bijvoorbeeld Spiro, Feltovich, Jacobson & Coulson, 1991), dat leerprocessen, die gekenmerkt worden door het (inter-)actief construeren en integreren van informatie en kennis, tot diepere en flexibeler inzichten, en dus tot hoogwaardiger leerresultaten, leiden. Overigens vond ook Vermunt (1992) in zijn onderzoek naar leerstijlen al, dat studenten met een betekenisgerichte leerstijl, onafhankelijk van de studie die zij volgden, in het algemeen betere studieresultaten boekten, dan studenten met een van de drie andere leerstijlen.

2 Probleemstelling

In het onderhavige onderzoek is nagegaan, of en in welke opzichten er verschillen bestaan in leerstijlen tussen oudere (senior) en jongere (reguliere) studenten in het Hoger Onderwijs. Meer in het bijzonder is onderzocht of, en zo ja, welke kwalitatieve verschillen er bestaan tussen beide groepen wat betreft (a) hun opvattingen over wat leren is (leerconcepties), (b) de motivationele oriëntaties die ten grondslag liggen aan het studiedrag, (c) de studie- en leer-

activiteiten waaraan de studenten de voorkeur geven en (d) de soort van sturing (intern versus extern) die studenten bij het leren prefereren.

Het studiegedrag van oudere studenten werd onderzocht bij cursisten die deelnemen aan het Hoger Onderwijs voor Ouderen (HOVO), waarbij gebruik werd gemaakt van onder meer de *Inventaris Leerstijlen (ILS)* voor het Hoger Onderwijs (Vermunt & Van Rijswijk, 1988). Deze gegevens werden vergeleken met 'ILS-profielen' van eerstejaars studenten afkomstig van verschillende opleidingen aan de Katholieke Universiteit Brabant.

Het onderzoek had een tweeledig doel: het eerste oogmerk was vast te stellen, in hoeverre de 'leerstijl-theorie' van Vermunt (1992) generaliseerbaar is naar andere populaties studenten binnen het Hoger Onderwijs; ten tweede ging het er ons om aanknopingspunten te vinden voor een optimale inrichting van het hoger onderwijs ten behoeve van senior-studenten.

3 Opzet van het onderzoek

Onderzoeksgroep

Bij drie instellingen van Hoger Onderwijs Voor Ouderen (HOVO) werden cursisten groepsgewijs benaderd met de vraag of zij mee wilden werken aan het onderzoek. De cursisten volgden allen onderwijs op het gebied van de psychologie (in Groningen: Humanistische Psychotherapiën en Ontwikkelingspsychologie, in Tilburg: Ontwikkelingspsychologie en Topics uit de Sociaal Cognitieve Psychologie), of op het gebied van wijsbegeerte (in Rotterdam: De Mythen van de Reconquista en Wijsgerige Antropologie). Van de in totaal 130 aldus benaderde cursisten reageerden er 88 (=68%) positief. De vragenlijsten zijn vervolgens thuis ingevuld, in een volgende bijeenkomst werden ze weer opgehaald. Een aantal achtergrondkenmerken van de onderzochte groep staat vermeld in Tabel 1 (zie verder).

Deelname aan het onderzoek was vrijwillig. Wel kregen alle deelnemers na afloop van het onderzoek hun persoonlijk leerstijlenprofiel uitgereikt.

Meetinstrumenten

Voor het verzamelen van gegevens is gekozen voor de *Inventaris leerstijlen (ILS)* voor het

hoger onderwijs van Vermunt en Van Rijswijk (1988). Deze vragenlijst bevat 120 uitspraken over leren en studeren, verdeeld over deel A: studie-activiteiten, en deel B: studiemotieven en studie-opvattingen. De respondenten moesten bij elke vraag op een vijfpuntsschaal aangeven in hoeverre zij zichzelf in de betreffende uitspraak herkennen.

In een toegevoegde *vragenlijst* is een aantal *persoonlijke gegevens* gevraagd, te weten geslacht, leeftijd, burgerlijke staat, opleidingsniveau, en werd een aantal vragen gesteld over deelname aan en keuze voor het HOVO. Voor dit laatste is gebruik gemaakt van de vragen die eerder reeds in een 'typeringsonderzoek' onder HOVO-cursisten te Groningen werden gesteld door een projectgroep gerontologie-studenten (Tempelman, 1988).

Analyses

Over de ILS-gegevens van HOVO-cursisten werden allereerst betrouwbaarheids- en itemanalyses verricht, vervolgens werden hierover factoranalyses uitgevoerd. Op grond hiervan is nagegaan of de gevonden patronen in de gegevens overeenkwamen met de patronen (leerstijlen) die Vermunt (1992) eerder in het Wetenschappelijk Onderwijs heeft gevonden.

Ten slotte werden op basis van de HOVO-gegevens zogenaamde normgroepen gedefinieerd, die werden vergeleken met de reeds beschikbare normscores van eerstejaars studenten in het Wetenschappelijk Onderwijs.

4 Resultaten

Persoonlijke en achtergrondkenmerken

In Tabel 1 zijn de achtergrondkenmerken van de proefgroep samengevat. Hieruit blijkt onder meer, dat in de steekproef vrouwen enigszins in de meerderheid zijn met 58% tegenover 42% mannen.

Dit heeft verschillende oorzaken (Snijders, 1990): er zijn op deze leeftijd nu eenmaal meer vrouwen dan mannen; zij hebben gemiddeld een betere gezondheid, zij komen vroeger buiten het arbeids- of maatschappelijk proces te staan; en ten slotte hebben zij naar hun eigen oordeel in de regel 'meer in te halen' dan mannen.

Dit laatste wordt ook in 'objectieve' zin dui-

Tabel 1
Achtergrondkenmerken van de onderzochte cursisten (N=88)

KENMERK	GEM	%
GESLACHT		
* Man		42%
* Vrouw		58%
LEEFTIJD		
* Totale groep	63	
* Mannen	66	
* Vrouwen	62	
OPLEIDINGSNIVEAU		
• Totale groep		
Lager onderwijs/LBO/MAVO		11%
HBS/HAVO/VWO/MBO		26%
HBO		38%
Universiteit		24%
• Mannen		
Lager onderwijs/LBO/MAVO		9%
HBS/HAVO/VWO/MBO		23%
HBO		31%
Universiteit		37%
• Vrouwen		
Lager onderwijs/LBO/MAVO		12%
HBS/HAVO/VWO/MBO		29%
HBO		45%
Universiteit		14%

delijk uit het feit, dat mannen over het algemeen een hoger opleidingsniveau blijken te hebben dan vrouwen binnen het HOVO: van de mannen beschikt 83% over een hogere opleiding (VWO, HBO, Universiteit), bij vrouwen is dit 67%.

Wat betreft leeftijd zijn vrouwen in de groep vijftigers die deelnemen aan HOVO ruimer vertegenwoordigd dan mannen: 81% in deze leeftijdsgroep bestaat uit vrouwen. Mannen zijn in die leeftijdscategorie waarschijnlijk vaker dan vrouwen nog in het arbeidsproces actief en werken vaak tot ongeveer hun zestigste levensjaar door. Daarom vangen zij in het algemeen pas later aan met cursussen in het HOVO. In samenhang hiermee speelt bij vrouwen waarschijnlijk een rol, dat hun kinderen, wanneer zij rond de vijftig zijn, al wat ouder of reeds het huis uit zijn, waardoor zij al op iets jongere leeftijd dan mannen weer meer tijd voor zichzelf en een eventuele studie krijgen.

Motieven om een HOVO-opleiding te volgen
Op grond van de ILS-gegevens zijn wij nage-

gaan, welke motieven er bij cursisten zijn om deel te nemen aan een HOVO-opleiding.

Uit een aan de ILS toegevoegde vragenlijst (Tempelman, 1988), bleken de belangrijkste motieven van HOVO-cursisten de volgende: interesse en verdieping in een onderwerp en actief blijven (respectievelijk 90% en 76%). Dat blijkt ook de conclusie van Ventura en Worthy: "that there is no question that the predominant motivation for older adults enrolling in education programs is to learn" (in Merriam & Cunningham, 1989).

Maar liefst 90% van de cursisten uit de steekproef houdt er een 'academische' leeroriëntatie op na, hetgeen wil zeggen dat zij graag studeren om zich verder te verdiepen en bekwamen in een of meerdere onderwerpen. Opvallend is dat het behalen van een certificaat door slechts 7% van de cursisten wordt genoemd; op grond hiervan lijkt het dan ook weinig zinvol om al teveel tijd te steken in het opzetten van certificaat-studies in het HOVO.

Motieven in relatie tot de arbeidsmarkt spelen - zoals nogal voor de hand ligt - bij HOVO-cursisten nauwelijks een rol van betekenis; het door hen gevolgde onderwijs is niet bedoeld als voorbereiding op een beroep, daarbij komt dat alle cursisten de vijftig gepasseerd zijn en dat de meesten hun beroeps- en maatschappelijke leven al hebben afgesloten.

Vervolgens zijn de factoren onderzocht die de deelname aan het Hoger Onderwijs Voor Ouderen beïnvloeden. De genoten vooropleiding kwam hierbij als belangrijkste factor naar voren: voor 70% van de cursisten had dit een stimulerende werking. Dit komt overeen met bevindingen uit eerder onderzoek (Herwijer, 1990). Dagelijkse activiteiten en beslommingen, zoals kinderen, leeftijd, partner en gezondheidssituatie oefenen slechts een geringe invloed uit. De overige factoren die in de aanvullende vragenlijst aan bod kwamen, zoals vervoersmogelijkheden, huisgenoten, financiële situatie en afstand, zijn voor de meeste cursisten niet of nauwelijks van invloed op hun deelname.

Psychometrische analyses van de ILS-vragenlijst
Over de ILS-gegevens werden per subschaal van de vragenlijst betrouwbaarheidsanalyses (interne consistentie) uitgevoerd. De resultaten hiervan zijn bevredigend. Op één na hebben

alle schalen een Cronbach $\alpha \geq .60$, hetgeen kan worden beschouwd als een indicatie voor een redelijke tot goede betrouwbaarheid van de ILS-(sub)schalen. Er is één uitzondering: de lage interne consistentie ($\alpha = .31$) voor de schaal persoonlijke interesse.

Factorstructuur ILS-data

Over de ILS-gegevens werd een factoranalyse (hoofdc componenten-analyse met Varimax-rotatie) uitgevoerd: allereerst om de factorstructuur van de ILS-gegevens van HOVO-studenten te kunnen vaststellen en deze, vervolgens te kunnen vergelijken met de gevonden structuren bij studenten in het reguliere wetenschappelijke onderwijs. De resultaten zijn samengevat in Tabel 2. In deze tabel zijn de relaties tussen de ILS-schalen weergegeven in een 4-factoren model.

Wanneer men in Tabel 2 de ladingen bekijkt

op de verschillende factoren, ziet men dat een aantal schalen op meerdere factoren laadt. Dit geeft de indruk dat de vier leerstijlen elkaar soms overlappen; ondanks dat op een aantal aspecten de leerstijlen wel herkenbaar zijn, mag niet zonder meer geconcludeerd worden dat de leerstijlen ook als zodanig kunnen worden onderscheiden.

Van de vier factoren die binnen het HOVO werden aangetroffen, komen met name de eerste twee overeen met de factoren die Vermunt heeft gevonden, te weten betekenisgericht en reproductiegericht (Vermunt, 1992). Toch zijn er ook hier enkele verschillen aan te geven, hetgeen is af te lezen uit de Tabellen 3 en 4, waarbij de ladingen op deze twee factoren van studenten in het reguliere Hoger Onderwijs enerzijds, en HOVO-cursisten anderzijds, met elkaar vergeleken worden.

Een tweetal voorbeelden van afwijkende

Tabel 2

Factorladingen van ILS-schalen in een 4-factoren Varimax-oplossing (hoofdc componentenanalyse, bij HOVO-cursisten (N=88))

ILS-schalen	F1	F2	F3	F4
Verwerkingsstrategieën				
Diepteverwerking				
Relateren en structureren	.88			
Kritisch verwerken	.76			
Stapsgewijze verwerking				
Memoriseren en herhalen	.34	.64		
Analyseren	.59	.43		.29
Concrete verwerking	.65		.28	.34
Regulatiestrategieën				
Zelfsturing				
Leerproces en -resultaat	.88			
Leerinhoud	.67		-.39	
Externe sturing				
Leerproces		.78		.26
Leerresultaat		.76		.30
Stuurloos			.61	
Leermotieven				
Certificaatgericht		.27		.79
Beroepsgericht				.76
Testgericht				.65
Persoonlijke interesse		.48		
Ambivalent	-.43	-.41		.44
Leeropvattingen				
Opname van kennis		.60	.56	.25
Opbouw van kennis	.71	.28	.30	
Gebruik van kennis	.31		.50	.52
Stimulerend onderwijs		.43	.71	
Samen studeren			.71	

NB: ladingen tussen -.25 en +.25 zijn niet opgenomen

Tabel 3

Vergelijking van factorladingen op de eerste factor (betekenisgericht) tussen studenten in het reguliere Hoger Onderwijs en HOVO-cursisten

HOGER ONDERWIJS		HOVO	
(Gebaseerd op gemiddelde scores van alle W.O.-KUB studenten uit het onderzoek van Vermunt, 1992)			
Relateren en structureren	.73	Relateren en structureren	.88
Kritisch verwerken	.71	Kritisch verwerken	.76
		Memoriseren en herhalen	.34
		Analyseren	.59
Concrete verwerking	.67	Concrete verwerking	.65
Zelfsturing leerproces en -resultaat	.74	Zelfsturing leerproces en -resultaat	.88
Zelfsturing leerinhoud	.71	Zelfsturing leerinhoud	.67
Certificaatgericht	-.40		
Persoonlijk geïnteresseerd	.52		
Ambivalent	-.26	Ambivalent	-.43
Opname van kennis	-.38		
Opbouw van kennis	.74	Opbouw van kennis	.71
		Gebruik van kennis	.31

Factor 1: Betekenisgerichte oriëntatie

combinaties bij HOVO-cursisten: op de eerste factor (Vermunt noemt dit de betekenisgerichte leerstijl), laden naast relateren en structureren en kritisch verwerken ook de andere drie verwerkingsstrategieën memoriseren en herhalen, analyseren en concrete verwerking. De

schaal persoonlijk geïnteresseerd ontbreekt bij het HOVO, terwijl de negatieve ladingen van de schalen certificaat-gericht en opname van kennis zich evenmin bij HOVO-cursisten manifesteren.

Op de tweede factor (bij Vermunt: de repro-

Tabel 4

Vergelijking van factorladingen op de tweede factor (reproductiegericht) tussen studenten in het reguliere Hoger Onderwijs en HOVO-cursisten

HOGER ONDERWIJS		HOVO	
(Gebaseerd op gemiddelde scores van alle W.O.-KUB studenten uit het onderzoek van Vermunt, 1992)			
Memoriseren en herhalen	.72	Memoriseren en herhalen	.64
Analyseren	.73	Analyseren	.43
Externe sturing leerproces	.74	Externe sturing leerproces	.78
Externe sturing leerresultaat	.71	Externe sturing leerresultaat	.76
Certificaatgericht	.46	Certificaatgericht	.27
		Persoonlijk geïnteresseerd	.48
		Ambivalent	-.41
Opname van kennis	.60	Opname van kennis	.60
		Opbouw van kennis	.28
		Stimulerend onderwijs	.43

Factor 2: Reproductiegerichte oriëntatie

ductiegerichte leerstijl), laden bij de studieopvattingen niet alleen opname van kennis, maar ook opbouw van kennis en stimulerend onderwijs. De schaal persoonlijk geïnteresseerd laadt bij HOVO-cursisten wèl op deze factor. Dit zou erop kunnen wijzen dat ouderen die persoonlijk geïnteresseerd zijn eerder overgaan tot het reproduceren van kennis, dan tot het opbouwen van kennis.

Wanneer we de gegevens bezien en de klassenmiddens inspecteren in het licht van de door Vermunt onderscheiden categorieën, dan valt het volgende op: Ten eerste scoren HOVO-cursisten hoger op de schalen diepteverwerking (relaties zoeken tussen onderdelen van de studiestof en de grote lijn proberen te volgen), persoonlijke interesse (studeren uit belangstelling voor bepaalde onderwerpen en om zich persoonlijk te ontwikkelen) en opbouw van kennis (verschillende studie-activiteiten, waaronder vooral relaterende en kritische verwerkingsactiviteiten, gebruiken om tot kennis en inzicht te komen). Deze scorepatronen hangen nauw samen met een betekenisgerichte leerstijl: HOVO-cursisten vertonen meer kenmerken van deze stijl dan reguliere studenten in het W.O. Een eerste verklaring hiervoor zou kunnen zijn, dat HOVO-cursisten in het algemeen een hogere afgeronde vooropleiding én meer studie-ervaring hebben.

Voor wat betreft persoonlijke interesse komt daar nog bij, dat de scores op deze schaal in vergelijking tot W.O.-studenten hoog zijn voor alle HOVO-cursisten. Hierbij moet echter wederom worden opgemerkt dat de interne consistentie van de ILS-items op met name deze schaal bij HOVO-cursisten relatief laag was. De toegevoegde vragenlijst met de biografische gegevens wees echter ook persoonlijke interesse aan als de belangrijkste reden om voor een bepaalde cursus bij een HOVO-instelling te kiezen.

Voorts blijken HOVO-cursisten vergelijkenderwijs meer behoefte aan externe sturing te hebben (= het zich laten sturen bij hun studie door aanwijzingen van docenten of in boeken). Dit geldt met name voor de hoger opgeleide HOVO-cursisten: zij scoren significant hoger op voorkeur voor externe sturing dan middelbaar en lager opgeleide cursisten ($t \geq -2.03 \leq 2.03$, $\alpha \leq 0.05$)¹.

Tevens vinden HOVO-cursisten het belangrijker dan W.O.-studenten dat ze door docenten gestimuleerd worden om bepaalde studie-activiteiten te ontplooiën. Een mogelijke verklaring hiervoor zou kunnen zijn, dat de verwachtingen van HOVO-cursisten ten aanzien van goed onderwijs gevormd zijn door het vroegere onderwijs, dat zij hebben genoten: wellicht traden docenten destijds meer op de voorgrond en was de toenmalige manier van lesgeven meer sturend, waardoor cursisten ook nu weer meer externe sturing verwachten.

HOVO-cursisten scoren naar verwachting lager op de schalen: certificaat-gericht; beroepsgericht en gebruik van kennis. Tevens hechten HOVO-cursisten minder waarde aan het samen bestuderen van studiestof. Dit wordt mogelijk veroorzaakt doordat de vragen in de ILS vooral gericht zijn op het samen voorbereiden van tentamens en het samen maken van opdrachten. HOVO-cursisten zijn echter niet verplicht tentamens te doen en de lesstof voor te bereiden. Daarbij komt dat de verschillen tussen cursisten onderling op het gebied van interesse vrij groot zijn; elke cursist heeft zo zijn of haar eigen ideeën over welke delen van de stof belangstelling behoeven. Dit maakt samenwerking bij de studie er om praktische redenen niet eenvoudig of voor de hand liggend op, ten gevolge waarvan HOVO-cursisten wellicht vaker kiezen voor zelfstudie.

Ook al geven HOVO-cursisten op de ILS aan, dat zij nauwelijks meerwaarde zien in het samen studeren, uit de gegevens van de aanvullende vragenlijst komt echter wel het beeld naar voren, dat HOVO-cursisten veel waarde toekennen aan het gezamenlijk volgen van onderwijs; sociale contacten zijn voor de meesten wel degelijk een belangrijke drijfveer om aan HOVO-onderwijs deel te nemen.

5 Discussie en aanbevelingen

Dit onderzoek ging over de vraag, of en in welke opzichten er verschillen bestaan in leerstijlen (in termen van de leerstijl-theorie van Vermunt, 1992) tussen cursisten die deelnemen aan het Hoger Onderwijs voor Ouderen (HOVO) enerzijds, en 'reguliere' studenten in het Wetenschappelijk Onderwijs anderzijds. Meer in het bijzonder is nagegaan of, en zo ja,

welke kwalitatieve verschillen er bestaan tussen beide groepen wat betreft (a) hun opvattingen over wat leren is (leerconcepties), (b) hun motivationele oriëntaties die ten grondslag liggen aan het studiegedrag, (c) de studie- en leeractiviteiten waaraan zij de voorkeur geven en (d) de soort van sturing (intern versus extern) die zij bij het leren prefereren. In samenhang hiermee is verder onderzocht, of de vier door Vermunt (1992) voor W.O.-studenten beschreven leerstijlen ook bij HOVO-cursisten konden worden teruggevonden.

Het onderzoek had een tweeledig doel: het eerste oogmerk was vast te stellen, in hoeverre de 'leerstijl-theorie' van Vermunt (1992) generaliseerbaar is naar andere populaties studenten binnen het Hoger Onderwijs; ten tweede ging het erom aanknopingspunten te vinden voor een adequate inrichting van het hoger onderwijs ten behoeve van senior-studenten.

Uit het onderzoek komt naar voren, dat de inhoud en de structuur van de vier leerstijlen, zoals die door Vermunt (1992) zijn beschreven voor studenten in het W.O., niet volledig konden worden teruggevonden bij studenten in het HOVO: het bleek op basis van de responsies van de HOVO-cursisten niet mogelijk om, met behulp van de werkwijze die Vermunt volgde (factoranalyse), de vier leerstijlen die hij heeft beschreven, i.c. de betekenisgerichte, de toepassingsgerichte, de reproductiegerichte en de stuurloze leerstijl, volledig te repliceren.

Waar Vermunt bijvoorbeeld vaste en elkaar uitsluitende samenhangen vond tussen bepaalde opvattingen over leren/studeren, motivationele oriëntaties, geprefereerde studie-activiteiten, en voorkeuren voor interne c.q. externe sturing, die te groeperen zijn in vier duidelijk onderscheiden leerstijlen (zie Figuur 1), blijken zulke combinaties bij HOVO-cursisten minder gearticuleerd en ook minder goed van elkaar te onderscheiden.

De belangrijkste verschillen tussen HOVO-cursisten en W.O.-studenten bestonden eruit, dat de eersten er veelal andere studiemotieven en studie-opvattingen op na houden. Voor wat betreft de aard en kwaliteit van de verwerkingsactiviteiten bij het leren en de aard van de regulatie van het leerproces bleken de verschillen minder sterk.

Sommige van de gevonden verschillen tussen HOVO- en 'reguliere' W.O.-studenten zijn

voor de hand liggend. Bijvoorbeeld de bevinding dat HOVO-studenten zich in het kader van hun studie veel minder laten leiden door het streven naar het behalen van certificaten, dat zij hun studie veel minder positioneren met het perspectief op een beroep, en ten slotte dat zij veel minder gericht zijn op het verwerven van kennis en vaardigheden met het oog om daar later gebruik van te maken.

HOVO-cursisten zijn daarbij vergeleken met reguliere studenten meer gericht op het dieper verwerken van de leerstof, stellen voorts een hogere mate van externe sturing op prijs, studeren vooral uitgaande van persoonlijke interesses en beschouwen hun leren vooral in het perspectief van het opbouwen van persoonlijke kennis (eigen kennisconstructie). Zij beoordelen het onderwijs vooral naar de mate waarin dit stimulerend voor hen is, en stellen minder belang in de mogelijkheden om met anderen te studeren.

De resultaten van dit onderzoek overziende, lijkt het erop, dat de beleving van het universitaire onderwijs door ouderen uitdrukkelijk anders is georiënteerd dan de beleving door jongere studenten. Ten eerste zijn natuurlijk de uitgangspunten van HOVO-cursisten en hun motieven om te gaan studeren van geheel andere aard dan die van jongere W.O.-studenten. Interesse of verdieping in een bepaald onderwerp is voor HOVO-cursisten een belangrijk motief. Ouderen leren meestal niet omdat ze een certificaat willen behalen of voor toepassing van het geleerde in hun beroep.

Een andere opvallende uitkomst van dit onderzoek was, dat ouderen de neiging hebben relatief veel te memoriseren en te herhalen, ook als zij de opvatting huldigen, dat leren het actief opbouwen van kennis omvat. Deze specifieke combinatie tussen geprefereerde verwerkingsactiviteiten (i.c. memoriseren/herhalen) en de leerconceptie 'opbouwen van kennis' is in vergelijking met het onderzoek van Vermunt zeker atypisch te noemen. De kwalitatief minder gewenste studie-activiteit van 'memoriseren/herhalen' die bij jongere studenten meestal karakteristiek is voor een 'reproductieve leerstijl', blijkt bij HOVO-cursisten te kunnen worden geassocieerd met een meer 'betekenisgerichte' instelling. Wellicht dient de combinatie 'herhalen/memoriseren' x 'leren als opbouwen van kennis' op oudere

leeftijd anders te worden geïnterpreteerd: om intellectueel goed te blijven functioneren is een bepaalde hoeveelheid parate kennis nodig. De senior-student zou meer moeite kunnen hebben om deze kennis op te bouwen en op peil te houden. Memoriseren en herhalen is in zo'n geval wellicht voor de oudere een noodzakelijke 'coping'-techniek. Er is over deze en soortgelijke kwesties echter aanzienlijk meer en diepgravender onderzoek nodig om tot andere dan tentatieve verklaringen voor de hier gevonden verschillen te komen.

Welke aanbevelingen kunnen er aan de resultaten van dit onderzoek worden verbonden voor de praktijk en de inrichting van het HOVO? Hoewel er – als gezegd – veel meer onderzoek nodig is, vooral om de hier gevonden verschillen tussen oudere en jongere studenten te kunnen *verklaren*, denken wij, dat het op grond van dit onderzoek gerechtvaardigd is om de volgende voorstellen te doen.

Aangezien senior-studenten relatief sterk gericht blijken op diepere verwerkingsactiviteiten, zouden vooral HOVO-cursussen moeten worden aangeboden, die zijn gericht op een grondige en diepgravende behandeling van een (specifieke) thematiek of vakgebied. Brede, algemene en inleidende cursussen lijken ons alleen zinvol, als hier een (verdiepend) vervolg op mogelijk is. Leergangen, in de vorm van een aantal samenhangende cursussen die een bepaald vakgebied bestrijken, zijn een uitstekende mogelijkheid om in de gesignaleerde behoefte aan diepgang te voldoen.

In verband met de door senior-studenten gewenste externe sturing, lijkt het ons raadzaam om docenten binnen het HOVO uitdrukkelijk op deze behoefte te attenderen. Waar deze behoefte aan externe sturing bij jongere studenten veelal duidt op onzekerheid omtrent aanpak van en kans op succes bij een studietoek, hebben wij de indruk, dat dit niet of veel minder geldt voor senior-studenten, die overigens in het algemeen ook weinig onzekerheid manifesteren (op de ILS). Deze behoefte aan externe sturing zou bij senioren wellicht kunnen berusten op de ervaring dat, wanneer men zich in korte tijd op een nieuw vakgebied wil inwerken, men beter kan profiteren van de vak-kennis en het inzicht van deskundigen.

Met betrekking tot het overwegende motief van senior-studenten om HOVO-cursussen te

volgen op basis van persoonlijke interesse, ligt het ons inziens voor de hand om vanuit de HOVO-opleidingen regelmatig en systematisch behoefte- c.q. interesse-onderzoek te verrichten onder (potentiële) cursisten om er enigszins zeker van te zijn, dat onderwijsaanbod en -behoefte nauw op elkaar aansluiten.

Omdat senioren over het algemeen een breed spectrum aan studie-activiteiten blijken te (willen) ontplooien om de aangeboden kennis te verwerken, verdient het ons inziens aanbeveling om het aantal contacturen als groep per week te beperken tot bijvoorbeeld maximaal 3 uur per week, vooral om ruimte te bieden aan de behoefte tot zelfstudie, zonder de gewenste mogelijkheden voor sociale contacten via gezamenlijk onderwijs teniet te doen. Uiteraard zijn goede bibliotheecare en andere informatievoorzieningen essentieel, ook voor succesvol HOVO-onderwijs.

Vanuit het oogpunt dat HOVO-cursisten een sterk stimulerende werking vanuit het onderwijs verlangen, zijn enthousiaste en stimulerende docenten die geen last hebben van stereotiepe opvattingen over (het leren en studeren van) ouderen, een noodzakelijke voorwaarde.

Hoewel HOVO-cursisten relatief weinig waarde toekennen aan het behalen van diploma's en certificaten, lijkt het ons vanuit een andere invalshoek wel van belang, dat de mogelijkheid tot het afleggen van tentamens wordt behouden. Deze tentamens zouden dan wel uitdrukkelijk de mogelijkheid aan de cursisten moeten bieden om na te gaan, in hoeverre en tot op welk niveau zij de studiestof zelf hebben *verwerkt* (diepteverwerking). Thuis-toetsen, portfolio-assessments en werkstukken lijken ons in dit verband goede en geschikte vormen van 'toetsing'.

En ten slotte nog dit: hoewel vanuit politieke en maatschappelijke organisaties sterk wordt aangedrongen op het inrichten van en voorzien in 'zinnvolle' cursussen voor ouderen (dat wil zeggen cursussen waar je in het leven buiten de onderwijssituatie mee uit de voeten kunt), kan op basis van dit onderzoek worden geconcludeerd dat dit voor het merendeel van de ouderen geen criterium uitmaakt bij de beslissing om al of niet een HOVO-cursus te gaan volgen. Het zinnvolle van een cursus ligt bij senior-studenten niet in het feit, dat de cur-

sus iets nuttigs te bieden heeft voor het leven van alledag, of uitzicht biedt op een diploma, en al evenmin dat deze een (gebruiks-) waarde vertegenwoordigt in relatie tot de professionele- of maatschappelijke praktijk. Waar het de HOVO-cursist vooral om gaat is om het eigen denken te verdiepen en bepaalde thematieken en disciplines beter te doorgronden. Welke docent zou zich met zo'n 'academische' instelling van de zijde van studenten niet in de handen wrijven?

Noot

- 1 Die bevinding (grotere behoefte aan externe sturing in functie van opleidingsniveau) komt ook uit ander door ons uitgevoerd onderzoek naar voren. Bijvoorbeeld Slaats, Van der Sanden en Lodewijks (1996) maken daarvan melding in onderzoek naar de samenhang tussen leerstijlen en persoonlijkheidskenmerken bij MBO-studenten.

Literatuur

- Bronneman-Helmers, H.M. (1993). *Opinies over onderwijs – Opvattingen van de bevolking en van opinieleiders over onderwijs(beleid)*. Cahiers Sociaal en Cultureel Planbureau, nr 102. 's-Gravenhage: Vuga.
- Candy, P.C. (1991). *Self-direction for lifelong learning: a comprehensive guide to theory and practice*. San Francisco/Oxford: Jossey Bass.
- Cramer, G., & Kamp, M. van der (1990). *Leren op latere leeftijd*. Lisse: Swets & Zeitlinger.
- Herwijer, L.J. (1990). *Educatie op leeftijd*. Rijswijk: Sociaal en Cultureel Planbureau.
- Linschoten, C.P., Gerritsen, J.C., Schaap, L., & Wolfensperger, E.W. (1989). *Samenhang in educatie voor ouderen, een overzicht van knelpunten in educatie voor ouderen*. Rijswijk: Voorlopige Raad van het Ouderenbeleid.
- Lodewijks, J.G.L.C. (1993). *De Kick van het Kunnen: over arrangementen en engagement bij het leren*. Tilburg: MesoConsult.
- Lodewijks, J.G.L.C., & Noij, L.G. (1994). *Activation of in-depth processing while learning*. Paper presented at the Annual Convention of the American Educational Research Association, New Orleans, April 1994.
- Maclure, S., & Davies, P. (Eds.) (1991). *Learning to think: Thinking to learn*. Oxford: Pergamon Press.
- Merriam, S.B., & Cunningham P.M. (1989). *Handbook of adult and continuing education*. London: Jossey-Bass.
- Noij, L.G., & Lodewijks, H.G.L.C. (1993). *Stimulating generative processing activities while studying: effects on learning*. Research Note. Katholieke Universiteit Nijmegen: Vakgroep Onderwijskunde/Katholieke Universiteit Brabant: Sectie Onderwijs- & Opleidingspsychologie.
- Ponti, E.E.C.M. de (1995). *Leren, een leven lang*. Doctoraal scriptie. Tilburg: Katholieke Universiteit Brabant, Onderwijs- & Opleidingspsychologie.
- Shuell, T.J. (1988). The role of the student in learning from instruction. *Contemporary Educational Psychology*, 13, 276-295.
- Simons, P.R.J. (1990). *Transfervermogen*. Inaugurele rede. Nijmegen: Katholieke Universiteit.
- Slaats, A., Sanden, J.M.M. van der, & Lodewijks, J.G.L.C. (1996). *Leerstijlen en persoonlijkheidskenmerken van MBO-studenten*. Paper gepresenteerd tijdens de OnderwijsResearchDagen. Tilburg: juni.
- Snijders, J.Th. (1990). Hoger onderwijs voor Ouderen in Nederland. *Universiteit en Hogeschool* 37(1), 1-12.
- Spiro, R.J., Feltovich, P.L., Jacobson, M.J., & Coulson, R.L. (1991). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill structured domains. *Educational Technology*, 31(5), 24-33.
- Steege, G. ter (1989). *Hoger onderwijs voor ouderen is het antwoord ... wat was de vraag?* In Congresbundel Erasmus Universiteit, Rotterdam: 29 november.
- Tuijnman, A.C., & Kamp, M. van der (1992). Learning for life: new ideas, new significance. In A.C. Tuijnman & M. van der Kamp (Eds.), *Learning across the lifespan. Theories, Research, Policies* (pp. 3-16). Oxford: Pergamon Press.
- Tempelman, C.C.J. (1988). *Typeringsonderzoek onder deelnemers aan cursussen Hoger Onderwijs Voor Ouderen aan de Rijksuniversiteit Groningen*. Verslag van een inleiding tijdens het congres "Oudewijs", Maastricht: 11-12 oktober 1988.
- Vermetten, Y., Vermunt, J., & Lodewijks, J.G.L.C. (1995) *Changes in learning styles as a result of student oriented education*. Paper presented at

the 6th EARLI Conference on Research in Learning and Instruction. Nijmegen, August.

Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs – naar procesgerichte instructie in zelfstandig denken*. Academisch Proefschrift, Katholieke Universiteit Brabant. Lisse: Swets & Zeitlinger.

Vermunt, J.D.H.M. & Rijswijk, F.A.W.M. van (1988). *Inventaris Leerstijlen*. Tilburg: Katholieke Universiteit Brabant, Onderwijs- & Opleidingspsychologie.

Manuscript aanvaard 20-3-1997.

Auteurs

E.C.M. De Ponti is als studie-adviseur verbonden aan de Faculteit Economische Wetenschappen van de Katholieke Universiteit Brabant te Tilburg.

J.G.L.C. Lodewijks is hoogleraar Onderwijs- & Opleidingspsychologie aan de Katholieke Universiteit Brabant te Tilburg.

N. Rögels is coördinator van de Senioren Academie Brabant te Tilburg.

Correspondentie-adres: J.G.L.C. Lodewijks, Katholieke Universiteit Brabant, Vakgroep Psychologie P-510, Postbus 90153, 5000 LE Tilburg

Abstract

Study motives and learning styles among students at an older age

E.C.M. De Ponti, J.G.L.C. Lodewijks & N. Rögels.
Pedagogische Studiën, 1997, 74, 197-209.

In this study we examined, if and to what extent 'regular' and 'senior' university students differ as to their learning conceptions, motivational orientations, preferences regarding the way learning processes are regulated and their preferred learning activities.

Departing from Vermunt's (1992) conceptualization of learning styles (i.e., self-consistent and pervasive orientations toward learning and studying), we compared learning style profiles of older students ($N = 88$), enrolled in different university courses for 'senior students', to profiles of first year university students from different departments of Tilburg University. These profiles were measured by a self-report questionnaire: the Inventory of Learning Styles ILS (Vermunt, 1992).

Striking differences were found, some of which were obvious, however. For instance, the result that the perspective that senior students take on their study, is less influenced by the need to acquire 'useful' knowledge and skills. Senior students prefer to study in a meaning oriented way, prefer external to self regulated learning, perceive knowledge acquisition generally from a constructive position and are stimulated mainly by genuine needs to know and understand.

Het einde van de klassenstrijd?

Achtergronden bij de discussie over klassengrootte in het basisonderwijs

R.J. Bosker*

Samenvatting

Recente peilingen laten zien dat de gemiddelde klassengrootte in basisscholen groot is vergeleken met andere landen en dat er bovendien een enorme variatie in klassengrootte is binnen, zowel als tussen scholen. Er wordt een samenvattend overzicht geboden van buitenlands onderzoek naar de effecten van klassengrootte op het functioneren van leerlingen en leerkrachten. Verder komt een grootschalig Nederlands onderzoek aan bod waarin klassengrootte centraal stond. Dit onderzoek laat zien dat er een samenhang bestaat tussen klassengrootte en leerlingprestaties op het terrein van taal en rekenen. Leerlingen uit leerjaar 2 die in klassen zitten met 25 of meer leerlingen hebben een grotere kans onder hun niveau te presteren dan vergelijkbare leerlingen uit kleinere klassen. Voor de leerjaren 4, 6 en 8 zijn de resultaten minder eenduidig. Zo is er bijvoorbeeld de bevinding dat in klassen met 30-34 leerlingen relatief hogere rekenprestaties behaald worden dan in kleinere klassen. Desalniettemin is er een duidelijke indicatie dat klassen met 35 of meer leerlingen vergelijken-derwijze niet gunstig zijn. Voor welbevinden van leerlingen, noch voor de onderwijsattitude, noch voor de arbeidssatisfactie van de leerkracht kan een verband met klassengrootte worden aangetoond. In de nabeschuiving, waarbij de effecten van klassengrootte vergele-

ken worden met de effecten van 'Reading Recovery' en het 'Succes for All' interventieprogramma, wordt de hypothese geopperd dat het niet zozeer de klassengrootte is die de genoemde effecten teweeg brengt, maar meer algemeen de leerling/leerkracht ratio.

Inleiding

Van oudsher verheugt het probleem van de klassengrootte zich in een bijzondere belangstelling (zie bijv. Van Gelder, 1959). Het ideaal dat een enkele onderwijskundige in het begin van deze eeuw schetst, namelijk 20 leerlingen per klas (idem, p. 500), komt steeds dichterbij: rond 1900 bevatte meer dan de helft der klassen 40 leerlingen of meer, in 1960 bedroeg de gemiddelde klassengrootte¹ nog 36 leerlingen en sinds 1972 ligt het gemiddelde onder de 30 (Warries, 1985, p. 7).

Met het verschijnen van het inspectierapport 'Groeps-grootte in het basisonderwijs' (Inspectie van het Onderwijs, 1995) werd evenwel duidelijk dat er zich zowel tussen als binnen basisscholen in Nederland enorme verschillen voordoen in klassengrootte. Er zijn klassen met minder dan 10 en klassen met meer dan 40 leerlingen. Op grote scholen (met meer dan 200 leerlingen) bevat de gemiddelde klas 10 leerlingen meer dan op de kleine scholen (met minder dan 100 leerlingen). De gemiddelde klassengrootte bedroeg in mei 1995 25,7 leerlingen. Een kwart van de klassen herbergt meer dan 30 leerlingen gedurende het laatste deel van het schooljaar. Dit impliceert dat, als we even afzien van leerkrachten met een part-time aanstelling, 1 op de 4 leerkrachten met een klas van meer dan 30 leerlingen te maken heeft, en dat – bij benadering – 1 op de 3 leerlingen in een klas van een dergelijke omvang zit.

In dit artikel zullen we ingaan op de volgende vragen:

1. Wat moeten we precies verstaan onder het – op het eerste gezicht zo eenduidig lijkende – begrip klassengrootte?
2. Wat is internationaal gezien de situatie met betrekking tot de grootte van klassen in het basisonderwijs?
3. Hoe komt het dat sommige klassen groter zijn dan andere?
4. Wat is bekend over de gevolgen van klassengrootte?

* Met dank aan J. Hox, K.-J. Polder en S. Veenman voor hun constructieve bijdragen aan de inhoud van dit artikel.

Klassengrootte en de leerling/ leerkracht ratio

Bij de (vaak slechts impliciete) definities van klassengrootte moeten we een onderscheid maken in de aggregatieniveaus en in de mate van specificiteit.

Op het hoogste aggregatieniveau, namelijk het macro-niveau, is de leerling/leerkracht ratio voor een bepaalde sector (in dit geval het basis-onderwijs) gebaseerd op een deling van het aantal leerlingen in de betreffende sector op het aantal personeelsleden in die sector. De kerngegevens primair onderwijs (Ministerie van OC&W, 1996) vermelden voor het basisonderwijs bijvoorbeeld een aantal van 19.3 leerlingen per full-time equivalent personeel in het onderwijs. Op het niveau van de school, het meso-niveau, wordt de leerling/leerkracht ratio analoog berekend. In het wat oudere, economisch getinte onderzoek wordt vaak deze indicator gebruikt, die vervolgens gerelateerd werd aan de prestaties van leerlingen uit een bepaald leerjaar (bijv. Coleman c.s., 1966). Op het micro-niveau is het een telling van het aantal leerlingen per klas, zoals dit ook in het inspectie-onderzoek is gebeurd. De inspectie definieerde een klas als het aantal leerlingen dat voor tenminste 50% van het aantal lessen in de week gelijktijdig les krijgt van één leerkracht (dat hoeft overigens niet één en dezelfde te zijn, gelet op duobanen, deeltijdarbeid, adv-ervanging e.d.). De macro- en mesoniveau benadering leveren een onderschatting van de situatie op: niet alle formatie wordt ook daadwerkelijk ingezet voor instructie en begeleiding (te denken valt aan: managementtaken, taakuren, e.d.). In de praktijk levert de micro-niveau benadering, althans zoals gehanteerd door de inspectie, een overschatting op: vaak worden bijvoorbeeld combinatieklassen gedurende een aantal lessen gesplitst in kleinere instructiegroepen.

Wat betreft specificiteit treden er in de praktijk ook grote verschillen in de gehanteerde definities op. Grofmazig zijn dan de micro-niveau definities (om nog maar te zwijgen over de meso- en macro-niveau definities) die niet tijdseenheden en vakinhouden specificeren.

Veel specifiekier zijn Glass et al. (1982). Deze auteurs hebben de bevindingen rond klassengrootte geïnventariseerd. Als we dit standaardwerk bekijken, dan blijkt dat het vraagstuk van de klassengrootte wordt begrepen als de vraag naar de instructiesituatie: gegeven een criterium (bijvoorbeeld rekenprestaties), wat is dan het effect van de omvang van de groep waarin een leerling in een bepaalde tijdsspanne instructie en begeleiding bij de verwerking van de rekenleerstof krijgt?

Zij die beweren dat op basis van het beschikbare onderzoek de conclusie gerechtvaardigd is dat reductie van de klassengrootte gunstig uitpakt voor de cognitieve ontwikkeling van leerlingen, verwijzen meestal naar de meta-analyse van Glass en Smith (1978). We komen hierop nog terug. Zij die beweren dat klassengrootte er niet zoveel toe doet, verwijzen meestal naar onderzoek waar veel grofmaziger te werk is gegaan; met name Hanushek (1986) wordt in dit verband veel geciteerd. Tenslotte zijn er nog die beweren dat grote klassen beter zijn en zij voelen zich daarbij gesteund door een onderzoek van Van Batenburg, Koster en Ros (1989).

Niet altijd echter wordt dezelfde definitie van klassengrootte gehanteerd. Soms is het de leerling/leerkracht ratio, soms is het de klassengrootte van een jaargroep en dan weer wordt het gepreciseerd naar vak en tijd en instructie-arrangementen. In het verloop zal ook duidelijk worden dat de conclusies van Batenburg c.s. alleen onder zeer bepaalde condities valide zijn, dat Hanushek de situatie onderschat, en dat Glass c.s. de positieve effecten van verkleining van de klassengrootte overschatten.

Internationale gegevens over klassengrootte

Het gemiddelde van bijna 26 leerlingen per klas in Nederland is internationaal gezien hoog. De IEA publiceerde op basis van de 'Reading Literacy' studie internationale gegevens over klassengrootte en leerling/leerkracht ratio (Ross & Postlethwaite, 1994), die in Tabel 1 gepresenteerd worden.

De eerste kolom met gegevens bevat de

Tabel 1
Enkele internationale gegevens over het basisonderwijs (bron: Ross & Postlethwaite, 1994, p. 129)

land	klassengrootte	leerling/ leerkracht ratio
België (Wallonië)	20.3	16.9
Denemarken	17.2	14.9
Duitsland (BRD)	22.5	22.0
Finland	24.6	19.6
Frankrijk	23.5	21.9
Hongarije	23.4	16.0
Ierland	31.0	30.6
Italië	16.0	16.8
Nederland	25.7	22.5
Nieuw-Zeeland	29.7	24.3
Noorwegen	17.7	12.1
Spanje	28.8	24.5
Verenigde Staten	24.6	20.6
Zweden	20.1	13.1
Zwitserland	19.6	14.5

gemiddelde klassengrootte. In een aantal landen (Denemarken, Italië, Noorwegen en Zwitserland) bevat de gemiddelde klas minder dan 20 leerlingen. In Ierland, Nieuw-Zeeland en Spanje ligt dit aantal bijna 10 hoger. Ook Nederland scoort met 25.7 leerlingen aan de hoge kant.

In de tweede kolom staat de leerling/leerkracht ratio. Gemiddeld is er in Nederland op elke 22.5 leerlingen één leerkracht werkzaam. De ongunstigste ratio wordt gerealiseerd in Ierland met 30.6, terwijl België, Denemarken, Hongarije, Italië, Zweden en Zwitserland ratio's van ongeveer 15 leerlingen op 1 leerkracht kennen.

Nu moeten deze cijfers met de nodige voorzichtigheid gehanteerd worden. Ierland, bijvoorbeeld, geeft relatief meer geld uit aan onderwijs dan Nederland, maar omdat de relatieve omvang van de populatie kinderen van 5 tot 14 jaar in Ierland zeer groot is, is er per leerling minder te besteden. Anderzijds hebben landen als Denemarken en Italië het speciaal onderwijs voor een belangrijk deel geïntegreerd in het reguliere onderwijs, zodat hier de klassen om deze reden kleiner kunnen uitvallen. Desalniettemin lijkt zich de conclusie op te dringen: vrij grote klassen in het Nederlandse basisonderwijs. Grotere klassen impliceren minder uitgaven voor leerkrachtsalarissen. Men zou hier dus van een grote doelmatigheid kunnen spreken, als Nederlandse leerlingen tenminste goed zouden presteren. Maar wat betreft het leesvaardigheidsniveau van onze 9-

jarigen verkeert Nederland in de achterhoede, blijktens recente IEA-onderzoeken (zie bijv. De Gloppe & Otter, 1993, p. 15) en de Commissie Evaluatie Basisonderwijs (1994) sprak haar zorgen uit over het rekenpeil.

De oorzaken van verschillen in klassengrootte

Hoe kan het dat sommige klassen groter zijn (of lijken) dan andere? De verschillen in leerling/leerkracht ratio's tussen landen, en dus ook in klassengroottes, zijn het directe gevolg van beleidskeuzen van de overheid. Sommige landen investeren meer in het onderwijs dan andere landen, stellen met andere woorden meer geld en dus meer formatie beschikbaar (Organization for Economic Co-operation and Development, 1995). Vergeleken met andere OESO-landen (gemiddeld 6.1%) besteedt Nederland relatief een gering percentage (5%) van het bruto nationaal product aan het onderwijs. De verdeling van deze middelen over de diverse sectoren volgt echter nauwgezet het gemiddelde OESO-patroon: Voor alle sectoren (primair, secundair en tertiair) geldt dat de uitgaven onder het OESO-gemiddelde liggen. De spreiding in klassengrootte binnen landen is vervolgens afhankelijk van een tweede beleidskeuze: Worden er wettelijke regels gesteld ten aanzien van de klassengrootte? In Tabel 2 zijn voor een aantal landen de wettelijke maximale klassengroottes aangegeven.

Klassen groter dan 35 kunnen in Grieken-

Tabel 2
Wettelijk bepaalde maximale klassengroottes in enkele landen (bron: RISE, 1994; OESO, 1995, p. 287)

land	maximale klassengrootte
Denemarken	28
Duitsland	25 voor jaargroep 1 en 2 29 daarna
Finland	25 voor jaargroep 3 en 4 32 daarna *
Frankrijk	25 jaargroep 1
Griekenland	35
Italië	20 voor klassen met een gehandicapt kind 25 anders
Noorwegen	28
Schotland*	25 voor combinatieklassen 33 anders

*) Niet bij wet geregeld, maar in de CAO vastgelegd.

Tabel 3

Aantal klassen en klassengrootte per type school (bron: Van der Werf & Guldemond, 1994, p. 35)

type school*	aantal klassen	gemiddelde klassengrootte	beschikbare extra formatie per klas
A. Zonder OVB-faciliteiten (schoolscore < 105)	7	25	0
B. Gebiedsscholen met $105 < \text{schoolscore} \leq 115$	6	22	0
C. Niet-gebiedsscholen met $105 < \text{schoolscore} \leq 115$	6	22	0
D. Gebiedsscholen met schoolscore > 115	8	21	0.2
E. Niet-gebiedsscholen met schoolscore > 115	7	22	0.1

*) schoolscore = $100 \times (\text{gewogen leerlingaantal} / \text{ongewogen leerlingaantal} \text{ minus een correctiefactor van } 9\%)$

land dus niet voorkomen, omdat dit wettelijk verboden is. Voor Denemarken en Noorwegen ligt het wettelijk bepaalde maximum zelfs op 28. Andere landen leggen de maxima vast afhankelijk van de jaargroep, met lagere maxima voor de jaargroepen die de jongere leerlingen herbergen. In Schotland is een en ander niet bij wet geregeld maar door middel van afspraken tussen de sociale partners vastgelegd. Bijzonder is bovendien dat in Schotland een apart lager maximum is overeengekomen voor combinatieklassen (25 tegenover 33 voor enkelvoudige klassen).

Dat er verschillen tussen scholen in Nederland bestaan in de gemiddelde klassengrootte, heeft te maken met de formatietoekenning, waarbij zowel de schoolgrootte als de samenstelling van de leerlingpopulatie een rol speelt. Kleine scholen, met minder dan 100 leerlingen, krijgen relatief veel, en grotere scholen relatief weinig formatie. Zouden scholen alle formatie behoeve van de minimalisering van de klassengrootte inzetten, dan kunnen scholen met minder dan 100 leerlingen hun gemiddelde klassengrootte op 18.1 stellen. Voor scholen met 100-200 leerlingen, respectievelijk met meer dan 200 leerlingen liggen deze getallen op 23.2, respectievelijk 24.2. Scholen met achterstandsleerlingen, dat wil zeggen kinderen uit de lagere sociale milieus van allochtone (1.9) of autochtone ouders (1.25), krijgen daarnaast extra formatie. Dit kan ertoe leiden dat op 'zwarte' scholen bijna twee keer zo veel formatie beschikbaar is als op 'witte' scholen. Deze extra formatie wordt in 46% van de gevallen geheel, en in 44% van de gevallen gedeeltelijk

ingezet voor verkleining van de klassengrootte (Inspectie van het Onderwijs, 1995, p. 28).

Scholen hebben de vrijheid de toegekende formatie naar eigen goeddunken in te zetten voor onderwijs, begeleiding, taakuren, managementtaken e.d. Verschillen in klassengroottes binnen een school zijn enkel en alleen het gevolg van door de school zelf gemaakte keuzen:

- werken met een leerstofjaarklassysteem;
- het formeren van combinatieklassen;
- het al dan niet doorbreken van het klassenverband gedurende enkele uren voor wezenlijke onderwijsactiviteiten;
- de specifieke inzet van formatie voor achterstandsbestrijding.

Van der Werf en Guldemond (1994) onderscheiden 5 typen scholen op basis van de leerlinggewichten en participatie in een OnderwijsVoorangsGebied. Per type school hebben ze vervolgens het gemiddeld aantal klassen, de gemiddelde klassengrootte en de hoeveelheid extra beschikbare formatie per klas berekend (zie Tabel 3).

De gegevens van Van der Werf en Guldemond suggereren dat scholen hun extra formatie met name gebruiken om de klassen klein te houden. Dat gebiedsscholen met een schoolscore groter dan 115 gemiddeld 8 klassen formeren, en vervolgens formatie overhouden (die dan wellicht ingezet wordt voor remedial teaching, NT2, OET, e.d.) lijkt erop te wijzen dat scholen een dergelijk formatiegebruik kennelijk effectiever achten dan het formeren van

kleinere combinatieklassen. Belangrijke overwegingen bij de indeling van de klassen zijn: zo veel mogelijk combinatieklassen vermijden, het kunnen opvangen van tussentijdse instroom, het klein houden van jaargroep 3, en het klein houden van klassen in verband met zorgverbreding (Inspectie van het Onderwijs, 1995, p. 31, 32).

Kennelijk achten scholen het in verband met achterstandsbestrijding effectief de klassen klein te houden, en achten ze grotere, enkelvoudige klassen effectiever en/of beter hanteerbaar dan kleinere combinatieklassen.

De gevolgen van klassengrootte voor leerlingprestaties

Enkele inleidende opmerkingen

In deze paragraaf brengen we de effecten van klassengrootte op leerprestaties en het welzijn van leerlingen en het functioneren van leerkrachten in kaart op basis van eerder en zelf (Bosker & Hox, 1996) verricht onderzoek.

De vraag of leerlingen in grotere klassen slechter presteren, meer faalangst hebben en minder gemotiveerd zijn, kan alleen door experimenteel onderzoek afdoende worden beantwoord. Hetzelfde geldt als we de gevolgen van klassengrootte voor het leerkrachtfunctioneren willen vaststellen. In zo'n experiment, met als condities grote en kleine klassen, zouden leerlingen en leerkrachten aselekt aan klassen moeten worden toegewezen. Op die manier worden andere storende invloeden (de kwaliteit van de leraar, de bewuste verdeling van goede en slechte leerlingen in grote en kleine klassen, etc.) geneutraliseerd. En dan nog zijn er problemen:

- er dreigt een Hawthorne-effect (als we gaan experimenteren met kleine klassen, dan gaan deze zich als gevolg van de aandacht die ze krijgen van de onderzoeker gunstiger ontwikkelen) of tegengesteld daaraan;
- een John Henry-effect (de leerkrachten in de controleconditie, d.w.z. met de grote klassen, gaan onder invloed van het onderzoek extra hun best doen).

Experimenteel onderzoek is een zeldzaamheid in het onderwijs. Meestal wordt gebruik gemaakt van observationele onderzoeksontwer-

pen. Dat wil zeggen: we kijken of er klassen zijn die verschillen in omvang en of de kinderen in de grotere klassen anders functioneren dan de kinderen in de kleinere klassen. In dat geval hebben we geen last van de net beschreven problemen, maar nu zijn er weer andere:

- de context van de klas, in casu de school speelt een rol. Stel dat bijvoorbeeld de ervaren leerkrachten de kleinere klassen hebben; als de kinderen in de kleinere klassen het beter doen, kan dat komen omdat de leerkrachten zo ervaren zijn, en niet omdat de klas zo klein is;
- de leerlingen zelf dragen een onderwijsverleden met zich mee. Als de school bijvoorbeeld besluit de minder gevorderde leerlingen in de kleinere klassen te zetten, dan zouden we kunnen vinden dat in de grotere klassen de kinderen beter presteren.

Aan het slot zullen we op dit probleem van causale attributie (wat is de 'echte' oorzaak?) terugkomen, wanneer de bevindingen worden samengevat en gewogen worden op hun implicaties.

De meta-analyse van Glass en Smith

Honderden studies naar klassengrootte werden door Glass en Smith (1978) geïnventariseerd en aan een nadere meta-analyse onderworpen om te bezien wat nu de beste samenvatting van al die resultaten zou zijn. Het resultaat ten aanzien van klassengrootte-effecten op leerprestaties lijkt in eerste instantie weinig veelzeggend: 40% van de 725 studies toont voordelen voor de grotere klassen en 60% voor de kleinere. De gemiddelde effectgrootte² die Glass en Smith rapporteren, is 1/10e van een standaarddeviatie (0.09 om precies te zijn) in het voordeel van de kleinere klassen. Maar dat voordeel kan tot stand zijn gekomen door de zeer kleine groepen.

De studie van Glass en Smith is niet zonder kritiek gebleven. Met name de in de meta-analyse dominerende rol van de onderzoeksresultaten van studies waar zeer kleine klassen met minder dan vijf leerlingen werden vergeleken met grote klassen werd bekritiseerd. Bovendien zouden meerdere studies door Glass en Smith meermalen zijn gebruikt (omdat bijvoorbeeld zowel reken- als taalprestaties werden onderzocht) en betroffen sommige studies irre-

Tabel 4

Klassengrootte effecten op leerprestaties in primair en secundair onderwijs bij experimentele onderzoeken (bron: Slavin, 1989, p. 249); tussen haakjes staan de uitgevoerde vergelijkingen

	Omvang van de kleine klas (vergelijking)			
	1	3	14-17	20-23
	.65 (1-32)	1.22 (3-25)	.17 (14-30)	.15 (20-28)
	.78 (1-30)		.17 (15-30)	.04 (23-27)
	1.52 (1-25)		.08 (16-37)	.04 (23-30)
	.72 (1-14)		.04 (16-30)	.00 (30-37)
	.30 (1-8)		.05 (16-23)	
	.22 (1-3)		-.29 (17-35)	
mediane effectgrootte	.69		.06	.04

levante domeinen (in één geval ging het om tennislessen). Reden voor Slavin (1989) nog eens naar de oorspronkelijke onderzoeken te kijken. Van de 77 oorspronkelijke studies bleken er slechts in veertien gevallen leerlingen aselekt aan grote dan wel kleine klassen toegevoerd te zijn. Slechts veertien experimentele studies dus. Na verwijdering van de tennislessen-studie, de studies die betrekking hadden op post-secundair onderwijs en studies waar het onderwijs slechts 30 minuten duurde, komt Slavin tot de volgende samenvattende resultaten (zie Tabel 4).

Tabel 4 laat zich als volgt lezen. In de eerste kolom staan de resultaten van studies vermeld waar individueel onderwijs vergeleken is met onderwijs in groepen. Voor de eerste studie waar individueel onderwijs wordt afgezet tegen onderwijs in een klas van 32 leerlingen (links bovenin de tabel terug te vinden als 1-32), is de effectgrootte 0.65.

In de onderste rij wordt de mediane effectgrootte gegeven als de kleine klas vergeleken wordt met de grotere klassen. Op basis van deze gegevens komt Slavin tot de conclusie dat de kleinere klassen (14 tot 17 leerlingen of 20 tot 23 leerlingen) het amper beter doen dan de grotere klassen: 0.04 respectievelijk 0.06 is de effectgrootte. En waar de kleinere klassen het beter doen, blijken, aldus Slavin, op de lange duur de effecten weg te ebbelen. De enige consequente effecten zijn te vinden als een leerling individueel in plaats van in een groep onderwijs ontvangt: 0.69 is hier de effectgrootte.

Sinds de meta-analyse van Glass en Smith en de kritiek die daarop is gekomen is er echter een aantal studies naar klassengrootte uitge-

voerd die nieuw licht werpen op de zaak en/of betrouwbaardere uitkomsten hebben opgeleverd. We gaan kort op deze studies in.

Een Nederlandse studie

In Nederland is in 1989 een onderzoek naar klassengrootte gerapporteerd door Van Batenburg e.a. (1989). De onderzoekers maken gebruik van de Cito-Eindtoetsgegevens van ruim 800 scholen, die ze vervolgens relateren aan de klassengrootte van jaargroep 8. Zij vinden in een analyse op het niveau van de school een positief verband tussen klassengrootte en de Cito-Eindtoetsprestaties, ook als er wordt gecorrigeerd voor de samenstelling van de schoolbevolking in termen van de sociaal-economische positie van de ouders der leerlingen. Zouden we de gerapporteerde resultaten omzetten in effectgroottes, dan zijn deze negatief: -0.10. Om twee redenen hebben deze resultaten een beperkte reikwijdte:

1. De Cito-Eindtoets meet het resultaat van acht jaar basisonderwijs, terwijl de klassengrootte van jaargroep 8 slechts een achtste deel omvat van de 'behandeling' die de kinderen acht jaar lang gekregen hebben. Wellicht worden grotere klassen in jaargroep 8 juist gevormd om kleine klassen in jongere jaargroepen te kunnen maken en vice versa. In dat geval zou het effect dus het gevolg kunnen zijn van de kleinere klasomvang in de jongere jaargroepen. Een dergelijke alternatieve verklaring is slechts uit te sluiten als alleen de leerwinst in het achtste leerjaar geanalyseerd zou worden.
2. Analyses op schoolniveau zijn uitermate gevaarlijk als we er valide conclusies uit

willen trekken over het individuele functioneren van leerlingen. Corrigeren voor de sociaal-economische samenstelling van de schoolbevolking heeft slechts een minimaal effect en is – conceptueel en statistisch – heel iets anders dan corrigeren voor de sociaal-economische achtergrond van individuele leerlingen (zie Bosker & Sniijders, 1990). Dergelijke analyses op schoolniveau worden wel getypeerd als ‘gevaarlijk op zijn best, rampzalig op zijn ergst’³.

Een Australische studie

Larkin en Keeves publiceerden in 1984 een onderzoek naar klassengrootte in Australië. Ze maakten gebruik van gegevens uit 1969 over de prestaties van jaargroep 8 leerlingen in ‘science’ (een vak met onderdelen van natuurkunde, biologie, scheikunde en aardrijkskunde) en rekenen/wiskunde. Het onderzoek is tamelijk standaard in die zin dat het een observationeel onderzoeksontwerp betreft (gekeken wordt of klassen verschillen in grootte – men is dus gebonden aan de natuurlijke variatie die zich in dit opzicht in het onderwijsveld voordoet – en vervolgens wordt nagegaan of in de kleinere klassen andere dingen gebeuren, en of dat samenhangt met prestaties van leerlingen). De conclusies:

- de vaardiger leerlingen werden meestal in de grotere klassen geplaatst (zodat in de kleinere klassen de minder vaardige leerlingen meer begeleiding konden krijgen);
- in grotere klassen lag het prestatieniveau uiteindelijk hoger;
- in kleinere klassen was het verschil tussen oorspronkelijk hoge en lage presteerders uiteindelijk minder groot dan in de grotere klassen.

Plaatsing van de leerlingen in klassen van verschillende omvang blijkt in dit onderzoek dus rechtstreeks gebaseerd te zijn op het vaardigheidsniveau. Iets dergelijks zou zich ook in de Nederlandse situatie kunnen voordoen, waar immers dankzij de gewichtenregeling achterstandsleerlingen een grotere kans hebben in de kleinere klassen terecht te komen. Het onderzoek maakt duidelijk hoe problematisch het is om onderzoek te doen naar de effecten van klassengrootte als de verschillende klassen systematisch verschillen in hun samenstelling.

Onderwijseconomisch onderzoek naar klassengrootte

In onderwijseconomische kringen wordt nogal eens verwezen naar een review van Hanushek (1986), waarin de auteur 112 studies naar leerkracht/leerling ratio op een rij zet (de economisten draaien de actoren om bij de ratioberekening): 14 studies leveren een significant negatief verband op en 9 een significant positief verband. Kortom: onbeslist. Hedges, Laine en Greenwald (1994) onderwierpen de studies aan een statistische meta-analyse. Deze auteurs komen op grond van hun analyses tot de slotsom, dat het niet is uit te sluiten dat een gunstige leerkracht/leerling ratio positieve effecten heeft. Zij trekken uiteindelijk de conclusie dat uitgaven per leerling een duidelijk verband met leerlingprestaties laten zien, maar dat de samenhang tussen de leerkracht/leerling ratio en de leerlingprestaties minder eenduidig is: het hangt kennelijk van de specifieke omstandigheden van een school af hoe het geld het best gebruikt kan worden.

De review van Hanushek was voor Akerhielm (1995) aanleiding te onderzoeken of in observationeel onderzoek kleinere klassen niet goed uit de verf komen, omdat zij vaker de minder vaardige leerlingen herbergen. Een eerste heranalyse van een dataset waarin meer dan 1000 scholen vertegenwoordigd zijn met elk 24 leerlingen uit jaargroep 8, laat zien dat voor klassen tot 20 leerlingen het klassengrootte-effect positief is. In het gebied tussen de 21 en 30 is er geen effect. Wanneer de onderzoekster echter controleert voor de verschillen in klassensamenstelling qua oorspronkelijk vaardigheidsniveau blijken er omgekeerde effecten op te treden voor twee van de vier onderzochte vakken (namelijk voor geschiedenis en science, maar niet voor rekenen en taal): kleinere klassen doen het dan iets beter in geschiedenis en science. De richting van het verband tussen klassengrootte en prestaties is dus soms inderdaad afhankelijk van het feit of rekening wordt gehouden met de samenstelling van de klas.

Het Prime Time onderzoek

In Indiana (VS) werd in 1984 Prime Time ingevoerd, een beleidsprogramma bedoeld om het kennisniveau van de leerlingen te verhogen door de klassen te verkleinen. In het schooljaar

84/85 werd begonnen met de reductie van de klassengrootte in jaargroep 3 ('1st grade'), waarna in 85/86 jaargroep 4 en 86/87 jaargroep 5 en jaargroep 2 ('Kindergarten') voor het eerst onder het beleidsprogramma vielen. Door de kleinere klassen (met 19.1 leerlingen gemiddeld) te vergelijken met de grotere klassen (met 26.4 leerlingen gemiddeld) van vóór de implementatie van Prime Time kan een schatting van het effect van reductie van de klassengrootte worden gemaakt. In de controlescholen die altijd al kleine klassen hadden gehad, bleek in twee jaar tijd niets veranderd te zijn (McGivern, Gilman & Tilitzki, 1989). In de scholen waar de klassengrootte was verkleind, deden de kleinere klassen het duidelijk beter na twee jaar. Als we de vastgestelde effectgrootte voor rekenen en taal samen bezien, dan is deze 0.34.

Het STAR-experiment

Een experimenteel onderzoek naar klassengrootte is uitgevoerd in Tennessee in de Verenigde Staten van Amerika (Word, 1990). De aanleiding vormde een politieke discussie over de enorme kosten van verkleining van de klassengrootte in verhouding tot de onzekere opbrengsten daarvan. Om uit de impasse te geraken werd besloten tot een nadere studie, die anders dan gebruikelijk in onderwijsonderzoek, als een echt experiment werd opgezet. 79 scholen werden – allesbehalve aselekt, namelijk op basis van vrijwilligheid en voldoende grootte van de jaargroepen – geselecteerd voor het onderzoek. Binnen elke school – dit is qua onderzoeksontwerp een sterk punt, omdat de grote en kleine klassen functioneren binnen dezelfde context – werden drie typen kleuterklassen ('Kindergarten') geformeerd: een klas van 13-17 leerlingen, een klas van 22-25 leerlingen en een klas van 22-25 leerlingen waar een full-time onderwijs-assistent zou gaan werken. Leerlingen en leerkrachten werden aselekt aan de klassen toegewezen. De klassen bleven vier jaar lang (tot en met '3rd grade', ofwel jaargroep 5) grotendeels intact (tussentijdse instroom of uitval had geen invloed op de resultaten van het experiment). De statistische analyses zijn geheel comme-il-faut, en doen dus recht aan het feit dat toeval een rol speelt op het niveau van leerlingen én klassen én scholen. De resultaten in termen van effectgroottes zijn weergegeven in Tabel 5.

Tabel 5

Het voordeel van kleine klassen volgens het STAR-experiment (naar Word, 1990, p. 24)

	taal	rekenen
jaargroep 2	.21	.17
jaargroep 3	.34	.33
jaargroep 4	.26	.23
jaargroep 5	.24	.21

Tabel 5 bevat de gegevens over één generatie leerlingen die vier jaar lang gevolgd zijn op hun weg door het onderwijs. De grote positieve effecten van de kleine klas aan het einde van jaargroep 3 (0.34 voor rekenen en 0.33 voor taal) houden niet ten volle stand tot aan het eind van jaargroep 5, maar zijn nog steeds aanzienlijk (0.24 voor rekenen en 0.21 voor taal).

Zowel voor rekenen als voor taal zijn de resultaten in het voordeel van kleine klassen (gemiddeld 15 leerlingen). Grote klassen (met gemiddeld 23 leerlingen) met en zonder onderwijsassistent (die onderwijsassistent heeft dus geen effect in deze studie) doen het beide slechter dan de kleine klassen. Vervolgens zijn analyses uitgevoerd om te onderzoeken of het effect van klassengrootte verschillend is:

- voor scholen in verstedelijkte en plattelandsgebieden,
- voor jongens en meisjes,
- voor oorspronkelijk (d.w.z. in het begin van de kleuterklas) hoog en laag presterende leerlingen,
- voor hoog- en laag-milieu leerlingen,
- voor alloctonen en autoctonen.

Kleine klassen blijken het onder alle condities beter te doen dan grote klassen. Finn en Achilles (1990) laten bovendien zien dat bij het STAR-experiment de effecten van kleine klassen voor kansarme leerlingen zelfs nog gunstiger uitpakken dan voor hun beter bedeelde medeleerlingen.

Vervolgonderzoek (Achilles et al., 1993) op het STAR-project bracht bovendien nog twee interessante zaken aan het licht:

1. Ook nadat de leerlingen uit de kleine klassen na het 5e leerjaar teruggekeerd waren naar klassen met een voor Tennessee normale omvang bleef hun voorsprong op de kinderen uit de van oorsprong grotere klassen behouden;
2. Omdat STAR succesvol bleek, ging men

over tot de invoering van zeer kleine klassen (13-17 leerlingen) in 17 arme districten met veel risicoleerlingen. Scoorden deze districten qua taal- en rekenprestaties altijd ver onder het gemiddelde van Tennessee, na de klassenverkleining scoorden ze daar (net) boven (Mosteller, 1995).

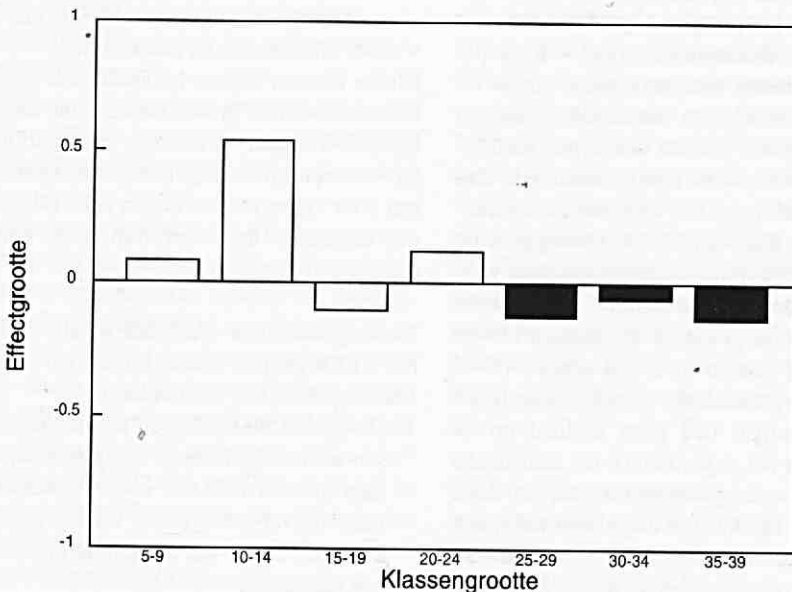
Klassengrootte en leerprestaties: de Nederlandse situatie

Om zicht te krijgen op de samenhang tussen klassengrootte en leerprestaties in Nederland is door Bosker en Hox (1996) een heranalyse uitgevoerd op de gegevens, zoals die verzameld worden in het kader van het zogenaamde PRIMA-cohort onderzoek (zie bijv. Jungbluth, Peetsma & Roeleveld, 1996). Het PRIMA-onderzoek heeft onder andere betrekking op 416 aselekt gekozen basisscholen die samen de zogenaamde referentiesteekproef vormen. In het schooljaar 94/95 werden de leerkrachten die het schooljaar daarvoor les hadden gegeven aan de jaargroepen 1, 3, 5 en 7 schriftelijk bevraagd, o.a. over de klassengrootte. Van alle leerlingen werd in het begin van het schooljaar 94/95 (maar voor de jaargroep-2 leerlingen pas halverwege het schooljaar) in kaart gebracht: leerlinggewicht, sekse, reken- en taalprestaties. Bovendien werd het IQ van de leerlingen in de

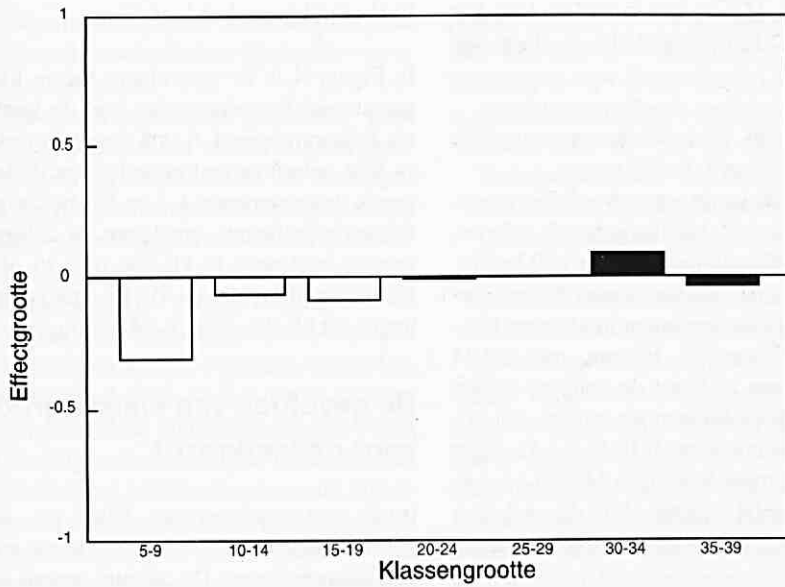
jaargroepen 4, 6 en 8 gemeten (voor de constructie van deze toetsen zie: Vierke, 1995). De samenhangen tussen klassengrootte en leerlingprestaties die hier gerapporteerd worden, zijn alle vastgesteld na correctie voor de achtergrondkenmerken van leerlingen en de samenstelling van de klas. In feite wordt dan de samenhang onderzocht tussen de klassengrootte en de mate waarin leerlingen boven of onder (de statistische) verwachting presteren, gegeven het IQ, het leerlinggewicht en de sekse van de leerling en gegeven de samenstelling van de klas qua IQ, leerlinggewicht en sekse. Nagegaan wordt dan of er een omslagpunt te vinden is in de klassengrootte waarboven de prestaties van leerlingen achterblijven bij vergelijkbare leerlingen uit de kleinere klassen.

In Figuur 1 is de samenhang tussen klassengrootte en rekenprestaties van de jaargroep-2 leerlingen afgebeeld.

In Figuur 1 zijn op de horizontale as 7 verschillende klassengroottes afgebeeld: 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, en 35 of meer leerlingen. Op de verticale as loopt de effectgrootte. Als referentiepunt is gekozen: het gemiddelde niveau van klassen ter grootte van 20-24 en 25-29 leerlingen. Daarom ligt de staaf voor klassen met 25-29 leerlingen ook net zoveel onder de as als de staaf voor klassen met 20-24 leerlingen er boven ligt. Witte balken



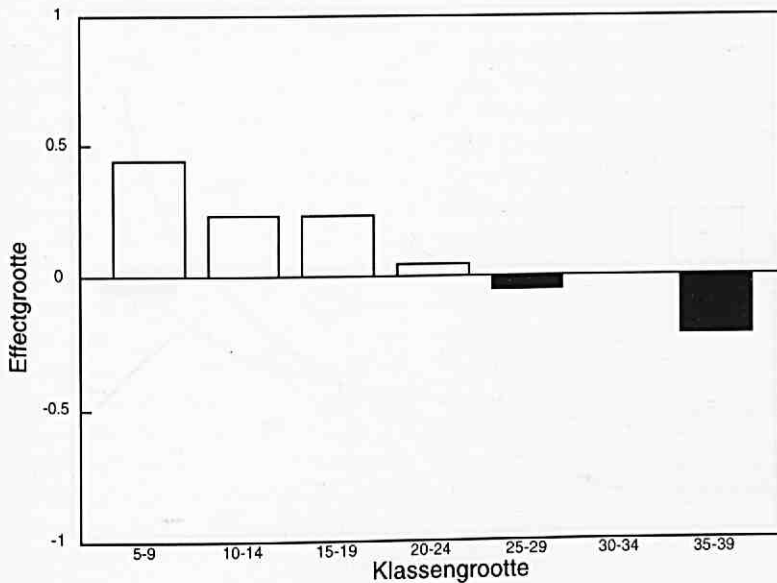
Figuur 1. Klassengrootte en rekenvaardigheid, jaargroep 2 (naar Bosker & Hox, 1996)



Figuur 2. Klassengrootte en rekenvaardigheid, jaargroep 4, 6 en 8 (naar Bosker & Hox, 1996)

geven aan dat de verschillen niet en zwarte bal-
 ken dat de verschillen wel statistisch signifi-
 cant zijn. Statistische analyses laten zien dat
 jaargroep-2 leerlingen in klassen ter grootte
 van 25 of meer leerlingen minder goed rekenen
 dan leerlingen in kleinere klassen dan we, op
 basis van hun IQ, leerlinggewicht, sekse en de
 samenstelling van de klas waarin ze zitten, zou-
 den mogen verwachten. Afgezet tegen de

rekenprestaties van de klassen met 20-24 leer-
 lingen blijven leerlingen in klassen met 25-29
 leerlingen -0.25 , leerlingen in klassen met 30-
 34 -0.18 en leerlingen in klassen met 35 of
 meer leerlingen -0.26 achter. Het rekenresul-
 taat van de kinderen in klassen met 10-14 leer-
 lingen lijkt boven dat van de overige klassen uit
 te steken, maar het gaat hier om een verschil
 dat niet de statistische toets der significantie



Figuur 3. Klassengrootte en taalvaardigheid, jaargroep 2 (naar Bosker & Hox, 1996)

kan doorstaan. Met andere woorden: hier kan heel goed toeval in het spel zijn in plaats van systematiek.

In Figuur 2 zijn de resultaten weergegeven voor de jaargroepen 4, 6 en 8.

Bij de hogere jaargroepen treedt het omgekeerde op van wat bij jaargroep 2 gebeurt. Leerlingen in klassen met meer dan 20 leerlingen doen het qua rekenprestaties boven verwachting beter dan leerlingen in kleinere klassen. De leerlingen in klassen met 30-34 leerlingen komen zelfs tot de hoogste scores: vergeleken met de klassen ter grootte van 20-24 leerlingen is hun score 0.10 hoger. De klassen met 35 of meer leerlingen blijven vervolgens weer -0.04 achter bij de gekozen referentieklassen (20-24 en 25-29 leerlingen).

In Figuur 3 zijn de taalprestaties van de leerlingen uit jaargroep 2 afgebeeld in samenhang met de klassengrootte.

Qua taalprestaties van de jaargroep-2 leerlingen blijken er twee breekpunten te zijn: bij klassen met 25-29 leerlingen (die -0.09 achterblijven bij klassen met 20-24 leerlingen) en bij klassen met 35 of meer leerlingen (de achterstand is hier -0.27). Ook hier geldt weer dat het verleidelijk is aan de taalprestaties in de hele kleine klassen aandacht te besteden, maar zulks is niet gerechtvaardigd, omdat het hier 'toeval-

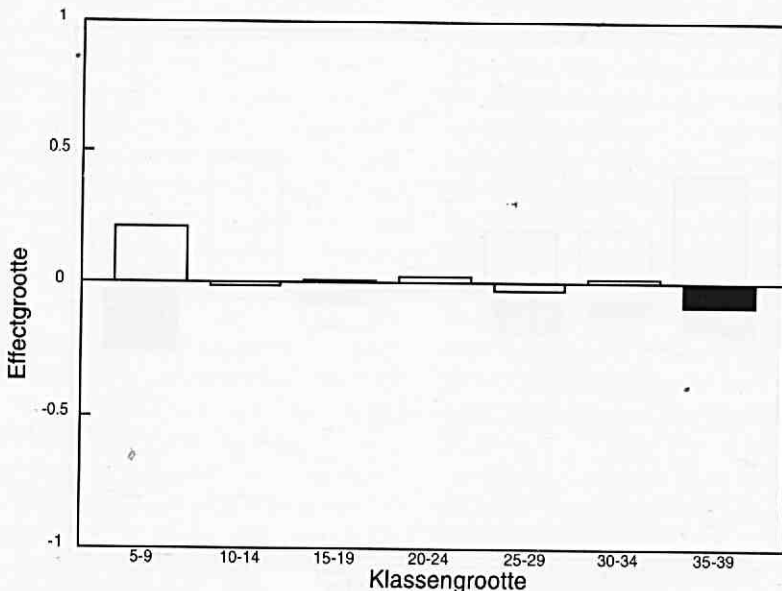
lige' afwijkingen kan betreffen.

In Figuur 4 is de samenhang tussen klassengrootte en de taalprestaties van de leerlingen uit de jaargroepen 4, 6 en 8 weergegeven.

Wat betreft de taalprestaties van de leerlingen in de jaargroepen 4, 6 en 8 is het enige statistisch significante breekpunt de achterstand van de leerlingen in klassen met 35 of meer leerlingen (deze blijven -0.11 achter op de leerlingen uit klassen met 20-24 leerlingen).

De gevolgen van klassengrootte voor de leerkracht

In de meta-analyses van Glass c.s. is ook gekeken naar de relatie klassengrootte en leerkrachtfunctioneren. De auteurs maken onderscheid tussen het algemene leerkrachtfunctioneren (waaronder: organisatie en planning, 'morale', houding tegenover leerlingen, tevredenheid, 'workload', afwezigheid, professionele groei) en instructiegedragingen (zoals: kennis omtrent leerlingen, hoeveelheid directe interactie, hoeveelheid aandacht voor individuele leerlingen, doelgerichtheid, gebruik van leermiddelen, adaptief onderwijs, contacten met ouders, gebruik van ruimte, innovatie e.d.). In totaal zijn 30 studies samengevat naar het algemene leerkrachtfunctioneren en 155 naar



Figuur 4. Klassengrootte en taalvaardigheid, jaargroep 4, 6 en 8 (naar Bosker & Hox, 1996)

instructiegedrag. In Figuur 5 zijn de resultaten van de geanalyseerde studies afgebeeld.

In Figuur 5 zijn twee lijnen getrokken. Voor beide lijnen geldt dat het willekeurige referentiepunt wederom een klas met 30 leerlingen betreft. De lijn die rechtsonder eindigt geeft de gevonden samenhang weer tussen klassengrootte en het algemene leerkrachtgedrag. Vergelijken met een klas van 30 leerlingen is de effectgrootte voor een leerkracht in een klas met 20 leerlingen qua tevredenheid, werkbelasting, absentie, e.d. 0.47. Anders dan in de voorgaande figuren blijkt klassengrootte ook boven de 30 leerlingen een duidelijke samenhang te vertonen met het leerkrachtgedrag: de effectgrootte voor een klas met 35 leerlingen is -0.13.

De tweede lijn in de grafiek geeft het verband tussen de klassengrootte en het instructiegedrag weer. De lijn loopt redelijk vlak in de buurt van de klassengroottes van 20 tot 40 leerlingen: er zijn verschillen ten gunste van de kleinere klassen binnen deze range, maar die verschillen zijn beperkt. De samenhang verandert echter snel naarmate de klassenomvang verder afneemt: de kleine klassen doen het duidelijk beter in termen van instructiegedrag.

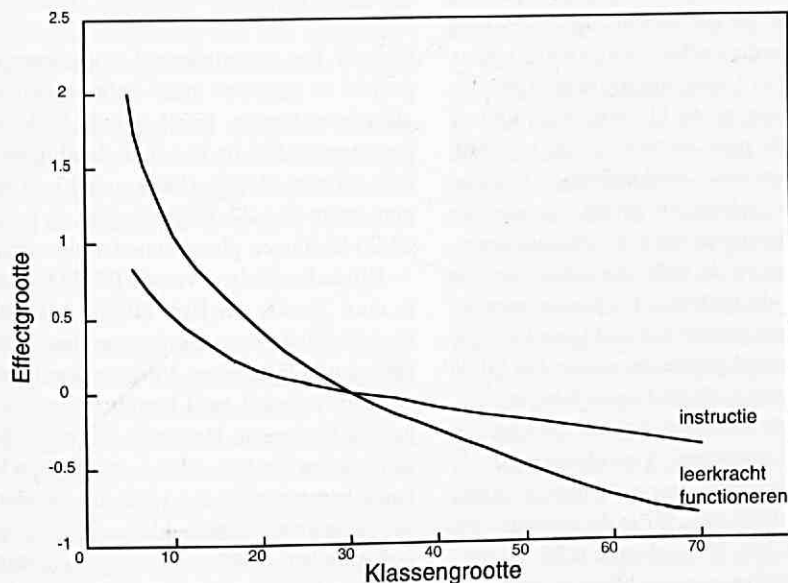
Ook ten aanzien van het leerkrachtfunctioneren zijn de PRIMA-gegevens geanalyseerd door Bosker en Hox (1996). De leerkrachten van de

jaargroepen 1, 3, 5 en 7 zijn namelijk bevraagd naar hun arbeidssatisfactie en hun onderwijsattitude (zie bijv. Overmaat & Ledoux, 1996):

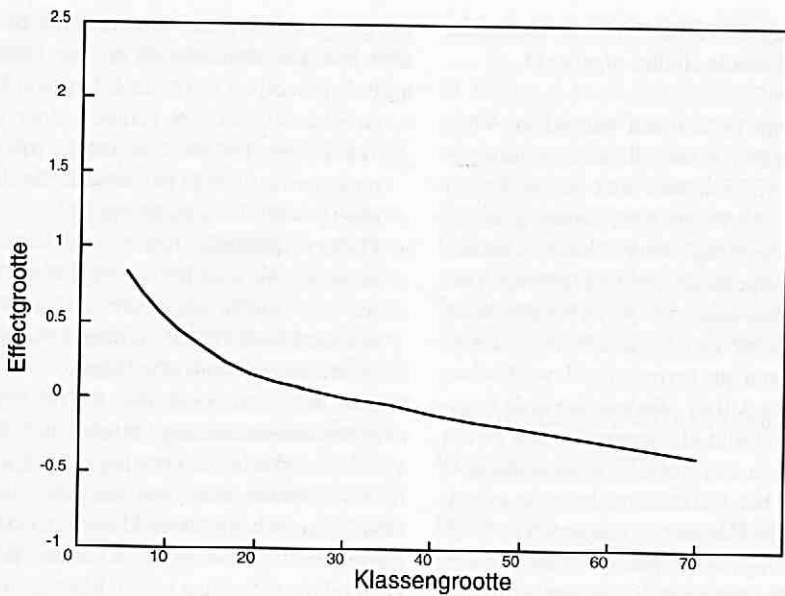
- Arbeidssatisfactie is gemeten door middel van vragen als 'ik voel me in mijn werk ondergewaardeerd' en 'ik vind de sfeer bij ons op school bijzonder prettig'.
- Onderwijsattitude betrof een schaal met items als 'ik vind het moeilijk eisen te stellen aan zwakke leerlingen' en 'ik heb vaak een hard hoofd in het vermogen van sommige leerlingen om wat te leren'.

Omdat niet is uit te sluiten dat de scores op deze schalen samenhang vertonen met de sekse van de leerkracht, de ervaring e.d., zijn dergelijke variabelen weer constant gehouden alvorens de samenhang tussen klassengrootte en het leerkrachtfunctioneren te schatten. Er blijkt geen relatie te bestaan tussen klassengrootte en de arbeidssatisfactie. Wel treedt er een verband op tussen de klassengrootte en de onderwijsattitude.

Het verband tussen klassengrootte en de onderwijsattitude blijkt met name veroorzaakt te worden door de afwijkende opvattingen van leerkrachten die klassen met minder dan 10 leerlingen onder hun hoede hebben. Aangezien dit een relatief zeldzame groep betreft, is er geen reden aan deze bevinding verder aandacht te schenken.



Figuur 5. Relatie tussen klassengrootte en instructie en klassengrootte en leerkracht functioneren (naar Glass c.s., 1982)



Figuur 6. Relatie tussen klassengrootte en affectief functioneren van leerlingen (naar Smith & Glass, 1980)

Klassengrootte en andere effecten

Klassengrootte en het affectief functioneren van leerlingen

In de meta-analyse van Glass c.s. (1982) is ook gekeken naar het affectief functioneren van leerlingen. Daarover is apart gerapporteerd door Smith en Glass (1980). Ze maken gebruik van 59 studies op dit terrein, die betrekking hebben op meerdere affectieve uitkomstmaten. Zo ontstaan er 172 vergelijkingsmogelijkheden tussen kleine en grote klassen, waarbij o.a. gekeken wordt naar motivatie, self-concept, houding tegenover leerkrachten, houding tegenover de school/de groep, geestelijke gezondheid, divergent denken, studiegewoonten, gerichtheid op de taak, discipline, apathie etc. 85% van de onderzochte studies rapporteert effecten ten gunste van de kleine klas. De resultaten zijn veel geprononceerder dan bij de cognitieve maten zoals uit Figuur 6 blijkt.

In Figuur 6 is als referentiepunt een klas van 30 leerlingen genomen. Vervolgens zijn de andere klassengroottes daartegen afgezet in termen van effectgroottes. Zo is de effectgrootte voor een klas met 20 leerlingen 0.13. Anders dan bij de cognitieve maten blijkt er geen volledige afname van het effect te zijn bij toene-

mende klassengrootte: hoe groter de klas des te lager de scores op de attitudeschalen (voor een klas met 40 leerlingen is de effectgrootte -0.08).

Voor de experimenteel opgezette onderzoeken die recent zijn uitgevoerd, zijn de resultaten minder spectaculair, maar dit effect wordt gecompenseerd doordat in het primair onderwijs de effecten sterker zijn dan in het secundair onderwijs.

Ook in het experimenteel opgezette STAR-project is gekeken naar opbrengsten in het affectieve domein. In dit opzicht traden echter geen verschillen op tussen de leerlingen uit de drie soorten klassen (klein met 13-17 leerlingen, groot met 22-25 leerlingen, en groot met 22-25 leerlingen plus een onderwijsassistent).

Bij de heranalyse van de PRIMA-gegevens is door Bosker en Hox (1996) ook gekeken naar de relatie klassengrootte en het welbevinden van de leerlingen. Voor de leerlingen van de jaargroepen 2 en 4 betreft het een oordeel van de leerkracht. De leerlingen uit de hogere jaargroepen hebben zelf een vragenlijst betreffende hun welbevinden ingevuld. Wederom is de samenhang tussen klassengrootte en het welbevinden onderzocht onder constanthouding van IQ, sekse, leerlinggewicht en de klas-samenstelling.

Voor de jaargroepen 2 en 4 blijkt er geen samenhang te bestaan tussen het welbevinden en de klassengrootte. Voor jaargroep 6 en 8 blijkt het welbevinden van de leerlingen in de grotere klassen (met 30 of meer leerlingen) boven verwachting positief te zijn. De effectgrootte bedraagt ca. 0.20.

Klassengrootte en de samenhang met zittenblijven en verwijzing naar het speciaal onderwijs

In verband met de kwestie van de kosteneffectiviteit van verkleining van de klassengrootte is recent uitgevoerd onderzoek in Nederland (Van der Werf, Reezigt & Guldmond, 1996) van belang. Dit laat zien dat juist in de kleinere klassen het zittenblijven vaker voorkomt, ook na controle voor achtergrondkenmerken van de leerlingen. De auteurs vermelden echter niet hoe groot de gevonden effecten zijn. Het tegenargument dat het hier niet om een direct effect van klassengrootte kan gaan, daar zittenblijven het gevolg is van een beslissing van de leerkracht waarbij objectieve criteria niet voor 100% de beslissing bepalen, geldt echter ook de volgende STAR-uitkomsten.

De STAR-onderzoekers betogen namelijk dat het aantal zittenblijvers in de kleine klassen geringer was (19.8 procent tegenover 27.4 in de grote klassen), zodat de kosten van de kleine klassen uiteindelijk weer gedrukt worden. De onderzoekers komen uiteindelijk tot de conclusie dat, alhoewel verkleining van de klassengrootte duurder is dan een onderwijs-assistent, het eerste kosteneffectiever is (Word, 1990, p. 29).

Het inspectie-onderzoek (Inspectie van het Onderwijs, 1995) liet zien dat verwijzingen naar het speciaal onderwijs iets frequenter voorkomen bij de scholen met de grotere klassen.

Conclusies

Welke conclusies zijn uit het hiervoor gepresenteerde overzicht aangaande klassengrootte in het basisonderwijs te trekken?

In de eerste plaats hebben we geconstateerd dat de klassen in het Nederlandse basisonderwijs, met 25.7 leerlingen gemiddeld, internationaal gezien groot zijn. Een belangrijke oorzaak lijkt gelegen te zijn in de geringe uitgaven voor het onderwijs in Nederland. Opvallend is voorts dat zeer grote klassen in enkele landen

niet in basisscholen kunnen voorkomen, omdat er wettelijk vastgestelde maxima worden gehanteerd. In Nederland zit één op de drie leerlingen nog in een klas die meer dan 30 leerlingen bevat.

Alhoewel meer recent onderzoek wenselijk zou zijn, kan op grond van Glass c.s. geconcludeerd worden, dat er sprake is van een negatieve samenhang tussen klassengrootte en het didactisch handelen van de leerkracht. Voorts is vermeldenswaard dat leerkrachten volgens het Nederlandse onderzoek niet echt lijken te 'lijden' onder grote klassen. Immers, hun arbeidssatisfactie en onderwijsattitude vertoont geen samenhang met de omvang van de klas.

Er zijn enige indicaties dat klassenverkleining of reductie van de leerling/leerkracht ratio gunstig uitpakt in termen van minder zittenblijven en minder verwijzing naar het speciaal onderwijs. Er zij echter op gewezen, dat beide verschijnselen sterk afhankelijk zijn van beslissingen van de leerkracht en niet voor 100% te herleiden zijn op het werkelijke functioneren van leerlingen.

Het overzicht van het onderzoek naar de samenhangen tussen klassengrootte en het affectief functioneren van leerlingen laat eenduidige conclusies niet toe. Directe relaties worden zelden of nooit gerapporteerd. Alhoewel de meta-analyse van Glass c.s. een verband laat zien, worden noch in het STAR-project noch in de analyse van het PRIMA-cohort samenhangen op dit gebied aangetoond.

Tenslotte vatten we de uitkomsten samen voor de relatie klassengrootte en leerprestaties. Er lijkt een trend waarneembaar dat met name voor de jongere leerlingen kleine klassen gunstig uitpakken. Maar is er hier werkelijk sprake van een oorzaak en gevolg relatie? De evidentie gevonden in het STAR-project is methodologisch gezien het sterkst en wijst duidelijk in het voordeel van kleine klassen. De tegenwerping dat juist het functioneren in een experimentele setting het effect veroorzaakt heeft en niet de klassenverkleining zelf, zou men kunnen pareren door te verwijzen naar de algemene trend die uit de meta-analyse van Glass c.s. blijkt, en naar de bevindingen uit Indiana (Prime Time) en Nederland (het PRIMA-cohort). Deze zijn immers ontleend aan gegevens die in een natuurlijke setting verzameld zijn.

De effecten die gerapporteerd zijn voor de jaargroep-2 leerlingen in Nederland zijn behoorlijk groot. Om een vergelijkingsbasis te geven: leerlingen in de grotere klassen (met 25 of meer leerlingen) hebben een achterstand die vergelijkbaar is met die van de arbeiderskinderen ten opzichte van de autochtone leerlingen uit de midden en hogere sociaal-economische milieus.

Zijn voor oudere leerlingen de resultaten minder geprononceerd, voor de jongere leerlingen lijkt een forse reductie van de klassengrootte samen te gaan met betere prestaties. Correa (1993) maakt aannemelijk dat het pas bij forse reducties van de klassengrootte voor rationeel handelende leerkrachten nuttig wordt meer tijd te besteden aan individuele hulp van leerlingen: tijd en vervolgens voldoende kwaliteit van de mogelijk geworden leerkracht-leerling interactie is dan de crux waarom het zou draaien.

Klassenverkleining in de onderbouw lijkt het meest kansrijk, maar onduidelijk is nog in welke jaargroep de klassenverkleining het meest effectief kan worden ingezet. De minder beklievende effecten van Prime Time ten opzichte van STAR lijken te wijzen op een verkleining van de klassengrootte voor de allerejongsten, maar bedacht dient te worden dat in STAR ook een veel sterkere reductie van de klassenomvang is ingezet (naar gemiddeld 15) dan in Prime Time (naar gemiddeld 19).

Desalniettemin zijn vele auteurs niet zeker over de kosteneffectiviteit van klassenverkleining. Met name Slavin (1989) spreekt de sterkste twijfels uit. Zelf propageert hij de omnibus aanpak van Success for All: een speciale lees-aanpak, 'tutors', een voorschools traject, coöperatief leren, regelmatige voortgangscontrole (eens per 8 weken), gezinsondersteuning en deskundige hulp voor het schoolteam. Het probleem met zijn alternatieve aanpak is echter, dat een onafhankelijke evaluatie van het Success for All programma nog niet heeft plaatsgevonden. De door de auteur zelf gerapporteerde resultaten lijken gunstig (Madden, Slavin, Karweit, Dolan & Wasik, 1993; Slavin, 1996), maar zijn wellicht sterk opgeblazen door mogelijke Hawthorne-effecten. Slavin (1996) spreekt zelf de veronderstelling uit dat voor risicoleerlingen op scholen met veel achterstandsleerlingen de omnibus aanpak van Suc-

cess for All het meest effectief is, terwijl voor risicoleerlingen op andere scholen enkel en alleen 'Reading Recovery' voldoende is. De benadering van Slavin heeft met klassenverkleining gemeen, dat een reductie van de leerling/leerkracht ratio noodzakelijk is. Kennelijk is dat toch de algemene cruciale factor.

Bosker en Hox (1996) laten zien dat scholen hun formatie, bedoeld voor achterstandsbestrijding, voor een belangrijk deel inzetten voor een algehele reductie van de klassenomvang in alle jaargroepen. Deze bevinding, gecombineerd met de magere resultaten van reductie van de klassengrootte voor de prestaties van leerlingen van 8 jaar en ouder, zou een mogelijke verklaring kunnen bieden voor de tot nu toe weinig bevredigende resultaten van het OnderwijsvoorrrangsBeleid (zie: Mulder, 1996). Het roept ook de vraag op of er niet toch scholen zijn die gericht hun formatie gebruiken, hetzij voor een forse klassenverkleining, hetzij voor een Slavin-achtige aanpak, voor de jongere leerlingen, en wat daar dan het effect van is.

Gericht – bij voorkeur experimenteel – onderzoek naar de effectiviteit van deze aanpakken op reguliere scholen en op scholen met veel achterstandsleerlingen zou hier uitsluitsel kunnen geven.

Voorts is nader onderzoek gewenst naar combinatieklassen in relatie tot de klassengrootte problematiek. De 'best-evidence' reviews van Veenman (1995, 1996) laten zien dat combinatieklassen niet ongunstig uitpakken, noch voor het cognitief noch voor het affectief functioneren van leerlingen. En onder verwijzing naar vernieuwende scholingsconcepten wordt regelmatig het einde gevraagd van het leerstofjaarklassensysteem. Maar zowel groot-schalig goed opgezet onderzoek naar combinatieklassen in Nederland (gebruik makend van adequate statistische analysemethoden) als goed opgezet evaluatie-onderzoek naar de effecten van vernieuwingsscholen als Dalton, Jenaplan, Montessori e.d. ontbreekt. Klassenverkleining lokt leeftijdsheterogene combinatieklassen of andere groeperingsvormen uit. Maar wat voor de leerkracht haalbaar en voor de leerling wenselijk is, blijft voor alsnog een open vraag. De discussie over klassengrootte in het basisonderwijs mag dan zijn voorlopig einde bijna lijken te hebben bereikt (Commissie Kwalitatieve Aspecten van Groeps-grootte,

1996), de discussie over en het onderzoek naar een pedagogisch-didactisch verstandig, efficiënt, effectief, en gelijke kansen bevorderend gebruik van formatie in basisscholen kan ons inziens nu beginnen.

Noten

1. Met de introductie van de Wet op het Basisonderwijs, is het woord klas vervangen door groep, zodat het usance is geworden om waar vroeger werd gesproken over klassengrootte (bijv. Van Gelder, 1959; Warries, 1985) nu te spreken over groeps-grootte. Deze term is echter misleidend, omdat één (combinatie)klas meerdere jaargroepen kan omvatten. Derhalve houden we vast aan de term klassengrootte.
2. De effectgrootte is in navolging van Cohen (1988), uitgedrukt als het effect van een conditie ten opzichte van een referentiegroep in termen van de residuele (dus niet-triviale) binnen-groepen variantie. Effectgroottes van ca. 0.10 worden door Cohen klein genoemd, maar kunnen desalniettemin relevant zijn (zie ook: Bosker & Witziers, 1996).
3. Citaat ontleend aan: Aitkin en Longford (1986, p. 42).

Literatuur

Achilles, C.M., Nye, B.A., Zaharias, J.B., & Fulton, B.D. (1993). Creating successful schools for all children: a proven step. *Journal of School Leadership*, 3, 606-621.

Aitkin, M.A., & Longford, N.L. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society, A*, 149, 1-43.

Akerhielm, K. (1995). Does class size matter? *Economics of Education Review*, 14(3), 229-241.

Batenburg, T. van, Koster, D., & Ros, A. (1989). Kleinere groepen, betere prestaties? In J. Scheerens & J.C. Verhoeven (Red.), *Schoolorganisatie, beleid en onderwijskwaliteit* (pp. 55-62). Bijdragen aan de onderwijsresearch. Lisse: Swets & Zeitlinger.

Bosker, R.J., & Hox, J. (1996). Klassengrootte en functioneren van leerlingen en leerkrachten; een onderzoek op basis van de PRIMA-cohort gegevens. In Commissie Kwalitatieve Aspecten

van Groeps-grootte, *Klassenverkleining* (pp. 107-118). Den Haag: SDU.

Bosker, R.J., & Sniijders, T.A.B. (1990). Statistische aspecten van multi-niveau onderzoek. *Tijdschrift voor Onderwijsresearch*, 15, 317-329.

Bosker, R.J., & Witziers, B. (1996) *The magnitude of school effects*. New York: AERA-paper.

Cohen, J. (1988) *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum.

Coleman, J.S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. Washington, DC: US Government Printing Office.

Commissie Evaluatie Basisonderwijs (1994). *Zicht op kwaliteit. Evaluatie van het basisonderwijs*. Den Haag: SDU.

Commissie Kwalitatieve Aspecten van Groeps-grootte (1996). *Klassenverkleining*. Den Haag: SDU.

Correa, H. (1993). An economic analysis of class size and achievement in education. *Education Economics*, 1(2), 129-135.

Finn, J.D., & Achilles, C.M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal*, 27(3), 557-577.

Gelder, L. van (1959). De klasgrootte bij het lager onderwijs. *Pedagogische Studiën*, 36, 498-519.

Glass, G.V., & Smith, M.L. (1978). *Meta-analysis of research on the relationship of class size and achievement*. San Francisco: Far West Laboratory.

Glass, G.V., Cahen, L.S., Smith, M.L., & Filby, N.N. (1982). *School class size*. London: Sage Publications.

Glopper, K. de, & Otter, M.E. (1993). *Nederlandse leesprestaties in internationaal perspectief*. Amsterdam: SCO.

Hanushek, E.A. (1986). The economics of schooling: production and efficiency in public schools. *Journal of Economic Literature*, 24, 1141-1177.

Hedges, L.V., Laine, R.D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher*, 23(3), 5-14.

Inspectie van het Onderwijs (1995). *Groeps-grootte in het basisonderwijs*. Den Haag: SDU.

Jungbluth, P., Peetsma, T., & Roeleveld, J. (1996). Leerlingprestaties en leerlinggedrag in het primair onderwijs. Ubbergen: ITS/SCO-Kohnstamm/Tandem Felix.

Larkin, A.I., & Keeves, J.P. (1984). *The class size*

- question: A study at different levels of analysis. Hawthorn: Australian Council for Educational Research.
- Madden, N.A., Slavin, R.E., Karweit, N.L., Dolan, L., & Wasik, B.A. (1993). Succes for all: Longitudinal effects of a restructuring program for inner city elementary schools. *American Educational Research Journal*, 30(2), 123-148.
- McGiverin, J., Gilman, D., & Tilitzki, C. (1989). A meta-analysis of the relation between class size and achievement. *The Elementary School Journal*, 90(1), 47-56.
- Ministerie van OC&W (1996). *Kerngegevens primair onderwijs*. Zoetermeer: MOCW.
- Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The Future of Children*, 5(2), 1-13.
- Mulder, L. (1996). *Meer voorrang, minder achterstand? Het onderwijsvoorrangsbeleid getoetst*. Academisch proefschrift, Katholieke Universiteit Nijmegen.
- Organization for Economic Co-operation and Development (1995). *Education at a Glance*. Parijs: OECD.
- Overmaat, M., & Ledoux, G. (1996). *School- en klaskenmerken basisonderwijs en speciaal onderwijs*. Ubbergen: SCO-Kohnstamm/Tandem Felix.
- RISE (1994). *Class size regulation. A dossier of international comparisons*. London: The Research and Information on State Education Trust.
- Ross, K.N., & Postlethwaite, T.N. (1994). Differences among countries in school resources and achievement. In W.B. Elley (Ed.), *The IEA study of reading literacy: Achievement and instruction in thirty-two school systems*. Oxford: Elsevier Science Ltd.
- Slavin, R.E. (1989). Achievement effects of substantial reductions in class size. In R.E. Slavin (Ed.), *School and classroom organization* (pp. 247-257). Hillsdale, N.J.: Lawrence Erlbaum.
- Slavin, R.E. (1996). *Education for All*. Lisse: Swets & Zeitlinger.
- Smith, M.L., & Glass, G.V. (1980). Meta-analysis of research on class size and its relationship to attitudes and instruction. *American Educational Research Journal*, 17(4), 419-433.
- Veenman, S. (1995). Cognitive and noncognitive effects of multigrade and multi-age classes: A best-evidence synthesis. *Review of Educational Research*, 65(4), 319-382.
- Veenman, S. (1996). Effects of multigrade and multi-age classes reconsidered. *Review of Educational Research*, 66(3), 323-340.
- Vierke, H. (1995). *De PRIMA-toetsen gecalibreerd*. Nijmegen: ITS.
- Warries, E. (1985). *Klassegrootte en onderwijskwaliteit*. Enschede: Universiteit Twente.
- Werf, M.P.C. van der, & Guldemon, H. (1994). *De tweede meting van de OVB evaluatie. School- en klaskenmerken in schooljaar 1990/1991*. Groningen: RION/ITS/LEO.
- Werf, M.P.C. van der, Reezigt, G.J., & Guldemon, H. (1996). *Zittenblijven in het basisonderwijs: omvang, effecten en relaties met leerling- en schoolkenmerken*. Groningen: GION.
- Word, E. (1990). *Student/Teacher Achievement Ratio (STAR). Tennessee's K-3 Class Size Study*. Washington: ERIC Report ED320 692.

Auteur

R.J. Bosker is als universitair hoofddocent werkzaam bij de vakgroep Onderwijsorganisatie & -management van de faculteit der Toegepaste Onderwijskunde van de Universiteit Twente.

Correspondentie-adres: Vakgroep O&M, Faculteit TO, Universiteit Twente, Postbus 217, 7500 AE Enschede.

Abstract

The end of class struggle? Background to the discussion on class size in primary education.

R.J. Bosker. *Pedagogische Studiën*, 1997, 74, 210-227.

Recent surveys show that class size in Dutch elementary schools is on average large as compared to other countries, and moreover varies enormously between as well as within schools. This article presents a review of foreign studies into the effects of class size on the functioning of teachers and students. Moreover, it also presents results from a recent large scale Dutch research study in this area. This study shows that there is an association between class size and student achievement in the areas of language and arithmetic. Grade-2 students (five year olds) in classes with more than twenty five students are more likely to be

underachievers than comparable students from smaller classes. For grades 4, 6, and 8 the results are less clear cut (showing, for example, a relatively high level of overachievement for arithmetic in classes sized 30-34), but there is an indication that class sizes of thirty five and up may be detrimental to student achievement.

Neither in the affective domain, nor with respect to the functioning of teachers this Dutch study could show a relationship with class size.

Comparing the results with programs like 'Reading Recovery' and 'Succes for All' it is hypothesized that not class size per se is the factor causing the effects mentioned, but a more general factor, namely the teacher-to-pupil ratio.

Boekbespreking

J.F.A. Braster

**De identiteit van het openbaar onderwijs
Wolters-Noordhoff, Groningen 1996**

495 pagina's f 69.50 ISBN 90 01 16020 4

De laatste jaren is de vraag naar de identiteit van het openbaar onderwijs sterk in de belangstelling gekomen. Deze belangstelling komt niet uit de lucht vallen. In de eerste plaats vragen wetenschappers en beleidsmakers zich al enige tijd af hoe het komt dat, ondanks de toenemende secularisatie, het marktaandeel van het openbaar onderwijs niet zichtbaar toeneemt. Blijkbaar heeft het bijzonder onderwijs zijn aantrekkingskracht, ook voor niet-kerkelijke ouders, niet verloren. In de tweede plaats heeft, naast een aantal andere demografische factoren, de toenemende culturele pluriformiteit van onze samenleving de druk op het openbaar onderwijs vergroot om 'kleur' te bekennen. Als gevolg daarvan spreekt men in kringen van het openbaar onderwijs bijvoorbeeld niet meer van 'neutraliteit', maar van 'actieve pluriformiteit'. En niet in de laatste plaats heeft de politiek weer aandacht voor de positie van het openbaar onderwijs. Al voor het aantreden van het paarse kabinet lagen er bijvoorbeeld verschillende voorstellen tot verdere bestuurlijke verzelfstandiging van het openbaar onderwijs.

Vanuit een betrokkenheid bij de vormgeving van het openbaar onderwijs, maar tegelijkertijd met voldoende wetenschappelijke distantie geeft Braster in zijn recent verschenen proefschrift een uitvoerig antwoord op de identiteitsvraag. Deze uitvoerigheid uit zich niet alleen in de omvang (rond de 500 pagina's), maar ook in de diepte van zijn studie. Daarmee is zijn studie in één klap een standaardwerk geworden waar betrokkenen bij het openbaar onderwijs en ook onderzoekers in de toekomst niet meer omheen kunnen.

Na twee inleidende hoofdstukken over de problemstelling en de onderzoeksopzet van zijn studie behandelt Braster in vier helder geschreven hoofdstukken de geschiedenis van de openbare school. Hij deelt die geschiedenis in vier perioden in. De eerste periode omvat de lange voorgeschiedenis tot 1806, waarin de openbare school nog een specifieke religieuze

signatuur had (eerst katholiek, later gereformeerd). Dan volgt een periode tot 1857, waarin de openbare school een algemeen-christelijk karakter droeg. In de derde periode, die duurt tot de jaren zestig van deze eeuw, sloeg de openbare school de weg in naar 'ware neutraliteit'. Tot slot, en dat is de laatste periode, doemen kansen op voor een pluriforme openbare school.

Wat Braster in die geschiedenis behandelt is niet verrassend, maar wel essentieel voor zijn verdere onderzoek. Hij ontwikkelt daaruit zijn kameleonhypothese. Hij laat namelijk zien dat de dominante verschijningsvorm van de openbare school in een periode zich ontwikkelt als reactie op de dominante denkbeelden van de tijd daarvoor. In lijn daarmee ontwikkelt hij zijn stelling dat de openbare school zich in het algemeen aanpast aan de sociaal-culturele context. De school neemt als een kameleon de kleur van de omgeving aan. Daarom kan men volgens hem niet van dé openbare school spreken, maar bestaan er in elke tijd steeds verschillende typen openbare onderwijsinstellingen naast elkaar. De specifieke identiteit daarvan kan steeds verklaard worden uit de karakteristieken van de omgeving.

In de daarop volgende hoofdstukken zet Braster deze stelling kracht bij door een breed opgezette empirische studie naar de identiteit van het openbaar onderwijs. Daarvoor maakt hij een onderscheid tussen de formele, de gepercipieerde en empirische identiteit. De formele identiteit heeft betrekking op de kenmerken zoals die in de wetgeving vastgelegd zijn. De gepercipieerde identiteit vervolgens slaat op de wijze waarop relevante betrokkenen de identiteit definiëren. Tenslotte omvat de empirische identiteit de concrete activiteiten waarmee in de school- of lespraktijk daadwerkelijk vorm gegeven wordt aan de identiteit van het openbaar onderwijs.

Uitgaande van de huidige formele identiteit, die bestaat uit de publiekrechtelijke bestuursvorm, algemene toegankelijkheid en de neutraliteit of pluriformiteit, schetst Braster eerst hoe beleidsvoerders en -beïnvloeders deze formele kenmerken percipiëren. Ook dit hoofdstuk brengt nog geen verrassende zaken aan het licht. De studie begint pas in het achtste hoofdstuk de zaken wat te kantelen. In dat

hoofdstuk wordt verslag gedaan van een onderzoek naar de identiteit van de basisschool onder 382 schoolleiders op basis van trefwoorden. Daarnaast is nog gekeken naar de naamgeving van de school als mogelijke indicator voor identiteit. Dit onderzoek leverde zo'n 1500 trefwoorden op, die nader geanalyseerd zijn. Daarnaast zijn aan elke school een aantal school- en contextkenmerken toegevoegd. Daaruit kwam naar voren dat de drie formele kenmerken in de ogen van de schoolleiders van de openbare scholen nauwelijks van betekenis zijn. De overgrote meerderheid schoolleiders van zowel openbare als bijzondere scholen definiëren de identiteit van hun school in onderwijskundige kenmerken. Andere definities worden slechts in beperkte mate naar voren gebracht. Dit geldt bijvoorbeeld voor omschrijvingen in levensbeschouwelijke grondslag. Alleen het protestants-christelijk onderwijs doet dit nog in zekere mate (23% van de schoolleiders). Alleen daar waar het openbaar onderwijs een minderheidspositie inneemt, is bij openbare schoolleiders een sterk besef van de openbare grondslag.

In de hoofdstukken negen en tien wordt aandacht geschonken aan de school- en lespraktijk. Door middel van mondelinge interviews met schoolleiders van 85 openbare basisscholen is onderzocht hoe deze scholen de verscheidenheid van levensbeschouwelijke en maatschappelijke waarden aan de orde stellen. De steekproef is getrokken op basis van een indeling naar negen verschillende sociaal-culturele contexten, waarin politieke kleur, etnische samenstelling en aantal inwoners van de gemeenten de criteria waren. Naast deze indeling naar sociaal-culturele context heeft Braster zijn gegevens ook nog vergeleken met het aanbod naar richting in de betreffende gemeente. Dit criterium is later opgenomen, omdat een aantal recente studies met dit criterium interessante resultaten heeft laten zien. Uit de analyse blijkt nu dat er vier typen openbare scholen onderscheiden kunnen worden: neutrale, traditionele, multiculturele en pluriforme. Hoewel men uit de wet zou kunnen opmaken, dat de pluriforme openbare school de dominante verschijningsvorm zou moeten zijn, is zij dat in werkelijkheid niet. Een kwart van de openbare basisscholen is als zodanig te typeren. Ook de drie andere typen nemen

meer of minder dan een kwart voor hun rekening. Als er naar verbanden wordt gekeken, dan blijkt dat het type openbare school, dat men in de praktijk aantreft, samenhangt met de specifieke sociaal-culturele context. Deze conclusie lijkt niet alleen de kameleonhypothese te bevestigen, maar is ook in lijn met recente organisatiekundige inzichten, zoals in de contingentiebenadering en het neo-institutionalisme naar voren komen.

Braster stopt niet bij de poort van de school, of liever gezegd bij de directiekamer. Hij heeft ook nog eens 237 leerkrachten van 82 openbare scholen, verspreid over het land, schriftelijk bevraagd naar hun lespraktijk. Daaruit komt een beeld naar voren dat niet alle openbare onderwijzers optimaal zijn toegerust voor hun taak in de plurale samenleving. Er kan evenmin volgens Braster gesteld worden dat de aandacht die leerkrachten op openbare basisscholen aan de maatschappelijke verscheidenheid besteden bijzonder sterk is. In totaal zijn zeven kenmerken gevonden die bepalend zijn voor de mate waarin leerkrachten aandacht besteden aan de verscheidenheid in de samenleving. Deze zijn een maatschappij-gerichte didactische aanpak, kennis van wereldgodsdiensten en andere culturen, specifieke opvattingen over onderwijs, een maatschappelijke oriëntatie, het lesgeven in hogere groepen, en het bezitten van cultureel en sociaal kapitaal.

Braster sluit af met de rapportage van een onderzoek onder een representatieve groep van 980 Nederlanders. Aan deze groep is gevraagd naar hun behoefte aan pluriformiteit in het onderwijs. Hoewel de behoefte aan pluriformiteit verschilt tussen mensen die een voorkeur hebben voor openbaar onderwijs of bijzonder onderwijs, is het prefereren van pluriform onderwijs geen synoniem voor het kiezen voor openbaar onderwijs. Hij constateert ook in de voorkeur voor het openbaar onderwijs een verschil in generaties. De oudere generatie blijkt een sterker besef van de identiteit van de openbare school te hebben dan de jongere generatie die is opgegroeid in een sterk ontzuilde maatschappij. Bij de jongere generatie daarentegen lijkt het opleidingsniveau een veel grotere rol te spelen bij de keuze voor een basisschool. Dit zou erop kunnen duiden dat er in het onderwijsbestel nieu-

we breuklijnen zullen ontstaan, die in tegenstelling tot het verleden niet langer verticaal (zuilen) maar horizontaal (sociale klassen) lopen.

Maken we een balans op van de gehele studie, dan kunnen we vaststellen dat dit proefschrift een rijkgevarieerd en degelijk stuk werk is. Braster maakt gebruik van allerlei soorten gegevens en technieken: historische bronnen, telefonische enquête en mondelinge interviews onder schoolleiders, schriftelijke enquête onder leerkrachten en een representatieve groep Nederlanders, gesprekken met beleidsmakers op verschillende niveaus. De enige opvallende afwezige actor vormt het schoolbestuur. Dat is aan de ene kant jammer, omdat uit (eigen) onderzoek is gebleken dat besturen belangrijke identiteitsbewakers zijn. Aan de andere kant wordt dit gemis gecompenseerd door (soms gewaagde) nieuwe benaderingen. Een voorbeeld daarvan is het trefwoorden-onderzoek onder schoolleiders. Niet in de laatste plaats is de studie in een prettige en heldere stijl geschreven, ook waar het meer technische aspecten van het onderzoek betreft. Hoewel een *single issue*-onderwerp, overstijgt de studie door zijn opzet en uitwerking het terrein van het openbaar onderwijs en geeft het een goed inzicht in de werking van het Nederlandse onderwijsbestel zoals dat historisch vorm gekregen heeft.

S. Karsten

Mededelingen

Promoties

Universiteit van Amsterdam

Promovendus: Drs. H.A. van der Stege

Titel proefschrift: Ideeën van ouders. Over de invloed van aanleg, opvoeding en omgeving op de ontwikkeling van hun kind

Promotor: Prof.dr. J.M.A. Hermanns

Datum: 28 april 1997

Universiteit van Amsterdam

Promovendus: Drs. O.C. McDaniel

Titel proefschrift: The effects of government policies on higher education

Promotores: Prof.dr. F.A. van Vught en

Prof. Dr. A.M.L. van Wieringen

Datum: 20 mei 1997

Rijksuniversiteit Leiden

Promovendus: Drs. W.J.M. Van den Broeck

Titel proefschrift: De rol van fonologische verwerking bij het automatiseren van de leesvaardigheid

Promotor: Prof.dr. A.J.J.M. Ruijsseenaars

Co-promotor: Dr. K.P. van den Bos

Datum: 4 juni 1997

Katholieke Universiteit Nijmegen

Promovendus: Drs. A.R.G. Jansen

Titel proefschrift: Construction monitoring systems in the behavioral sciences

Promotor: Prof.dr. E.E.J. de Bruyn

Co-promotores: Dr. J.H.L. Oud en

Dr. J.F.J. van Leeuwe

Datum: 12 mei 1997

Katholieke Universiteit Nijmegen

Promovendus: Drs. G.J.Th. Haselager

Titel proefschrift: Classmates. Studies on the development of their relationships and personality in middle childhood

Promotor: Prof.dr. C.F.M. van Lieshout

Co-promotor: Dr. J.M.A. Riksen-Walraven

Datum: 23 mei 1997

Katholieke Universiteit Nijmegen

Promovendus: Drs. B. Stellingenwerf

Titel proefschrift: Learning the rules of the game: teaching strategies in arithmetic word problem solving.

Promotor: Prof.dr. L. Verhoeven

Co-promotor: Dr. E.C.D.M. van Lieshout

Datum: 24 juni 1997

Universiteit Utrecht

Promovendus: Drs. G. Erkens

Titel proefschrift: Coöperatief probleemoplossen met computers in het onderwijs. Het modelleren van coöperatieve dialogen voor de ontwikkeling van intelligente onderwijs-systemen

Promotor: Prof.dr. G. Kanselaar

Co-promotor: Dr. J.L. van der Linden

Datum: 28 mei 1997

Recent gestart onderzoek

Universiteit van Amsterdam

Titel project: Geïntegreerd taalonderwijs in Amsterdam

Contactpersoon: Drs. Y. Emmelot, SCO-Kohnstamm Instituut

Titel project: PRIMA-cohort, tweede ronde

Contactpersoon: Drs. M. van Erp, SCO-Kohnstamm Instituut

Titel project: Validering Alvabavo

Contactpersoon: Dr. J. Meijer, SCO-Kohnstamm Instituut

Titel project: Succesfactoren terugdringen schoolverzuim

Projectleiding: Drs. M. Derriks en

Drs. E. de Kat

Vrije Universiteit

Titel project: Wanneer is kennis van jonge

kinderen inzake seksualiteit leeftijdsadequaat?

Onderzoekster: Drs. S.N. Brilleslijper-Kater

Universiteit Utrecht

Titel project: Argumentation in computer

supported collaborative environments

Onderzoekster: Drs. A. Veerman

Oratie

Dr. J.J. Hox, benoemd tot hoogleraar in de faculteit der Sociale Wetenschappen aan de Universiteit Utrecht, om werkzaam te zijn op

het vakgebied van de Methoden en Technieken van sociaal-wetenschappelijk onderzoek, heeft op 1 mei 1997 zijn ambt aanvaard met het houden van een oratie, getiteld: *Er is nieuws onder de zon: nieuwe oplossingen voor oude problemen.*

Afscheidscollege

Prof.dr. K. Doornbos heeft op 23 mei 1997 zijn werkzaamheden als hoogleraar Orthopedagogiek aan de Universiteit van Amsterdam beëindigd met een afscheidscollege, getiteld: *Weg van Onderwijs.*

PAO-prijs 1996

Mevr. drs. C.J. Rieffe, prof.dr. W. Koops en dr. M. Meerum Terwogt hebben de PAO-prijs 1996 ontvangen van de Stichting Nederlands Instituut voor Postacademisch onderwijs. Zij ontvingen deze prijs voor hun bijdrage aan het Jaarboek Ontwikkelingspsychologie, Orthopedagogiek en Kinderpsychiatrie 2 (1996-1997) getiteld: *Vroegkinderlijk begrip van mentale processen: de Child's Theory of Mind.*

Inhoud andere tijdschriften

Comenius

17e jaargang, nr. 1, 1997

Thema: Onderwijsbeleid voor de 21e eeuw

Onderwijsbeleid voor de 21e eeuw. Inleiding op het thema, door F. Meijers, W. Wardekker en A. Wesselingh

Onderwijs op weg naar de 21e eeuw; trends en contradicties, door H. Leune

Regeren of dereguleren. Dilemma's van onderwijsbeleid in een nieuwe tijd, door P. Slegers en A. Wesselingh

Op zoek naar een heuristiek voor funderend onderwijsbeleid in de 21e eeuw, door S. Blom en verder:

Seksualiteit als opvoedings- en ontwikkelings-terrein: nieuwe oriëntaties en dilemma's, door Y. te Poel en J. Ravesloot

Een nieuw vergezicht? Een reactie op Y. Leeman, H. Lutz en W. Wardekker 'Inter-cultureel onderwijs en culturele identiteit' door W. Fase

Tijdschrift voor Orthopedagogiek
36e jaargang, nr. 4, 1997

Kinderbescherming als bijdrage tot herstel: op zoek naar een realistisch optimisme, door J.E. Doek

De rol van het basisonderwijs in de groei van lom en mlk, door Y.J. Pijl

Orthopedagoog versus gezondheidszorg-psycholoog, door A.E.J. Burger-Veltmeijer

Ontvangen boeken

Bouwers, H., & Goor, H. van, *Problemen met begrijpend lezen*. Uitgeverij Intro, Baarn, 1997, f 29,90.

Broeck, W. Van den, *De rol van fonologische verwerking bij het automatiseren van de leesvaardigheid* (dissertatie). Rijksuniversiteit Leiden, Leiden, 1997.

Haafte, W. van, Korthals, M., & Wren, Th.(Eds.), *Philosophy of development; reconstructing the foundations of human development and education*. Kluwer Academic Publishers, Dordrecht/Boston/London, 1997, f 79,50.

Haenen, J., *Piotr Gal'perin: Psychologist in Vygotskij's Footsteps*. Nova Science Publishers, New York, 1996.

Heer-Dehue, J. de (Ed.), *Intercultural education and education of migrant children. Practice and perspectives*. Stichting Promotie Talen, Utrecht, 1997, f 37,50.

Hoop, F. de, & Janson, D.J., *Omggaan met verschillen. Adaptief werken aan basisonderwijs*. Uitgeverij Intro, Baarn, 1997, f 29,90.

Klomp, H.A., *De relativiteitstheorie in Nederland. Breekijzer voor democratisering in het interbellum* (dissertatie). Epsilon Uitgaven, Utrecht, 1997, f 42,50.

Luit, J.E.H. van, *Kinderen en vroege (leer)problemen; stimulerende activiteiten voor nul- tot twaalfjarigen met een achterblijvende ontwikkeling*. Graviant Educatieve Uitgaven, Doetinchem, 1997, f 25,-.

Schouwenburg, H.C. & Groenewoud, J.T., *Studieplanning*. Wolters-Noordhoff, Groningen, 1997, f 24,50.

Verhey, F., *Multidisciplinair verslagleggen en interdisciplinair denken*. Uitgeverij SWP, Utrecht, 1997, f 64,50.

Voorlezen op school aan jonge kinderen: een review*

H. Blok

Samenvatting

In kindercentra en in kleutergroepen wordt praktisch dagelijks door leidsters en leerkrachten voorgelezen. Eén van de oogmerken is de bevordering van de taal- en leesontwikkeling van de kinderen. Dit artikel beschrijft een review naar de effecten die het voorlezen op school heeft op de taal- en leesontwikkeling van jonge kinderen (tot acht jaar). In totaal werden elf studies (met twaalf steekproeven) gevonden die het voorlezen op een kindercentrum of op school tot onderwerp hebben. De gecombineerde effectgrootte voor 'mondelinge taal' bedraagt .64, voor 'lezen' .22. Hoewel met name het eerste effect niet te verwaarlozen is, wordt gewaarschuwd voor een al te optimistische interpretatie. De empirische basis is erg smal. Niet alleen is het aantal studies beperkt, maar ook laat een kritische analyse van de studies ernstige tekorten zien, zowel in de opzet als in de rapportage.

1 Inleiding

Op basisscholen wordt veelvuldig aan jonge kinderen voorgelezen. Exacte cijfers zijn niet gemakkelijk te geven. Maar de indruk bestaat dat in de meeste kleuterklassen (groepen 1 en 2) tenminste één keer per dag door de leerkracht voorgelezen wordt (Veen, Overmars & De Glopper, 1995; p. 181). Waarschijnlijk neemt de frequentie van voorlezen af naarmate de leerlingen ouder worden. Hierbij speelt een rol dat kinderen in de middenbouw (groepen 3, 4 en 5) steeds beter zelf gaan lezen. In de bovenbouw is de leesvaardigheid van de mees-

te leerlingen voldoende hoog om zelfstandig een breed scala van teksten te kunnen lezen. Toch wordt ook in de groepen 7 en 8 nog veelvuldig aan kinderen voorgelezen. In een steekproef van 28 leerkrachten (deels groep 7, deels groep 8) uit Amsterdam en omgeving bleek dat het merendeel van hen meermalen per week naar een voorleesboek greep (Blok, 1992).

In dit artikel beperk ik me tot het voorlezen – in een geïnstitutionaliseerde context – aan jonge kinderen, geoperationaliseerd als kinderen tot ongeveer 8 jaar. Een strikte ondergrens is er niet, omdat deze afhangt van de leeftijd waarop kinderen naar een kindercentrum of naar school gaan. In ons land is dat als regel 4 jaar, maar soms ook eerder. In de Verenigde Staten, waar het merendeel van het wetenschappelijke onderzoek verricht wordt, gaan kinderen als regel vanaf hun vijfde naar school, maar soms eerder.

Werkers op kindercentra en leerkrachten – gemakshalve vatten we vanaf hier beide groepen samen met de term leerkrachten – hebben met het voorlezen diverse doelen voor ogen. De belangrijkste doelen zijn onder drie noemers samen te vatten: bevordering van de taal- en leesontwikkeling, cultuuroverdracht en persoonlijkheidsgroei (Baudoin et al., 1990; Nijmeegse Werkgroep Taaldidactiek, 1992). Voorbeelden van doelen op het terrein van de taal- en leesontwikkeling zijn: vergroting van de taalbeheersing, kennismaking met de functie van geschreven taal en met de opbouw van boeken en verhalen. Voorbeelden van doelen op het terrein van de cultuuroverdracht zijn: wereldoriëntatie, overdracht van waarden en normen (incl. omgangsvormen) en inwijding in de wereld van de literatuur. Een voorbeeld op het terrein van persoonlijkheidsgroei: kinderen inzicht geven in de eigen gevoels- en ervaringswereld.

De vraag kan worden gesteld of voorlezen een geschikt middel is voor het eerste doel:

* Deze literatuurstudie is uitgevoerd met financiële steun van het Instituut voor Onderzoek van Onderwijs (SVO), project 95729.

bevordering van de taal- en leesontwikkeling. Er is uitgebreid empirisch onderzoek gedaan naar deze vraag. Het grootste deel van de studies heeft betrekking op het voorlezen in gezinsverband. Recent is een tweetal meta-analyses verschenen, namelijk Bus, IJzendoorn en Pellegrini (1995) en Scarborough en Dobrich (1994). Deze overzichtstudies komen in belangrijke opzichten overeen. De bestudeerde onafhankelijke variabele is het voorlezen door ouders aan hun kinderen in de voorschoolse leeftijd. De gemeenschappelijke vraag is welke effecten deze variabele heeft op de taal- en leesontwikkeling van kinderen. Gemeenschappelijk is ook het grootste deel van het bronnenmateriaal: ruim 30 studies die in de afgelopen dertig jaar verschenen zijn.

Een verschil is er in de manier waarop de resultaten van de studies gecombineerd zijn. Bus et al., (o.c.) voerden een voornamelijk kwantitatieve meta-analyse uit waarbij studies met een verschillende opzet (correlatieel, longitudinaal, experimenteel) bij elkaar gevoegd zijn. Scarborough en Dobrich (o.c.) hanteerden naast kwantitatieve maten ook kwalitatieve overwegingen en analyseerden de studies per afzonderlijk onderzoekstype. Ondanks het methodische verschil stemmen de conclusies sterk met elkaar overeen. Bus et al. rapporteren een algemene effectgrootte (Cohens d) van .59. Scarborough en Dobrich rapporteren een mediane effectgrootte (pmc) van .26. Beide conclusies samengenomen duiden erop dat 7 tot 10 procent van de variantie in taal- en leesvaardigheden samenhangt met het voorlezen door ouders in de voorschoolse periode.

In beide meta-analyses wordt tevens de conclusie getrokken dat de onafhankelijke variabele, het voorlezen door de ouders, in de meeste studies onvoldoende precies beschreven is. Er bestaat daardoor onzekerheid over de modale 'treatment' die nodig is om, bijvoorbeeld, een effectgrootte van .6 te realiseren. Die onzekerheid bestaat er in drie belangrijke opzichten: de frequentie waarmee voorgelezen wordt, de leeftijd waarop met voorgelezen wordt aangevangen (en bijgevolg de duur van de 'treatment') en de stijl van voorlezen. Wat de voorleesstijl betreft hebben alleen Scarborough en Dobrich een vijftal relevante studies apart geanalyseerd. Zij vonden hierin geen aanwijzingen dat de voorleesstijl sterker met de taal-

en leesontwikkeling samenhang dan de frequentie van voorlezen.

Hoewel beide besproken overzichtstudies het inzicht in de mogelijke effecten van voorlezen verdiepen, kan men de besproken conclusies niet zonder meer generaliseren naar het voorlezen op school. Tussen voorlezen op school en voorlezen in gezinsverband bestaan niet onaanzienlijke verschillen. Het meest kenmerkende verschil is dat ouders als regel aan maar één kind, of hoogstens enkele kinderen tegelijk voorlezen, terwijl op school meestal aan een hele klas tegelijk wordt voorgelezen. Dit verschil heeft consequenties voor het aantal interacties tussen voorlezer en luisteraar en de aard van deze interacties. Leerkrachten hebben tijdens het voorlezen veel minder mogelijkheden hun gedrag af te stemmen op de individuele leerling dan ouders dat hebben ten opzichte van hun eigen kind(eren). Op zijn best bevatten de meta-analyses van Bus et al. en Scarborough en Dobrich alleen indirecte evidentie aangaande het vraagstuk van de effectiviteit van voorlezen op school: de 7 tot 10 procent gemeenschappelijke variantie is mogelijksterwijs een bovengrensschatting.

Een meer direct antwoord kan men proberen te krijgen door een analyse van de literatuur over voorlezen op school. Maar hierbij stuit men op twee moeilijkheden. De eerste is dat er nog maar weinig studies gedaan zijn naar de effecten van voorlezen op school. De tweede moeilijkheid is dat over deze studies geen review beschikbaar is. Het onderhavige artikel is bedoeld om in deze leemte te voorzien. Het doet verslag van een review van elf studies die het resultaat zijn van een uitgebreide speurtocht.

2 Probleemstelling

2.1 Kernbegrippen

Idealiter dicteert een theorie de hypothesen die in een review worden geëvalueerd. In het geval er geen sprake is van een coherente theorie, moet men volstaan met een minder samenhangend stelsel van uitspraken over de factoren die een beschrijving of verklaring geven van de bestudeerde verschijnselen. Hier wordt een poging gedaan zo'n theorie, in een noodzakelijk nog rudimentaire vorm, te specificeren.

Taalontwikkeling is het object van de te ontwikkelen theorie. De snelheid waarmee taalontwikkeling plaatsvindt, en met name de individuele variatie daarin, is de te verklaren grootheid. Taalontwikkeling wordt hier in een ruime betekenis gehanteerd, betrekking hebbend op zowel de gesproken als geschreven taal. Onderscheiden componenten zijn: fonologische vaardigheden, woordkennis, kennis van woord- en zinsbouw, lees- en schrijfvaardigheid. De volgorde van deze opsomming representeert grofweg een ontwikkelingsvolgorde, althans wat betreft het beginpunt van elk der componenten. De ontwikkeling van vaardigheden in het herkennen en produceren van spraakklanken kent bijvoorbeeld een vroegere start dan de ontwikkeling van een woordenschat. Evenzo kent de ontwikkeling van een woordenschat een eerdere start dan de ontwikkeling van syntactische en morfologische vaardigheden. Maar wat betreft de periode waarin de ontwikkelingen plaatsvinden, is er een ruime overlap. De ontwikkeling van een woordenschat verloopt simultaan met ontwikkelingen van vaardigheden waarvan het startpunt later ligt.

Waarvan is de taalontwikkeling, preciezer gezegd de snelheid waarmee het kind zijn taalbeheersing ontwikkelt, afhankelijk? De hoofdstelling is dat de taalontwikkeling van kinderen een functie is van twee grootheden: het mentale vermogen om taal te leren en de ervaringen die kinderen tijdens hun leven met taal opdoen.

Op het *taalleervermogen* wordt hier slechts kort ingegaan. Het is in dit verband van minder belang, omdat voorlezen een exponent is van de andere grootheid, de ervaringen met taal. Men veronderstelt dat kinderen door hun aanleg over een aangeboren vermogen beschikken om taal te leren: kinderen zijn daartoe als het ware voorgeprogrammeerd. Biologische groei en ervaringen met taal (het taalaanbod) zorgen ervoor dat het taalleervermogen tot ontwikkeling komt. Verschillen tussen kinderen wat betreft hun taalleervermogen zijn derhalve op drie bronnen terug te voeren: ze zijn deels door aanleg en biologische groei bepaald, maar deels ook het resultaat van de mate waarin het taalaanbod het taalleervermogen stimuleert. Deze zienswijze impliceert dat de beide grootheden, het taalleervermogen en de ervaringen met taal, niet onafhankelijk van elkaar opere-

ren maar in een interactie de taalontwikkeling bepalen.

De *ervaringen die kinderen met taal opdoen* is de tweede belangrijke grootheid. De aard en de hoeveelheid van deze ervaringen bepalen mede hoe snel taalontwikkeling plaatsvindt. Wat betreft de aard van de ervaring is het nuttig enkele onderscheidingen aan te brengen, omdat deze laten zien dat de ervaringen met taal van grote diversiteit zijn.

Een eerste onderscheiding is die tussen taal die mondeling en taal die schriftelijk aangeboden wordt. Vanzelfsprekend is in de aanvangsfase alleen het mondelinge taalaanbod van belang. Maar als kinderen eenmaal zelf kunnen lezen, neemt hun vaardigheid toe om via schriftelijk taalaanbod hun taalontwikkeling te versnellen. In feite is er bij voorlezen nog een derde vorm van taalaanbod in het geding, geschreven taal die mondeling wordt aangeboden. Door voorlezen ontdekken kinderen allerlei eigenschappen aangaande de vorm en functie van geschreven taal en komen ze in aanraking met het typische schrijftaalregister. Onderzoek naar ontluikende geletterdheid heeft duidelijk gemaakt dat veel kinderen al vóórdat ze als vierjarige de basisschool betreden, een zeker inzicht hebben in vorm en gebruik van geschreven taal (Sulzby & Teale, 1991; Verhoeven, 1994).

Een tweede onderscheiding is die tussen gezin en school. Traditioneel is wat betreft de eerste taalontwikkeling – zeg tot ongeveer vier jaar – de aandacht vooral naar het gezin uitgegaan. Gezien het feit dat jonge kinderen wat betreft hun taalontwikkeling onder sterke invloed van hun ouders staan, valt dat te begrijpen. Bij de bestudering van de latere taalontwikkeling is er vooral aandacht voor de beïnvloeding vanuit de school. Men vergeet daarbij wel eens dat kinderen ook in de schoolleeftijd nog steeds een invloed ondervinden van de talige omgeving thuis. Het lijkt waarschijnlijk dat het gezin, óók voor kinderen in de schoolleeftijd, een belangrijke determinant van de taalontwikkeling is (Wells, 1986).

In het verleden werd wel een derde onderscheid gemaakt, namelijk tussen ongestuurde taalverwerving, die zich buiten de schoolmuren zou afspelen en gestuurde taalverwerving, die uitsluitend binnen school zou plaatsvinden. Maar deze beide termen – gestuurd en onge-

stuurd – suggereren een tegenstelling die er in werkelijkheid niet is. Immers, ook ouders ondersteunen bewust en actief de taalverwerking van hun kinderen, zo goed als kinderen op school ook via ongestuurde situaties hun taalbeheersing vergroten.

2.2 Verbale interactie

Een kind ontwikkelt zijn taalvaardigheid door middel van de taal die hem wordt aangeboden. In dat taalaanbod vallen twee uitersten te onderscheiden: taal die in een directe interactiesituatie rechtstreeks tot het kind wordt gericht en taal die het kind uit de omgeving oppikt, maar die niet rechtstreeks tot het kind is gericht. Algemeen heerst het inzicht dat het eerste type – de verbale interactie – voor de taalontwikkeling van veel groter belang is dan het tweede type, de omgevingstaal. Anders gezegd, verbale interacties vormen de sleutel tot de taalontwikkeling. Er zijn vier redenen aan te geven waarom de verbale interactie een zoveel krachtiger ‘motor’ is dan de omgevings-taal:

- a. door de interactie wordt de aandacht van het kind gericht;
- b. de noodzaak de interactie op gang te houden, dwingt volwassenen tot begrijpelijke taal, aangepast aan het taalniveau van het kind;
- c. de interactie fungeert als een sterke uitnodiging aan het kind om zelf taal te gebruiken en met taalvormen te experimenteren;
- d. de interactie voorziet het kind van uiteenlopende vormen van feedback.

In de literatuur worden uiteenlopende voorbeelden gegeven van effectieve interactievormen of -strategieën. Zo maken ouders in de allereerste fase van de taalontwikkeling gebruik van een specifieke taalvariant, het ‘motherese’, waarvoor in ons land als term het ‘Moeders’ of ‘Mammiaans’ is voorgesteld. Deze variant kenmerkt zich onder meer door een grote expressiviteit, een nadrukkelijke articulatie, een laag spreektempo, een sterk vereenvoudigde syntaxis en een specifiek idioom (Gleitman, Newport & Gleitman, 1984). Bruner (1983) wijst op het grote belang van routinematige procedures (‘routines’), die als het ware uitnodigen tot herhaling van interacties: spelletjes, vaste gebeurtenissen zoals aankleden, wassen, eten en dergelijke. Die

routinematige procedures lijken ook een essentieel onderdeel te vormen bij het gebruik van aanwijsprentenboeken, zoals beschreven door Ninio en Bruner (1978). Schaerlaekens en Gillis (1987) geven een ander belangrijk aspect aan: semantische contingentie. Ouders proberen in de interactie als regel de inhoudelijke relatie te bewaken; zij borduren – inhoudelijk gezien – voort op de bijdragen van het kind. Strategieën die zij hierbij gebruiken zijn onder andere: expansie (herhaling van de uiting van het kind, maar in een meer volledige vorm), bevestigen van beweringen van het kind, vragen om verduidelijkingen, aanbrenge van correcties, enzovoort. Dergelijke strategieën vormen als het ware de tijdelijke steigers die kinderen in staat stellen aan hun taal te bouwen. Bruner (1983) spreekt in dit verband van ‘scaffolding’, een techniek waarmee ouders uitbreiding geven aan wat Vygotskij de ‘zone van de naaste ontwikkeling’ heeft genoemd. Zodra het kind laat blijken de aangeboden steun niet meer nodig te hebben, wordt deze steun ook niet meer aangeboden.

2.3 Taalontwikkeling door voorlezen

Om te begrijpen waarom voorlezen de taalontwikkeling van jonge kinderen kan versnellen, is het nodig aandacht te schenken aan de verschillende manieren van voorlezen. Hierin zijn twee uitersten te onderscheiden: simpel gezegd, ‘praten met het kind’ en ‘praten tegen het kind’. Bij de eerste manier van voorlezen is er zeer veel interactie tussen voorlezer en kind over de tekst, bij de tweede manier is die interactie nagenoeg afwezig; de voorlezer draagt de tekst voor en de rol van het kind is teruggebracht tot luisteren. De eerste manier duiden we aan als een interactieve voorleesstijl, de tweede als een passieve voorleesstijl.

Er zijn redenen om van een interactieve voorleesstijl een veel sterker effect op de taalontwikkeling te verwachten dan van een passieve voorleesstijl. In feite zijn dit ongeveer dezelfde redenen als waarom het aangaan van verbale interacties voor de taalontwikkeling veel gunstiger is dan het aanbieden van omgevingstaal. Voorlezen vindt veelal plaats in een specifieke ‘setting’ waarin de handelingen met een grote mate van routine vervuld worden. Die routinematigheid zorgt voor een bepaalde mate van voorspelbaarheid. Door die voorspel-

baarheid worden kinderen in staat gesteld langzamerhand een steeds groter aantal handelingen, eerder nog verricht door de voorlezer, zelfstandig te verrichten. Men kan denken aan het aanreiken van het boek, het omslaan van de bladzijden, het aanvullen van het laatste woord, of, in een later stadium, het parafraseren van delen van het verhaal. Ondanks het routinematige karakter is voorlezen geen rigide activiteit. De 'setting' biedt voldoende mogelijkheden voor een zekere variatie in de herhaling.

Een ander belangrijk kenmerk is dat de interactieve voorleesstijl volop gelegenheid biedt tot de al genoemde 'scaffolding'. De voorlezer biedt het kind ondersteuning bij de betekenisreconstructie, waartoe het kind zelfstandig nog niet in staat is. Die ondersteuning biedt het kind als het ware de mogelijkheid de eigen beperkingen te overstijgen. Van belang is dat de voorlezer in opeenvolgende situaties de mate van ondersteuning vermindert, waardoor het kind de mogelijkheid krijgt steeds grotere delen van de betekenisreconstructie zelfstandig te verrichten.

Rest hier de bespreking van de vraag welke taalontwikkelingscomponenten bij voorlezen aan bod kunnen komen. De verwachting is dat de effecten zich op een breed terrein kunnen manifesteren: inzicht in verhaalstructuren, woordenschat, vaardigheden op het gebied van woord- en zinsbouw, meta-linguïstisch bewustzijn en vaardigheden op het gebied van de vroege geletterdheid. Op voorhand is het moeilijk om aan te geven op welk gebied de grootste effecten te verwachten zijn. De effectgrootte zou kunnen samenhangen met de leeftijd van de kinderen en mogelijk ook met de taalcomponent die in de interactie over de tekst centraal gesteld wordt. Een betrekkelijk algemene verwachting is dat de effecten op mondelinge taalvaardigheidsaspecten groter zullen zijn dan op de schriftelijke. Deze verwachting is ingegeven door het feit dat voorlezen – althans voor de kinderen – een voornamelijk mondelinge aanleggenheid is. Ze oefenen veel meer hun luister- en spreekvaardigheid (inclusief woordenschat en grammaticale vaardigheden) dan hun geletterdheid. Enige steun voor deze algemene verwachting werd gevonden door Bus et al. (1995), die voor mondelinge taalvaardigheid een iets sterker effect rapporteren ($d = .67$) dan voor ontlukende geletterdheid ($d = .58$) of

voor lezen ($d = .55$).

2.4 Hypothesen

Het voorgaande leidt tot de algemene verwachting dat het effect van voorlezen op school op de taalontwikkeling van jonge kinderen aan diverse invloeden bloot staat. Deze algemene verwachting is in vijf hypothesen gespecificeerd.

Hypothese 1: voorlezen is gunstiger voor de taalontwikkeling naarmate er tussen voorlezer en kind meer interactie is over de voorgelezen tekst.

Hypothese 2: in vergelijking met het voorlezen door ouders thuis valt te verwachten dat het effect van voorlezen door leerkrachten op school geringer zal zijn, geringer derhalve dan .6 (Cohens d), het resultaat uit de meta-analyses van Bus et al. (1995) en Scarborough en Dobrich (1994).

Hypothese 3: de grootte van de groep waaraan wordt voorgelezen hangt negatief samen met de grootte van het effect.

Hypothese 4: het effect van voorlezen zal groter zijn op de mondelinge taalontwikkeling dan op de schriftelijke taalontwikkeling.

Hypothese 5: naarmate het voorlezen op school vaker gebeurt, zal de effectgrootte toenemen.

3 Opzet

3.1 Verzameling en selectie van de studies

Searches in elektronische literatuurbestanden hebben een belangrijke basis gelegd voor de verzamelde literatuurverwijzingen. Deze searches zijn uitgevoerd aan de hand van zoekprofielen op basis van drie groepen descriptorren: een groep ter aanduiding van de onafhankelijke variabele ('storybook reading, story reading, reading aloud to others'), een groep ter aanduiding van de afhankelijke variabelen ('reading comprehension, listening comprehension, vocabulary development') en een groep ter aanduiding van de leeftijd van de kinderen ('early childhood education, preschool education, primary education'). Deze keuze van descriptorren is mede afgestemd op de descriptorren die aan reeds beschikbare kernartikelen bleken te zijn toegekend. De searches zijn uitgevoerd in juni 1995 op vier bestanden:

ERIC (vanaf 1965 tot en met maart 1995), Dissertation Abstracts International (vanaf 1980, het jaar waarin men begonnen is met het opne- men van abstracts, tot en met mei 1995), Language Learning en Behavior Abstracts (1973 tot en met december 1994) en PsycLit (vanaf 1974 tot en met maart 1995). Het leeu- wendeel van de relevante referenties werd in ERIC gevonden. De andere drie bestanden leverden weinig bruikbare verwijzingen op.

Naast de searches is ook handmatig gezocht in literatuurlijsten van reeds beschikbare publi- caties, waaronder de eerdere bibliografieën van McCormick (1981) en Radecki (1987). Naar- mate de aan opsporing bestede tijd toenam, nam het aantal gevonden referenties af. Omdat is doorgezocht totdat geen nieuwe referenties meer gevonden werden, bestaat de indruk dat de populatie van relevante referenties bij bena- dering volledig verzameld is. In totaal handel-

Tabel 1
Overzicht van gecodeerde studiegenmerken

Studiegenmerk	Schaal
1. Jaar van publiceren	jaartal
2. Publicatietype	1. artikel 2. ongepubliceerd (these, rapport, congresbijdrage, etc.)
3. Omvang van de experimentele groep	aantal leerlingen
4. Omvang van de controlegroep	aantal leerlingen
5. Sociaal-economische status van de leerlingen	1. achterstandssituatie 2. gemengd 3. voornamelijk niet uit achterstandssituatie
6. Natie waar het onderzoek verricht is	1. Verenigde Staten 2. elders
7. Leeftijd leerlingen, gemiddeld bij de start van het onderzoek	aantal maanden
8. Onderzoeksopzet	1. experimenteel 2. correlatieel
9. Controle op samenstelling experi- mentele en controle groep	1. matching 2. randomisatie
10. Duur van het experimentele pro- gramma	aantal weken
11. Aantal contacturen in het experi- mentele programma	aantal uren
12. Gemiddelde groeps grootte tijdens voorlezen	aantal leerlingen
13. Scholing van de voorlezers	1. geen speciale scholing 2. alleen schriftelijke scholing 3. schrifte- lijke en mondelinge vormen van scholing
14. Aard van voorlezen	1. alleen voorlezen 2. 'aangekleed' met activiteiten direct voor of na het voorlezen 3. ingebed in verdergaande activiteiten
15. Herhaalde aanbieding van hetzelf- de boek	0. nee 1. ja
16. Bijzondere aandacht voor de selec- tie van aan te bieden boeken	0. nee 1. ja
17. Aard van de voorlezers	1. leerkrachten aan eigen leerlingen 2. 'vreemden'
18. Bezigheden controlegroep	1. standaardprogramma 2. leesrijheidscurriculum
19. 'Bias' in effectgrootte	0. geen 'bias' of 'bias' onbekend 1. positieve 'bias', bijvoorbeeld als gevolg van matching of covariantie-analyse
20. Domein	1. mondelinge taal 2. lezen
21. Geschatte effectgrootte	uitgedrukt als het gestandaardiseerde verschil in gemiddelde tus- sen experimentele en controle groep
22. Standaardfout van de geschatte effectgrootte	uitgedrukt op dezelfde schaal als de effectgrootte

de het om 63 literatuurverwijzingen.

Niet alle betrokken literatuur is daadwerkelijk verzameld, bestudeerd of in de review opgenomen. Soms bleek de taal een beletsel (drie publicaties waren in het Hebreeuws), zes publicaties (met name oudere dissertaties en congresbijdragen) konden niet geleverd worden, drie onderzoeksrapporten beschreven onderzoek dat later ook als tijdschriftartikel verscheen, in zeven gevallen bleek voorlezen op school geen onderdeel van de 'treatment' te zijn, twee publicaties betroffen onderzoek onder uitsluitend migrantenkinderen. Bij nadere bestudering vielen nog 22 publicaties af, ofwel omdat er van empirisch onderzoek geen sprake was (in zeven gevallen), ofwel omdat er geen controlegroep in de onderzoeksopzet aanwezig was (in twaalf gevallen), ofwel omdat de rapportage zo onvolledig was dat er geen effectgroottes geschat konden worden (in drie gevallen). Acht publicaties zijn buiten beschouwing gebleven, omdat als 'treatment' slechts één prentenboek was voorgelezen. Ten slotte, Morrow (1988) is verwijderd vanwege het afwijkende karakter van haar afhankelijke variabelen. Zij meet het effect van voorlezen op de taalontwikkeling uitsluitend af aan de 'responsiveness', geoperationaliseerd als het aantal opmerkingen van kinderen tijdens het voorlezen.

In totaal zijn derhalve 52 van de 63 publicaties terzijde geschoven en bleven er elf bruikbare studies over. Deze zijn in de literatuurlijst met een asterisk aangegeven.

3.2 Selectie en codering van variabelen

Er zijn 22 variabelen geselecteerd voor codering (zie Tabel 1). De variabelen kunnen als volgt worden gerubriceerd: publicatiestatus (variabelen 1 en 2), beschrijving steekproef (variabelen 3 tot en met 7), onderzoekstype (variabelen 8 en 9), typering van voorlezen (variabelen 10 tot en met 17), typering activiteiten in controlegroep (variabele 18).

Variabele 19 ('bias' in effectgrootte) is toegevoegd, omdat in een aantal onderzoeken gebruik is gemaakt van technieken om de variantie in de afhankelijke variabele te reduceren. Het gaat met name om onderzoeken waarbij leerlingen gematcht zijn of waarin covariantieanalyse is toegepast. Zulke technieken zijn weliswaar verdedigbaar vanuit de wens de

gevoeligheid van de onderzoeksopzet te vergroten. Maar het effect is een positieve 'bias' van de effectgrootte, een 'bias' die groter is naarmate de matchingsvariabelen of covariaten sterker samenhangen met de afhankelijke variabelen. Opname van variabele 19 maakt het mogelijk om achteraf een schatting te doen van de grootte van de 'bias'.

Variabele 20 (domein) is het resultaat van een post-hoc classificatie. De afhankelijke variabelen zijn onderverdeeld naar het domein waarop ze betrekking hebben: mondelinge taal (inclusief woordenschat) en lezen. In beginsel verdient een fijnere indeling de voorkeur, maar het geringe aantal studies maakte een fijnere indeling onpraktisch.

Variabelen 21 en 22 betreffen respectievelijk de effectgrootte en de bijbehorende standaardfout. Gebruik is gemaakt van Cohens d , waarbij de standaardisatie is uitgevoerd via de 'within group' standaardafwijking. De standaardfout is geschat volgens Hedges en Olkin (1985, p. 86, Eq. 15). De standaardfout is een functie van de steekproefgrootte en de effectgrootte. Naarmate de steekproef kleiner is of de effectgrootte groter, is de standaardfout van de effectgrootte kleiner.

De codering van de studies volgens de in Tabel 1 aangegeven variabelen bleek in het algemeen eenvoudig. Eén kwestie vroeg evenwel extra aandacht. Aangezien de meeste publicaties geen schatting geven van de effectgrootte, moest deze dikwijls worden afgeleid uit de wèl vermelde gegevens. Hierbij is zoveel mogelijk gebruik gemaakt van omzettingstabellen in leerboeken voor statistiek. In de enkele gevallen dat zulks onmogelijk bleek, is gebruik gemaakt van benaderingsformules zoals gegeven in Mullen (1989).

3.3 Integratie van effecten

Het resultaat van de coderingsfase was een bestand met resultaten uit 12 studies¹. In dit bestand waren 59 geschatte effectgroottes aanwezig. Om te vermijden dat studies met meer uitkomsten een zwaarder stempel zouden drukken, zijn de effectgroottes in twee stappen naar een hoger aggregatieniveau gebracht. In de eerste stap is de 'studie x domein' combinatie als eenheid gekozen. Anders gezegd, de uitkomsten zijn binnen elke studie per domein vastgesteld. Dit resulteerde in een bestand met 12 stu-

dies en 20 uitkomsten (elf voor het domein 'mondelinge taal', negen voor het domein 'lezen'. In de tweede stap zijn de 20 uitkomsten geaggregeerd door per studie de uitkomsten samen te nemen. Dit resulteerde in een bestand met twaalf uitkomsten, voor elke studie één gecombineerde uitkomst.

Beide aggregaties zijn verricht volgens het model van de 'weighted integration method', waarbij de uitkomsten worden gewogen naar hun standaardfout: hoe groter de standaardfout, des te geringer het gewicht. De geaggregeerde effectgroottes en de bijbehorende standaardfouten zijn geschat volgens Hedges en Olkin (1985, p. 112, Eqs. 8 en 9). Aanname bij deze manier van aggregeren is dat de uitkomsten binnen één studie homogeen zijn en alleen van elkaar verschillen op basis van toevallige verschillen tussen de afhankelijke variabelen.

Voor het integreren van de uitkomsten van verschillende studies lijkt deze aanname te beperkend. Verschillen in uitkomsten kunnen dan immers ook het gevolg zijn van systematische verschillen in de experimentele opzet of verschillen tussen de populaties waaruit steekproeven getrokken zijn. Daarom is voor het integreren van de uitkomsten van verschillende studies het 'random effects' model toegepast (Hedges & Olkin, 1985, p. 189 e.v.). In dit model wordt de effectgrootte voor studie j , d_j , gesplitst in twee componenten: de ware effectgrootte δ_j en de foutencomponent ϵ_j :

$$d_j = \delta_j + \epsilon_j \quad (1)$$

Dankzij het feit dat de varianties van d_j bekend zijn, biedt het model de mogelijkheid de variantie van δ_j , de variantie van de ware effectgrootte te onderscheiden van de variantie van ϵ_j , de foutenvariantie. Het is vervolgens mogelijk te toetsen of de ware parametervariantie statistisch significant van nul verschilt. Indien dat inderdaad het geval blijkt te zijn, wordt de verzameling effectgroottes als heterogeen beschouwd. Het loont dan de moeite te trachten de heterogeniteit in verband te brengen met de diverse studiekekenmerken.

De beide aggregaties op basis van de 'weighted integration method' zijn verricht met het programma META (Schwarzer, 1989). De analyses volgens het 'random effects' model zijn uitgevoerd met behulp van een spe-

ciale versie van HLM (namelijk VKHLM, zie Bryk, Raudenbusch & Congdon, 1994). In tegenstelling tot META biedt VKHLM de mogelijkheid studiekekenmerken als predictoren in de analyse op te nemen.

De resultaten van de beide aggregaties zijn in Tabel 2 opgenomen. De gerapporteerde effectgroottes zijn erg heterogeen: ze vertonen een groot bereik (van -1.24 tot en met 1.59).

4 Resultaten

Allereerst wordt, in paragraaf 4.1, een kwalitatieve beschrijving en evaluatie gegeven van de geselecteerde studies. In paragraaf 4.2 worden de studieuitkomsten kwantitatief samengevat en in verband gebracht met de studiekekenmerken.

4.1 De studies kwalitatief beschouwd

De geselecteerde studies zijn over het algemeen van recente datum. Ze zijn gepubliceerd tussen 1968 en 1994 (mediaan: 1990). Zeven studies zijn als artikel gepubliceerd, vier behoren er tot wat wel als grijze literatuur wordt aangeduid.

Qua onderzoekstype zijn de studies tamelijk homogeen: in elf van de twaalf gevallen is er sprake van een experimentele opzet met een controlegroep. Meyer et al. (1994) zijn de enigen die over een survey-opzet rapporteren. De studies zijn over het algemeen tamelijk kleinschalig. De gemiddelde omvang van de experimentele groep bedraagt 50 kinderen (minimum: 10, maximum: 155). In de meeste studies wordt extra zorg besteed aan de vergelijkbaarheid van de experimentele en controlegroep, soms via 'matching', maar vaker post-hoc via covariantieanalyses.

Desalniettemin vertonen diverse studies op het vlak van de onderzoeksoptzet duidelijke tekortkomingen. De studie van Cohen (1968) schiet tekort doordat er geen controle was op de uitvoering van de experimentele conditie. Maar er zijn meer studies die weinig inzicht geven in de implementatie van het experimentele voorleesprogramma. In twee studies – Feitelson, Kita & Goldstein, 1986; Feitelson et al., 1993 – wordt een sterke oppositie van de leerkrachten tegen de voorleesconditie gemeld. In de eerste van de genoemde studies zijn om deze reden

Tabel 2

De geselecteerde studies: aantal proefpersonen in experimentele en controlegroep (N_{exc}), domein (1 = mondelinge taal, 2 = lezen), effectgrootte (d) en standaardfout van de effectgrootte (se_d)

Studie	N_{exc}	Domein	d	se_d
Cohen, 1968	155+130	1	.31	.07
		2	.15	.08
		gecombineerd	.23	.08
Feitelson et al., 1991	19+19	1	.91	.15
Feitelson et al., 1993	86+49	1	.53	.07
Feitelson et al., 1986	31+57	1	.19	.10
		2	.67	.16
		gecombineerd	.42	.16
Karweit, 1989; pre-K	43+43	1	.52	.16
		2	.52	.22
		gecombineerd	.52	.13
idem; Kindergarten	60+60	1	.43	.13
		2	1.33	.12
		gecombineerd	.70	.11
Khare, 1992	38+38	2	.30	.23
Meyer et al., 1994	325	1	.45	.11
		2	-1.24	.05
		gecombineerd	-.33	.08
Morrow et al., 1990	32+30	1	1.59	.21
		2	.08	.10
		gecombineerd	.72	.19
Peterman, 1988	60+60	1	.61	.13
		2	.16	.13
		gecombineerd	.38	.13
Sullivan, 1988	14+14	1	.57	.28
		2	.00	.38
		gecombineerd	.27	.27
Valdez-Menchaca & Whitehurst, 1992	10+10	1	1.49	.36

zelfs twee van de drie experimentele klassen voortijdig uit het onderzoek gestapt. De rapportage heeft alleen op de overblijvende klas betrekking. Zoiets kan de uitkomsten sterk vertekenen, indien bijvoorbeeld de overblijvende leerkracht over bijzondere capaciteiten zou beschikken. De correlatieve studie van Meyer et al. (1994) schiet ernstig tekort op het vlak van de data-analyse. In de eerste plaats wordt voorbij gegaan aan de multiniveau structuur van de data. Ernstiger is dat er alleen 'zero-order' correlaties worden gerapporteerd. Veel krachtiger multivariate analysetechnieken – waaronder multipale regressie en lineair-structurele modellen – worden niet toegepast.

Over de proefpersonen valt het volgende te melden. In acht van de twaalf studies kwamen de kinderen uit een achterstandssituatie. Het accent ligt op kinderen van vijf jaar oud. De gemiddelde leeftijd bedraagt 66 maanden (minimum: 31 maanden, maximum: 90 maanden). Acht van de twaalf studies zijn in de Verenigde Staten verricht. De steekproef van Val-

dez-Menchaca en Whitehurst (1992) steekt af bij die in andere onderzoeken. Het betreft tweejarige Mexicaanse kinderen uit achterstandssituaties die een kinderdagopvang bezochten.

De gemiddelde duur van het experimentele programma bedroeg 24 weken (minimum: 3 weken, maximum: 40 weken). Uitgedrukt in contacturen bedraagt de gemiddelde duur 42 uur (minimum: 5 uur, maximum 140 uur). De gemiddelde grootte van de groep waaraan werd voorgelezen bedroeg 19 kinderen. Slechts in drie studies was de gemiddelde groepsgrootte minder dan tien kinderen. In zes studies werd alleen maar aan de kinderen voorgelezen, zonder 'follow-up' activiteiten. In acht studies ontvingen de voorlezers geen aparte scholing. Meestal waren het de eigen leerkrachten die aan de kinderen voorlezen. Over de bezigheden van de controlegroep bevatten tien studies geen of bijna geen nadere aanduidingen.

Wat betreft de afhankelijke variabelen is een algemene tekortkoming in de rapportages dat over de betrouwbaarheid van de afhankelijk-

ke variabelen nagenoeg geen gegevens vermeld worden. Een tweede probleem is de aard of de betekenis van de variabelen. De hier gehanteerde indeling in twee domeinen (mondelinge taal, lezen) verbergt in feite een enorme variëteit aan afhankelijke variabelen. De indruk bestaat dat met name sommige instrumenten voor de mondelinge taalbeheersing in standaardisatie tekort schieten. Morrow, O'Connor en Smith (1990) en Peterman (1988) maken gebruik van 'free recall' en 'probed recall' taken, waarbij kinderen een voorgelezen boek zo goed mogelijk moeten proberen na te vertellen. Behalve een gebrek aan standaardisatie bestaat tegen deze instrumenten nog een tweede bezwaar. Ze leunen sterk op wat de kinderen in de experimentele conditie is aangeleerd: luisteren naar en praten over voorgelezen boeken. De standaardisatie van de maten voor het domein 'lezen' is van een hoger niveau. Maar hier doet zich het verschijnsel voor dat de maten erg sterk uiteenlopen: van toetsen voor ontluikende geletterdheid (o.a. 'Sulzby's scale for reenactments' en de toets 'Concepts about print') tot en met testen voor technisch en begrijpend lezen. Die diversiteit zou gemakkelijk kunnen verbergen dat voorlezen voor sommige aspecten van geletterdheid veel gunstiger is dan voor andere aspecten.

Als laatste punt van deze paragraaf wordt aandacht besteed aan afwijkende effectgroottes, waaronder hier de effectgroottes begrepen worden die groter dan 1.00 (absoluut) zijn. Het zijn er vier, twee voor het domein mondelinge taal en twee voor het domein lezen. De studie van Morrow, O'Connor en Smith (1990) resulteerde in een effectgrootte van 1.59 voor het domein mondelinge taal. In het voorgaande is al naar voren gebracht dat de afhankelijke maten in deze studie dicht tegen de experimentele conditie aanliggen. De effectgrootte voor het domein lezen in dezelfde studie, geoperationaliseerd via maten die veel verder afliggen van de experimentele conditie, bedroeg slechts .08. Het sterke effect (domein mondelinge taal; effectgrootte 1.49) in de studie van Valdez-Menchaca en Whitehurst (1992) is gerealiseerd in een experimentele groep van slechts tien kinderen aan wie in een één-op-één situatie werd voorgelezen. Het is de enige studie waarin één-op-één voorgelezen is. Opvallend is verder dat de experimentele conditie niet erg

intensief is geweest: 30 sessies van circa tien minuten interactief voorlezen. In feite is in deze studie het laagste aantal voorleesuren gerealiseerd.

De studie van Karweit (1989) in de kindergarten-steekproef resulteerde in een effectgrootte van 1.33 voor lezen. Het effect is sterk beïnvloed door een effectgrootte van 3.75 op de 'Woodcock word attack test', waarbij de auteur aangeeft dat de controlegroep in de door de test gemeten vaardigheid – in tegenstelling tot de experimentele groep – geen enkele oefening heeft gehad. Deze opmerking suggereert dat de verschillen tussen de experimentele en controle conditie veel verder reikten dan alleen het voorlees-programma, een suggestie die wordt versterkt door het feit dat de experimentele scholen proefscholen waren voor het 'Success for All' programma van Slavin. De vierde afwijkende effectgrootte (domein lezen; effectgrootte -1.24) is gebaseerd op de correlatieve studie van Meyer et al. (1994). Het is in feite een gecombineerde effectgrootte, op basis van vier tests voor technisch lezen die alle vier sterk negatief correleerden met de tijd die leerkrachten aan voorlezen besteedden. Diezelfde tests correleerden sterk positief met de tijd die leerkrachten aan specifieke oefeningen voor aanvankelijk lezen besteedden.

4.2 De gerapporteerde effecten

Voor het domein 'mondelinge taal' blijkt de gecombineerde effectgrootte .64 te bedragen (zie Tabel 3). Gezien de bijbehorende standaardfout verschilt deze waarde statistisch significant van nul. Voor het domein 'lezen' bedraagt de gecombineerde effectgrootte .22, een waarde die statistisch niet significant van nul verschilt. De 'overall' effectgrootte, waarin de resultaten over de twee domeinen samengevoegd zijn, bedraagt .47. Deze waarde verschilt statistisch significant van nul. In alle drie de gevallen wijst de heterogeniteitstoets uit dat de uitkomsten heterogeen zijn. Dit rechtvaardigt het zoeken naar studiekenmerken die met de effectgrootte samenhangen.

Gezien het geringe aantal studies zijn de heterogeniteitsanalyses op een terughoudende manier verricht. Per afzonderlijke analyse is steeds maar één studiekenmerk als predictor ingevoerd. Bestudeerd is of opname van een predictor leidde tot een reductie van de hete-

Tabel 3

Uitkomsten van de analyse volgens het 'Random effects' model (model zonder predictoren), waarbij δ de ware effectgrootte is

Domein (aantal studies)	'Fixed effect'		'Random effect'			
	δ	se_{δ}	$Var(\delta_j)$	df	χ^2	p -waarde
Mondelinge taal (11)	.64	.12	.14	10	80.7	.00
Lezen (9)	.22	.24	.48	8	946.9	.00
Gecombineerd (12)	.47	.11	.13	11	135.0	.00

rogeniteit (i.e. de variantie van δ). De reductie is uitgedrukt als een percentage, waarbij de heterogeniteit van het nulmodel – het model zonder predictoren – op 100 procent gesteld is. Bedraagt de reductie voor een predictor bijvoorbeeld 10 procent, dan betekent dit dat toevoeging van de predictor aan het nulmodel de heterogeniteit van de ware effectgrootte met 10 procent doet afnemen. De analyses zijn verricht voor bijna alle in Tabel 1 opgenomen studiekekenmerken, uitgezonderd kenmerk 4 (de omvang van de controlegroep; deze variabele overlapt sterk met kenmerk 3, de omvang van de experimentele groep) en kenmerk 8 (er was slechts één studie met een correlatieve opzet).

Uit Tabel 4 komt naar voren dat diverse studiekekenmerken sterk samenhangen met de effectgrootte. Maar tevens blijkt dat het beeld voor de drie domeinen (mondelinge taal, lezen, gecombineerd) niet gelijk is. Opvallend voor

het domein 'mondelinge taal' is dat het kenmerk 'groeps-grootte' sterk negatief geassocieerd is met de effectgrootte: de sterkere effecten worden gevonden in de studies waarin aan kleinere groepen kinderen wordt voorgelezen. Dit kenmerk verklaart maar liefst 98 procent van de heterogeniteit. Nog één ander studiekekenmerk verklaart meer dan 25 procent van de variantie: de leeftijd van de kinderen waaraan wordt voorgelezen (grotere effecten bij jongere kinderen).

Voor het domein 'lezen' geldt dat er vier studiekekenmerken zijn die elk op zich omstreeks 25 procent van de heterogeniteit kunnen verklaren. Eén daarvan is de omvang van de experimentele groep (kenmerk 3), die – evenals bij het domein 'mondelinge taal' – negatief geassocieerd is met de effectgrootte. De andere drie zijn de sociaal-economische status van de leerlingen (grotere effecten bij leerlingen met een lagere status), kenmerk 9 (grotere effecten

Tabel 4

Reductie van de heterogeniteit (in procenten) die het gevolg is van de opname van een studiekekenmerk in het nulmodel; tussen haakjes wordt de richting van het effect aangegeven

Studiekekenmerk ^a	Mondelinge taal (n=11)	Lezen (n=9)	Gecomb. (n=12)
1. Jaar van publiceren	0	0	0
2. Publicatietype	0	0	0
3. Omvang exp. groep	20 (-)	28 (-)	85 (-)
5. SES leerlingen	0	24 (-)	41 (-)
6. Natie	0	0	15 (+)
7. Leeftijd	28 (-)	0	12 (-)
9. Matching/randomisatie	0	25 (-)	34 (-)
10. Duur programma	0	0	0
11. Aantal contacturen	0	0	0
12. Groeps-grootte	98 (-)	0	31 (-)
13. Scholing	0	0	0
14. Aard van voorlezen	0	13 (+)	0
15. Boekherhaling	0	9 (+)	0
16. Boekselectie	0	29 (+)	9 (+)
17. Aard voorlezers	20 (+)	0	15 (+)
18. Bezigheden C-groep	11 (+)	0	0
19. 'Bias' in effectgrootte	0	9 (+)	33 (+)

^a Voor een toelichting op de studiekekenmerken raadplege men Tabel 1

in de studies waarin de leerlingen gematcht zijn) en kenmerk 16 (grotere effecten in de studies waarin expliciete aandacht gegeven werd aan de selectie van de voor te lezen boeken).

Bezien we de gecombineerde uitkomsten, dan valt op dat de steekproefomvang een sterk negatieve relatie vertoont met de effectgrootte (sterkere effecten bij onderzoeken met een kleinere experimentele groep). Dit kenmerk verklaart de heterogeniteit voor 85 procent. Nog vier andere kenmerken vertonen op dit niveau van aggregatie een sterke relatie met de effectgrootte: de sociaal-economische status van de leerlingen (grotere effecten bij leerlingen met een lagere status), kenmerk 9 (grotere effecten in de studies waarin de leerlingen gematcht zijn), de groepsgrootte (grotere effecten in studies waarin aan kleinere groepen kinderen wordt voorgelezen) en kenmerk 19 (grotere effecten in de studies waarin de effect-schattingen waarschijnlijk positief onzuiver zijn).

Al met al is het in Tabel 4 gepresenteerde beeld niet erg eenduidig. Het kleine aantal studies, het kleine aantal leerlingen in de meeste studies en de hiervoor besproken tekorten in de kwaliteit van de afzonderlijke studies dragen hieraan waarschijnlijk bij. In feite is er slechts één kenmerk dat consistent negatief samenhangt met de effectgrootte, namelijk de omvang van de experimentele groep: hoe kleiner de groep, des te groter het gerealiseerde effect.

5 Conclusies en discussie

5.1 Conclusies

Tot welke conclusies leidt deze review van onderzoek naar de effecten van voorlezen op school aan jonge kinderen? Allereerst moet worden vastgesteld dat de empirische basis tamelijk smal is. De basis wordt gevormd door twaalf onafhankelijke steekproeven, verzameld in elf studies. Van belang is dat een aantal van deze studies tamelijk kleinschalig is en qua opzet ernstige tekortkomingen vertoont.

De meest algemene, over de twaalf steekproeven en de twee domeinen gegeneraliseerde effectschatting bedraagt .47, met een standaardfout van .11. Uitgesplitst naar domein bedraagt de effectschatting voor mondelinge

taal .64 (standaardfout .12) en voor lezen .22 (standaardfout .24). In termen van Cohen (1988): voorlezen op school heeft een middelmatig effect op de mondelinge taalontwikkeling van kinderen en een zwak effect op de leesontwikkeling van kinderen. Niet uitgesloten kan worden dat het effect op lezen nul bedraagt, aangezien de nul binnen het 95-procents betrouwbaarheidsinterval valt.

De gegeneraliseerde schattingen verbergen een aanzienlijke variatie in onderzoeksuitkomsten. Bij statistische toetsing blijkt er in alle drie de gevallen (mondelinge taal, lezen, gecombineerd) sprake te zijn van heterogeniteit. Het is een ongunstige omstandigheid dat de heterogeniteit slechts voor een beperkt gedeelte samen blijkt te hangen met de in de review betrokken studiekekenmerken. Voor het domein mondelinge taal zijn er twee kenmerken die de effectgrootte modereren: de effecten zijn sterker bij voorlezen aan kleinere groepen en aan jongere kinderen. Voor het domein lezen zijn er vier modererende kenmerken: de effecten zijn sterker in studies met een kleinere steekproef en in studies waarin kinderen gematcht zijn. Bovendien worden sterkere effecten gerapporteerd in studies met kinderen uit achterstandssituaties en in studies waarin de voorgelezen boeken met zorg gekozen zijn.

De conclusies verlenen steun aan twee van de vijf in paragraaf 2.4 geformuleerde hypothesen. Gesteund wordt de hypothese dat er een 'overall' positief effect is, evenwel zwakker dan het effect van voorlezen door ouders. Men vergelijk de hier gerapporteerde effectgrootte van .47 met de door Bus et al. (1995) gerapporteerde effectgrootte van .59 en de door Scarborough en Dobrich (1994) gerapporteerde mediane correlatie van .26. Bevestigd wordt ten tweede ook de hypothese dat het effect van voorlezen op school voor mondelinge taal sterker is dan voor lezen.

Maar er zijn ook twee hypothesen die in de data geen steun vinden. Het verwachte modererende effect van de grootte van de groep waaraan wordt voorgelezen is alleen voor het domein mondelinge taal aangetroffen. Het verwachte modererende effect van de duur en de intensiteit van het voorleesprogramma is zelfs in geen van beide domeinen waargenomen. Met name de laatste bevinding roept serieuze twijfel op. Immers, als voorlezen op school de

taal- en leesontwikkeling van kinderen ondersteunt, dan mag men toch verwachten dat er een relatie bestaat tussen de duur van de 'treatment' en het daarmee gerealiseerde effect.

Tenslotte, er is één hypothese die in de onderhavige review niet getoetst kon worden. Dat is de verwachting dat interactief voorlezen tot een sterker effect zal leiden dan een passieve voorleesstijl. Vrijwel alle rapportages gaven onvoldoende inzicht in de gehanteerde voorleesstijl. Dit punt brengt ons naar het discussiegedeelte van deze paragraaf, waarin drie onderwerpen worden besproken.

5.2 Discussie

Elke review – dus ook deze – wordt serieus bedreigd door het feit dat de gevonden studies geen getrouwe afspiegeling van al het verrichte onderzoek hoeven te vormen. Rosenthal (1979) spreekt in dit verband over het 'file drawer' probleem, het verschijnsel dat studies met een statistisch significant resultaat een grotere kans hebben om in een tijdschrift gepubliceerd te worden dan studies met een nul-uitkomst. Hoe ernstig deze bedreiging voor de onderhavige review is, valt moeilijk te zeggen. Weliswaar is gebleken dat het studiekekenmerk 'publicatietype' geen modererend effect had op de effectgrootte. Maar hierbij moet worden aangetekend dat met een aantal van slechts twaalf steekproeven geen gevoelige analyses mogelijk zijn. Een andere benadering is een schatting te maken van het zogenaamde 'fail-safe number', het aantal studies met nul-uitkomsten dat nodig is om de effectgrootte van .64 (zoals vastgesteld voor het domein mondelinge taal) terug te brengen tot bijvoorbeeld .20, een waarde die correspondeert met nog slechts een gering effect. Volgens een door Orwin (1983) gegeven formule bedraagt dit 'fail-safe' aantal in het onderhavige voorbeeld 24 studies. Anders gezegd, er zouden 24 niet gepubliceerde studies met een nul-resultaat moeten zijn om het gemiddelde effect van .64 terug te brengen tot .20. Hoe exact dit aantal ook is, het is moeilijk te evalueren bij gebrek aan inzicht in het werkelijke aantal ongepubliceerde studies en de uitkomsten daarvan.

Een andere bedreiging voor de onderhavige review betreft de betrouwbaarheid van de afzonderlijke rapportages. In een aantal gevallen is er reden om hier aan te twifelen. De stu-

die van Morrow et al. (1990) is ook gepubliceerd als 'doctoral thesis' van O'Connor (1989). Waar in het artikel sprake is van een steekproef van 62 kinderen, rept de dissertatie stelselmatig over een steekproef van 53 kinderen. Zo'n verschil kan een onschuldige verklaring hebben. Maar vreemder is dat de gerapporteerde gemiddelden en standaarddeviaties tot op twee decimalen nauwkeurig gelijk zijn. Feitelson et al. (1993) rapporteren drie grootheden waaruit effectgroottes te destilleren zijn: gemiddelden en standaarddeviaties, F-waarden en percentages verklaarde varianties. Als men ze evenwel alle drie gebruikt, komt men tot verschillende geschatte effectgroottes. In de onderhavige review is overigens gebruik gemaakt van de (na covariantieanalyse) gecorrigeerde gemiddelden en standaarddeviaties. Karweit (1989) maakt een opmerking – waarnaar in het voorafgaande al is verwezen – die twijfel oproept over de verschillen tussen de programma's voor de experimentele en de controlegroep. Het lijkt erop alsof die verschillen meer omvatten dan alleen het voorlezen. Meyer et al. (1994) schrijven dat ze beschikken over de gegevens uit twee cohorten, elk ruim 300 leerlingen bevattend, maar rapporteren vervolgens alleen over het eerste cohort. Al met al roept de kwaliteit van vele rapportages vraagtekens op. Hier ligt zeker een taak voor tijdschriftredacties en beroepsverenigingen om met meer kracht te streven naar uniformiteit en kwaliteit van rapportage.

Het is te hopen dat de onderhavige review geen aanleiding zal geven tot de voldane constatering dat voorlezen op school gunstig is voor de mondelinge taalontwikkeling van kinderen. De empirische basis voor deze constatering is nog veel te smal: te weinig onderzoek, te veel onderzoek met kwalitatieve beperkingen, te veel onbegrepen variatie in de onderzoeksuitkomsten. Beter zou het zijn de onderzoeksinspanningen te vergroten en beter op elkaar af te stemmen. In dat licht worden – ter afsluiting – nog de volgende suggesties gedaan.

Er is behoefte aan studies waarin de duur en/of de intensiteit van de voorleesprogramma's experimenteel wordt gevarieerd. Is het bijvoorbeeld zo dat voorleesprogramma's die zich over een heel schooljaar uitstrekken tot grotere effecten aanleiding geven dan programma's die slechts enkele maanden beslaan? Het

zou verontrustend zijn als de duur van de programma's de effecten niet zou modereren, zoals op dit moment nog de conclusie lijkt te zijn.

Er is behoefte aan studies waarin de groeps-grootte systematisch wordt gevarieerd. Elsäcker (1994) vergeleek de effecten van het voorlezen aan groepen van vijf kleuters met het voorlezen aan klassikale groepen. Morrow en Smith (1990) vergeleken drie condities met elkaar: één-op-één, groepjes van drie kleuters en integrale kleuterklassen. Beide studies leidden tot de conclusie dat voorlezen in kleine groepen (van vijf, respectievelijk drie kleuters) het gunstigst is.

Er is behoefte aan studies waarin de stijl van voorlezen experimenteel wordt gevarieerd. Corcoran Nielsen (1990) vergeleek de effecten van vier verschillende voorleesstijlen, zonder overigens een nul-conditie in haar onderzoek te betrekken. Dickinson en Smith (1994) vergeleken in een survey-opzet drie voorleesstijlen wat betreft de effecten op de taalontwikkeling van kinderen. Burton (1993) demonstreerde in een aantrekkelijk 'multiple baseline design' dat de voorleesstijl van leerkrachten beïnvloedbaar is.

Naast de gesignaleerde behoeften aan experimenteel onderzoek is er ook behoefte aan onderzoek naar processen op micro-niveau. Hoe groot is bijvoorbeeld de kans dat voorgelezen kinderen de betekenis van een onbekend woord oppikken en vasthouden? Welke factoren beïnvloeden die kans? Relevant onderzoek in deze richting is gerapporteerd door Eller, Pappas en Brown (1988), Elley (1989), Leung en Pikulski (1990), Robbins en Ehri (1994), Sénéchal en Cornell (1993) en Sénéchal, Thomas en Monker (1995). Deze onderzoeken tonen aan dat voorlezen – veelal in een één-op-één situatie – inderdaad aanleiding geeft tot woordenschatuitbreiding. Maar ze roepen tevens de vraag op of hetzelfde effect op zou treden bij het voorlezen aan grotere groepen.

Noot

- 1 De studie van Karweit (1989), die over twee steekproeven rapporteert – een 'pre Kindergarten' en een 'Kindergarten' steekproef – is in het gegevensbestand als twee onafhankelijke studies beschouwd.

Literatuur

(De in de review opgenomen studies zijn met een asterisk aangeduid.)

- Baudoin, T., Meelis-Voorma, T., Moolenaar, P. Overmeijer, H., & Ros, B. (1990). *Jeugdliteratuur voor de beroepspraktijk; een leergang voor het HBO*. Groningen: Wolters-Noordhoff.
- Blok, H. (1992). *Lezen voor je plezier, hoe staat het daarmee bij leerlingen op de basisschool?* (SCO-rapport 286). Amsterdam: Stichting Centrum voor Onderwijsonderzoek van de Universiteit van Amsterdam.
- Bruner, J. (1983). *Child's talk: learning to use language*. New York: Norton.
- Bryk, A.S., Raudenbush, S.W., & Congdon, R.T. (1994). *HLM 2/3; Hierarchical linear modeling with the HLM/2L and HLM/3L programs*. Chicago: Scientific Software International.
- Burton, Sister A.M. (1993). *Influence of teacher-mediated strategies on the verbal responses of at risk first grade students during small group storybook read-aloud* (doct. diss.). Temple University.
- Bus, A.G., Ijzendoorn, M.H. van, & Pellegrini, A.D. (1995). Joint book reading makes for success in learning to read: a meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, 65, 1-21.
- *Cohen, D.H. (1968). The effects of literacy on vocabulary and reading achievement. *Elementary English*, 45, 209-213, 217.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Corcoran Nielsen, D.E. (1990). *The effects of four approaches to group storybook reading in kindergarten on story comprehension, story structure knowledge, and concepts about print* (doct. diss.). University of Minnesota.
- Dickinson, D.K., & Smith, M.W. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly*, 29, 104-122.
- Eller, R.G., Pappas, C.C., & Brown, E. (1988). The lexical development of kindergartners: learning from written context. *Journal of Reading Behavior*, 20, 5-24.
- Elley, W.B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, 24, 174-187.

- Elsäcker, W. van (1994). *Prentenboeken voorlezen aan allochtone en autochtone kleuters; een onderzoek naar het effect van groeps grootte op de interactie tijdens het voorlezen en de mate waarin het voorgelezen verhaal begrepen en onthouden wordt*. Tilburg: Werkverband Taal en Minderheden, Letterenfaculteit K.U.B. Tilburg.
- *Feitelson, D., Goldstein, Z., Eshel, M., Flasher, A., Levin, M., & Sharon, S. (1991). *Effects of exposure to literary language on kindergartner's listening comprehension and use of language* (unpublished manuscript). Haifa: University of Haifa.
- *Feitelson, D., Goldstein, Z., Iraqi, J., & Share, D. (1993). Effects of listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly*, 28, 70-79.
- *Feitelson, D., Kita, B., & Goldstein, Z. (1986). Effects of listening to series stories on first graders' comprehension and use of language. *Research in the Teaching of English*, 20, 339-356.
- Gleitman, L., Newport, E., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, 11, 43-79.
- Hedges, L.V., & Olkin, I. 1985. *Statistical methods for meta-analysis*. San Diego: Academic Press.
- *Karweit, N. (1989). The effects of a story-reading program on the vocabulary and story comprehension skills of disadvantaged prekindergarten and kindergarten students. *Early Education and Development*, 1, 105-114.
- *Khare, P. (1992). *Effects of oral versus read stories on children's creativity and sense of story structure* (doct. diss.). Birmingham, AL: University of Alabama.
- Leung, C.B., & Pikulski, J.J. (1990). Incidental learning of word meanings by kindergarten and first-grade children through repeated read-aloud events. In J. Zutell & S. McCormick (Eds.), *Literacy theory and research: analyses from multiple paradigms, thirty-ninth yearbook of the National Reading Conference* (pp. 231-240). Chicago: National Reading Conference.
- McCormick, S. (1981). *Reading aloud to preschoolers age 3-6: a review of the research* (paper presented at the Annual Meeting of the Southwest Regional Conference of the International Reading Association; ED 199 657).
- *Meyer, L.A., Stahl, S.A., Linn, R.L., & Wardrop, J.L. (1994). Effects of reading storybooks aloud to children. *Journal of Educational Research*, 88, 69-85.
- Morrow, L.M. (1988). Young children's responses to one-to-one story readings in school settings. *Reading Research Quarterly*, 23, 89-107.
- *Morrow, L.M., O'Connor, E.M., & Smith, J.K. (1990). Effects of a story reading program on the literacy development of at-risk kindergarten children. *Journal of Reading Behavior*, 22(3), 255-275.
- Morrow, L.M., & Smith, J.K. (1990). The effects of group setting on interactive story book reading. *Reading Research Quarterly*, 25, 213-231.
- Mullen, B. (1989). *Advanced BASIC meta-analysis*. Hillsdale (NJ): Lawrence Erlbaum Associates.
- Nijmeegse Werkgroep Taal didactiek (1992). *Taal didactiek aan de basis* (4e druk). Groningen: Wolters-Noordhoff.
- Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language*, 5, 1-15.
- O'Connor, E.M. (1989). *The effect of story reading as an instructional strategy on kindergarten children's literacy skill development*. New Brunswick, NJ: Rutgers State University of New Jersey.
- Orwin, R.G. (1983). A fail-safe N for effect size in meta-analysis. *Journal for Educational Statistics*, 8, 157-159.
- *Peterman, C.L. (1988). *Successful storyreading procedures: Working with kindergarten teachers to improve children's story understanding* (paper presented at the Annual Meeting of the National Reading Conference (38th, Tucson, AZ, November 29-December 3, 1988; ED 314 739).
- Radecki, K.K. (1987). *An annotated bibliography of the literature examining the importance of adults reading aloud to children* (ED 296 274). South Bend, IN: Indiana University.
- Robbins, C., & Ehri, L.C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology*, 86, 54-64.
- Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, 86, 638-641.
- Scarborough, H.S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review*, 14, 245-302.
- Schaerlaekens, A.M., & Gillis, S. (1987). *De taalverwerving van het kind*. Groningen: Wolters-Noordhoff.
- Schwarzer, R. (1989). *Meta-analysis program; program manual*. Berlin: Institut für Psychologie, Freie Universität Berlin.
- Sénéchal, M., & Cornell, E.H. (1993). Vocabulary acquisition through shared reading experiences.

Reading Research Quarterly, 28, 360-374.

Sénéchal, M., Thomas, E., & Monker, J. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology*, 87, 218-229.

*Sullivan, J.L. (1988). *Reading aloud to children and its effect on their attention span* (Master's Thesis, ED 292 077). Kean College (NJ).

Sulzby, E., & Teale, W. (1991). Emergent literacy. In R. Barr, M.L. Kamil, P. Mosenthal, & P.D. Pearson (Eds.), *Handbook of Reading Research, Vol. II* (pp. 727-757). White Plains (NY): Longman.

*Valdez-Menchaca, M.C., & Whitehurst, G.J. (1992). Accelerating language development through picture book reading: A systematic extension to day-care. *Developmental Psychology*, 28, 106-114.

Veen, A., Overmars, A., & Glopper, K. de (1995). *Boekenpret belicht: onderzoek naar een project Leesbevordering voor 0- tot 6-jarigen*. Utrecht: ISOR/Onderwijsonderzoek, Universiteit Utrecht.

Verhoeven, L. (1994). *Ontluikende geletterdheid*. Lisse: Swets & Zeitlinger.

Wells, G. (1986). *The meaning makers: children learning language and using language to learn*. Portsmouth: Heinemann.

Manuscript aanvaard 7-1-1997

Auteur

H. Blok is verbonden aan het SCO-Kohnstamm Instituut van de Faculteit Pedagogische en Onderwijskundige Wetenschappen van de Universiteit van Amsterdam.

Adres: Wibautstraat 4, 1091 GM Amsterdam. E-mail: henkb@educ.uva.nl

Abstract

Reading to young children in educational settings: a review of research

H. Blok. *Pedagogische Studiën*, 1997, 74, 233-248.

Reading to young children is an almost daily activity in kindergartens. One of the perceived benefits is the facilitation of literacy acquisition. This article reviews 11 studies into the effects of reading to young children in schools. The age of the children varied between 31 and 90 months. Dependent variables were classified in two domains: oral language and reading skills. The combined effect size for the oral language domain was $d = .64$, and for the reading domain $d = .22$. Although the effect in the oral language domain looks promising, caution is needed because the empirical evidence appears to be meager. Not only is the number of studies small, but a critical analysis of the design of the studies reveals general poor quality.

Handelingsaspecten bij het wederzijds leren van ervaren docenten

L.J. Galesloot, C.P. Koetsier en Th. Wubbels

Samenvatting

Samenwerking met collega's in netwerken van scholen lijkt een goede manier voor ervaren docenten om van elkaar te leren en kan worden gezien als een vorm van professionele ontwikkeling. In dit artikel bespreken we een aantal aspecten van het handelen van docenten die dit leren van elkaar kunnen bevorderen. Voor het beschrijven van de handelingsaspecten maken we gebruik van de resultaten van een exploratief ontwikkelingsonderzoek over door deelnemers zelf gestuurde scholennetwerken (Galesloot, 1994). Ook recentere bevindingen in literatuur en praktijk zijn verwerkt. Mede gezien de recente 'explosie van scholennetwerken' in Nederland lijkt het ons goed van deze 'leren bevorderende handelingsaspecten' verslag te doen. We belichten juist deze aspecten omdat ze naar onze mening ook van belang zijn in andere, niet door deelnemers gestuurde, netwerken. Bovendien blijven ons inziens deze – schijnbaar vanzelfsprekende – handelingsaspecten in de huidige netwerkpraktijk onderbelicht. De door ons verkregen gegevens over genoemd de handelingsaspecten kaderen we binnen de onderwijskundige principes voor het leren en opleiden van (oudere) volwassenen (Thijssen, 1988, 1992, 1996).

1 Inleiding

Scholennetwerken als een vorm van scholing hebben in het basisonderwijs in enkele regio's in Nederland al een zekere bekendheid (zie onder meer Galesloot, Jansen, Schildmeijer & Treffers, 1995). In het voortgezet onderwijs vond er tot voor kort vrijwel geen systematische uitwisseling plaats van praktijkervaringen en deskundigheden in netwerken van scholen (Eisner, 1984; Holly, 1989; Van Tulder, 1992). Toch beschikken ook goed functionerende

ervaren docenten in het voortgezet onderwijs over veel praktijkervaring en grote deskundigheden. Veel van die kennis en ervaring blijft echter binnen de school of het klaslokaal. Met enige regelmaat veronderstelden auteurs dat netwerken van scholen een nuttige en aantrekkelijke constructie zouden kunnen zijn voor nascholing (zie bijvoorbeeld Matthijssen, 1973; Louis & Rosenblum, 1981; Bollen, 1983, 1991; Eisner, 1984; Van den Berg & Vandenbergh, 1981, 1984, 1988; Ten Brinke & Galesloot, 1991; Ten Brinke, Galesloot & Schilperoord, 1991; Ten Brinke & Galesloot, 1992). Over de redenen waarom tot voor kort in het voortgezet onderwijs slechts geringe inhoudelijke uitwisseling in netwerkverband plaatsvindt, is weinig bekend (Louis, 1986; Pelkmans & Van Kuijk, 1986; Gould & Letven, 1987; Dalin, 1989; Pepels, Bastiaans & Van der Vegt, 1989; Van der Vegt & Pepels, 1991).

De term 'scholennetwerken' verdient eerst verheldering. We kunnen een onderscheid maken tussen 'netwerken' als een menselijke activiteit, een proces dat plaatsvindt tussen mensen (networking) en 'netwerken' als een organisatievorm (networks). Met 'networking' bedoelen we hier *een wijze van systematisch leren met en van elkaar als collega's*, die wij benoemen als 'horizontaal leren'. Met 'networks' duiden we op scholennetwerken als *meer of minder geformaliseerde structuren*, waarin deelnemers gedurende een bepaalde tijdsperiode van tevoren vastgelegde doelen nastreven. In een ontwikkeld model voor 'beroepsgericht horizontaal leren' via inhoudelijke samenwerking, als vorm van professionele ontwikkeling in netwerken van scholen (Galesloot 1994), verwijst 'beroepsgericht' naar dat deel van de professionele ontwikkeling dat niet is gericht op het schoolvak dat de docent geeft. Horizontaal leren onderscheidt zich van het 'verticale leren', het binnen de

nascholing meer gebruikelijke leren onder leiding van een expert. 'Beroepsgericht horizontaal leren', zoals hiervoor omschreven heeft veel kenmerken van intervisie (vergelijk De Hoop, 1992; Weerman, 1993). Het ontwikkelde (nascholings)arrangement (Galesloot, 1994) kenmerkt zich door vrijwillige deelname van ervaren docenten en schoolleiders van overeenkomstige delen van een beperkt aantal scholen, die gedurende langere tijd werken aan een gezamenlijk afgesproken thema, ondersteund door externe coördinatie. Deelnemers werken aan concrete onderwerpen die relevant zijn voor de eigen praktijk en tevens binnen de eigen school herkenbaar zijn en bepalen in onderling overleg zelf de frequentie, de duur en de plaats van de bijeenkomsten.

We bepleiten in dit artikel niet dat beroepsgericht horizontaal leren in de plaats moet komen van het leren onder leiding van een deskundige op een bepaald gebied. Ons inziens kan het evenwel binnen de mogelijkheden voor professionele ontwikkeling een plaats krijgen naast het leren van een expert.

Uit onderzoek blijkt dat er een aantal factoren/omstandigheden zijn, die stimulerend dan wel remmend werken op het functioneren van een scholennetwerk (zoals aantal en soort scholen, aantal deelnemers, de wijze van externe begeleiding). Een daarvan is een 'netwerkhouding' (Galesloot, 1994). Het (gaan) beschikken over een 'netwerkhouding' is bevorderend voor het leren met en van elkaar als collega's. Aan deze netwerkhouding kunnen verschillende handelingsaspecten worden onderscheiden. Dergelijke aspecten van het handelen van deelnemers in een scholennetwerk lijken van belang voor het wederzijds leren van ervaren docenten. Tegelijkertijd zijn deze aspecten schijnbaar vanzelfsprekend, bijna triviaal, edoch blijken ze in de netwerkpraktijk lang niet altijd onderdeel uit te maken van het handelen van de deelnemers. Vanuit het handelings-theoretisch paradigma beschouwen we hier intern handelen (cognities) en extern handelen (gedrag) als een twee-eenheid; beide componenten kunnen worden onderscheiden, niet gescheiden (vergelijk Peters, 1984; Corporaal, 1984; De Jong, 1988; Koetsier, 1991).

In dit artikel gaan we in op de vraag: 'Welke – schijnbaar vanzelfsprekende – handelingsas-

pecten van ervaren docenten bevorderen dat zij in netwerken van elkaars ervaringen kunnen leren?' Tevens gaan we na 'of de beschreven handelingsaspecten in verband gebracht kunnen worden met de onderwijskundige principes voor het leren en opleiden van (oudere) volwassenen van Thijssen, 1988, 1992, 1996' (zie paragraaf 2). Teneinde genoemde vragen te kunnen beantwoorden, baseren we ons op de resultaten van het exploratieve ontwikkelingsonderzoek van Galesloot (1994). Hij onderzocht – via een veldonderzoek (1988-1994) gecombineerd met een literatuurstudie – aan welke kenmerken inhoudelijke samenwerking in scholennetwerken moet voldoen om te kunnen functioneren als een vorm van professionele ontwikkeling. Het onderzoek levert gegevens over de vraag hoe ervaren docenten met en van elkaar kunnen leren.

In paragraaf 2 beschrijven we als theoretisch kader zes principes voor een nascholingsarrangement. Deze principes lijken bruikbaar om de resultaten van ons onderzoek in te kaderen, waardoor het vertrouwen in de geldigheid van de gevonden resultaten wordt vergroot. In paragraaf 3 geven we een samenvatting van de opzet van het exploratieve ontwikkelingsonderzoek. In paragraaf 4 gaan we vervolgens in op de resultaten van dit exploratieve onderzoek die van belang zijn voor het beantwoorden van bovengenoemde vraagstellingen, aangevuld met ervaringen uit recentere ontstane netwerken, zowel in door deelnemers zelf gestuurde als niet door henzelf gestuurde netwerken. Ook vermelden we recentere literatuurgegevens die het belang van de beschreven handelingsaspecten voor het leren met en van elkaar als collega's ondersteunen. We besluiten het artikel met het kaderen van de resultaten van ons onderzoek binnen de onderwijskundige principes van Thijssen, 1988, 1992, 1996 en een stellingname ten aanzien van horizontaal leren in scholennetwerken.

2 Theoretisch kader

In dit artikel proberen we handelingsaspecten, die het wederzijds leren van ervaren docenten kunnen bevorderen, te kaderen binnen de principes die Thijssen (1988, 1992, 1996) formuleert voor het leren en opleiden van (oudere)

volwassenen. Hij ontleent deze principes aan de praktijktheoretische noties van ervaren bedrijfsopleiders. Het gaat om twee principes ten aanzien van de *inhoud* in de scholennetwerken, twee principes ten aanzien van de *gewenste werkwijze* en twee principes ten aanzien van de *sfeer*.

Ten aanzien van de inhoud:

1. *Het herkenbaarheidsprincipe*. Dit principe houdt in dat nieuw aangeboden ideeën of ervaringen niet te ver van de eigen werkervaring dienen af te liggen. Hierbij gaat het erom dat de lerende, vanuit in het verleden opgedane ervaringen, de kern van nieuw aangeboden zaken althans gedeeltelijk kan herkennen, kan plaatsen. Zo ontstaat een basaal oriënteringskader waaraan door de lerende steeds gerefereerd kan worden als een 'bekende kapstok' voor nieuwe kennis. Dit principe is volgens Thijssen met name voor volwassenen met veel werkervaring van belang.

2. *Het toepasbaarheidsprincipe*. Dit principe houdt in de vraag naar de zinvolheid voor de eigen beroepspraktijk van hetgeen geleerd kan worden. Volwassenen vragen zich eerder dan een kind af of iets nieuws leren zinvol is, of ze er feitelijk iets mee kunnen doen en liever niet slechts op lange termijn. Illustraties van concrete toepassingsmogelijkheden zijn voor volwassenen van meer belang dan algemeen geldige, abstracte regels en modellen.

Teneinde het effectief leren te bevorderen, formuleert Thijssen tevens twee principes voor gewenste werkwijzen:

3. *Actief-structurende verwerking*. Dit principe houdt in de wenselijkheid dat volwassenen zelf actief en structurend met de stof bezig dienen te zijn. Het is voor volwassenen met het toenemen van leeftijd en levenservaring steeds belangrijker dat nieuwe leerstof actief wordt verwerkt. Ofschoon volwassenen, zeker in onzekere situaties, nogal eens 'de kat uit de boom kijken' en dan dus receptief, passief leren, is het van belang dat zij mentaal actief bezig zijn.

4. *Het feedbackprincipe*. Dit principe houdt in

dat aan een persoon informatie over zijn gedrag wordt verschaft, teneinde hem te helpen meer inzicht in het eigen functioneren te krijgen. Vanuit bepaalde, in de loop der jaren ontstane preconcepties (opgebouwd door ervaringen) kijken volwassenen naar nieuwe leersituaties. Door vanuit 'ingeslepen' preconcepties te kijken kunnen gauw vertekeningen en misverstanden ontstaan. Daarom is systematische feedback tijdens het leerproces van groot belang.

Daarnaast onderscheidt Thijssen (1992, 1996) nog twee principes, die wenselijkheden formuleren ten aanzien van de 'sfeer' van scholingsbijeenkomsten voor (oudere) volwassenen, namelijk:

5. *Aangepaste sociale context*. Dit principe houdt in dat het voor volwassenen met veel levenservaring van belang is dat het leren van nieuwe zaken vanuit een vriendelijke sociale sfeer gebeurt. Het is eveneens wenselijk om gelijkwaardigheid tussen docent en cursist en cursisten onderling na te streven en competitie en concurrentie te vermijden.

Volwassenen zijn doorgaans kritischer ten opzichte van de sociale context waarin zij leren dan kinderen en achten zich eerder belemmerd in een sociaal-onvriendelijke leersituatie.

6. *Aangepaste logistieke context*. Dit principe houdt in dat het voor volwassenen met veel levenservaring, die geconfronteerd worden met het leren van nieuwe zaken, belangrijk is dat zij de leersituatie aankunnen, dat wil zeggen kunnen overzien en beheersen c.q. beïnvloeden, bijvoorbeeld door een flexibele planning van tijd en inhoud. Dat heeft alles te maken met de onzekerheid die mensen met veel ervaring kennen, indien zij met nieuwe leersituaties worden geconfronteerd.

Thijssen meent, dat de theoretische verklaring achter de genoemde principes schuilt in het begrip 'ervaringsconcentratie', dit is het verschijnsel dat in het algemeen met de jaren de hoeveelheid ervaringen zal toenemen, maar de verscheidenheid ervan zal afnemen. Ervaringsconcentratie impliceert dat mensen zich allengs beter gaan 'thuisvoelen' in een steeds kleiner domein. Het wordt daarmee steeds moeilijker

de overstap te maken naar een nieuw ervaringsdomein buiten de vertrouwde sfeer. Kenmerkend voor deze theoretische invalshoek is, dat niet zozeer de volwassenheid geldt als verklarende factor, maar de mate waarin men niet (meer) gewend is met een verscheidenheid van nieuwe (leer)situaties om te gaan. Deze theorie geldt dus met name voor degenen die minder flexibel en 'open' omgaan met (geheel) nieuwe leersituaties. Als gevolg daarvan lijkt het van belang om in leersituaties van volwassenen met veel werkervaring, meer aandacht te besteden aan de hiervoor genoemde onderwijskundige principes om het leren te bevorderen. In het kader van dit artikel gaan wij niet nader in op dit begrip 'ervaringsconcentratie'.

In paragraaf 4 en 5 relateren we de uit ons onderzoek volgende handelingsaspecten aan vier van de zes onderwijskundige principes van Thijssen.

Principe 1 'het herkenbaarheidsprincipe' dat inhoudt dat nieuw aangeboden ideeën niet te ver van de eigen werkervaring dienen af te liggen, vormt een uitgangspunt van het gebruikte concept scholennetwerk. Immers, de bedoeling van de onderzochte netwerken is juist dat deelnemers concrete ervaringen inbrengen uit eigen school – en klaspraktijk. Evenzo vormt principe 6 'aangepaste logistieke context' een uitgangspunt van de onderzochte scholennetwerken, waarbij deelnemers in onderling overleg – naast de inhoud – zelf frequentie, duur en plaats van de bijeenkomsten bepalen. Doordat de principes 1 en 6 feitelijk als uitgangspunten van het onderzochte (nascholings)arrangement golden, is het niet relevant de handelingsaspecten met deze principes in verband te brengen. We schenken derhalve geen aandacht meer aan deze twee principes in het vervolg van dit artikel.

3 Exploratief ontwikkelingsonderzoek

Hieronder beschrijven we kort de opzet van en de werkwijze in het exploratief ontwikkelingsonderzoek naar inhoudelijke samenwerking van ervaren docenten en schoolleiders in scholennetwerken als een vorm van professionele ontwikkeling (Galesloot, 1994), voor zover dit

van belang is voor het beantwoorden van de vraagstelling van dit artikel.

3.1 Opzet van het onderzoek

In het veldonderzoek werden op basis van een 'naïef model' tussen 1988 en 1991 vier scholennetwerken geïnitieerd, waaraan docenten en schoolleiders deelnamen van in totaal 14 scholen. Deze netwerken waren gericht op beroepsgericht horizontaal leren via inhoudelijke samenwerking. Met 'naïef' bedoelen we dat het ontwikkelde arrangement niet was gebaseerd op bestaande wetenschappelijke theorieën, maar op redelijk lijkende uitgangspunten uit de schoolpraktijk (vergelijk Wester, 1991).

Daarnaast werd een literatuurstudie uitgevoerd naar professionele ontwikkeling via scholennetwerken met externe coördinatie. Bovendien werden gesprekken gevoerd met een aantal praktijkdeskundigen op dit terrein. Het doel van de literatuurstudie en de raadpleging van praktijkdeskundigen was om gegevens te verzamelen over de fenomenen netwerk(modellen), beroepsgerichte professionele ontwikkeling en (externe) netwerkcoördinatie.

3.2 Onderzoeksmethoden

De empirische gegevens over het ontwikkelde arrangement werden verkregen uit participerende observatie tijdens 90 netwerkbijeenkomsten, interviews met deelnemers aan de netwerken (in totaal 49) en analyse van documenten van 11 scholen.

De gegevens van de participerende observaties werden vastgelegd in verslagen. De interviews werden vastgelegd op audioband en woordelijk omgezet in interviewprotocollen. Het totaal aantal pagina's in documenten waarop informatie was vastgelegd in relatie met de onderzochte netwerken, bedroeg 235. Het aantal doorgelezen pagina's aan documenten varieerde per school van vele tientallen tot vele honderden.

De empirische gegevens over 'handelingsaspecten die het leren met en van elkaar als collega's stimuleerden' kwamen vooral voort uit de interviews en de participerende observaties. Tijdens de interviews werd in verband met deze handelingsaspecten uitgegaan van de startvragen: 'Bij het netwerk spelen remmende en bevorderende factoren van verschillende aard mogelijk een rol. Kunt u factoren noemen

die volgens u bevorderend of remmend zijn op het functioneren van het scholennetwerk?' en 'Ziet u zo'n inhoudelijk scholennetwerk als het uwe, als een mogelijk waardevolle zelfscholingsvorm?'. De gegevens leverden antwoorden op de onderzoeksvraag. In verband met de gewenste betrouwbaarheid gebruikten we alleen deze interviewgegevens, wanneer deze bevestigd werden door gegevens uit minstens één andere bron: de door de deelnemers goedgekeurde verslagen van de participerende observaties en/of de schooldocumenten. We merken hierbij op dat in de schooldocumenten slechts sporadisch een opmerking over genoemde handelingsaspecten werd gevonden. Categorieën om de verschillende handelingsaspecten te beschrijven werden ontleend aan het materiaal.

Bij het verzamelen en verwerken van de gegevens tijdens het veldonderzoek zijn verder diverse strategieën toegepast, teneinde de betrouwbaarheid en validiteit van de resultaten te bevorderen (zie Galesloot, 1994, pp. 42-54). Zo werden zowel de interviewprotocollen als de verslagen van de participerende observaties aan de deelnemers voorgelegd ter controle op volledigheid en ter validering van de interpretatie. Voorts zijn bij de verwerking van een deel van de gegevens, de interpretaties van de onderzoeker systematisch vergeleken met die van een tweede onderzoeker (vergelijk Brun-schot, 1990). De intersubjectieve betrouwbaarheid bleek hoog. In geval van niet-overeenstemming volgde overleg tussen hen. Dit overleg leidde in de meeste gevallen tot consensus. In de enkele gevallen waar er geen consensus werd bereikt, handhaafde de eerste onderzoeker op basis van de hem bekende contextinformatie zijn score.

4 Leren stimulerend handelen

Uit de analyse van de onderzoeksgegevens volgde onder meer een aantal categorieën, die we benoemd hebben als aspecten van 'leren stimulerend handelen'. Het betreft de volgende aspecten: 'vertrouwen in de bruikbaarheid van eigen ervaring', 'open opstelling', 'voorbereiden op vertellen over eigen schoolpraktijk', 'vertellen over eigen schoolpraktijk', 'actief luisteren en doorvragen', 'afspraken maken

over concreet uitproberen in eigen schoolpraktijk' en 'vertellen over nieuwe ervaringen'. Dit lijkt een triviale opsomming van gedrag te zijn. Uit de interviews en de participerende observaties bleek echter het grote belang van dit gedrag voor professionele ontwikkeling. Tevens bleek dat deelnemers niet zo bewust bezig zijn met dergelijk gedrag, wellicht door *het vanzelfsprekend lijkende karakter* ervan.

We bespreken hierna aspecten van 'leren stimulerend handelen' die bij docenten in de onderzochte netwerken zijn geconstateerd. Hoewel het handelen van *alle* deelnemers aan een netwerk, dat wil zeggen docenten, schoolleiders en externe coördinator in interactie met elkaar invloed lijkt te hebben op het leren met en van elkaar als collega's, beperken we ons in het kader van dit artikel tot de ervaren docenten.

In deze paragraaf beschrijven we steeds kort het handelingsaspect dat stimulerend dan wel remmend kan werken op het leren, gevolgd door een effect uit de praktijk van de deelnemers waaruit blijkt dat dit handelingsaspect tot wederzijds leren lijkt te leiden.

Vervolgens proberen we bij wijze van hypothese waar mogelijk aan te geven, waardoor het komt dat het besproken handelingsaspect bij ervaren docenten in de netwerkpraktijk, minder dan wenselijk voorkomt. Ten slotte geven we steeds aan welke principes van Thijssen (1992, 1996) een rol kunnen spelen bij de werkzaamheid van de handelingsaspecten.

Vertrouwen in de bruikbaarheid van eigen ervaring

Deelnemers gaven aan dat beroepsgericht horizontaal leren vereist dat zij zelf de eigen kennis en ervaring als potentieel nuttig voor collega's zien. Dit nu blijkt lang niet altijd het geval. Ook ander onderzoek wijst uit dat docenten 'het bij collega's in de keuken kijken' hoog op hun verlanglijst hebben staan, maar dat velen van hen twijfelen aan de waarde van eigen kennis en ervaring voor collega's van andere scholen (Holly, 1989; Smyth, 1989; Galesloot, 1994).

Citaat van een netwerkdeelnemer:

"Zo gauw er een ander - officieel - project of een boekje is, dan ontstaat er in de werkgroep bij ons op school een soort collectief minderwaardigheidscomplex, dat vind ik een heel merkwaardig soort ervaring."

Blijkbaar moeten docenten meer vertrouwen krijgen in de eigen mogelijkheden met betrekking tot het leren van elkaars ervaringsdeskundigheid en inzien dat zij de architect van eigen kennis zijn (vergelijk Wubbels, 1991). Het blijkt dat het accepteren van dit inzicht een in de tijd langzaam verlopend proces is (zie ook Engel, 1995; Galesloot, 1995; Bergen, 1996). Leraren (en anderen) zijn veelal gewend aan de verticale wijze van leren, het leren onder leiding van een expert. De wijze waarop we zelf in het verleden veel zaken hebben geleerd, bepaalt in sterke mate de wijze waarop we nu willen leren (vergelijk Bolhuis, 1988; Thijssen & De Greef, 1989). Dit verklaart wellicht deels de door deelnemers aangegeven noodzakelijke mentaliteitsverandering bij beroepsgericht horizontaal leren en het feit dat deze niet snel tot stand komt (Galesloot & Koetsier, 1996).

Het gaat bij dit handelingsaspect om het vertrouwen in de zinvolheid voor de eigen beroepspraktijk van eigen en andermans ervaring en om uitwisseling van concrete toepassingsmogelijkheden en daarmee om het tweede principe van Thijssen, 'het toepassingsprincipe'.

Open opstelling

De resultaten van de interviews en de participerende observaties lieten zien dat het werken met collega's van *andere* scholen vruchtbaarder en met minder remmingen verloopt dan met collega's van de eigen school. Ze leidt tot een open opstelling van de deelnemers. Delen van aanpakken van collega's van een andere school lijken eerder in de eigen school uitgeprobeerd te worden. 'Vreemde collega's' brengen andere ervaringen en andere benaderingswijzen in van eenzelfde problematiek. Ze halen eerder de door de eigen schoolcontext veroorzaakte 'blinde vlekken' weg, waardoor problemen c.q. mogelijke veranderingen in eigen praktijk eerder zichtbaar worden en verruiming van het eigen blikveld kan optreden (vergelijk Koot, 1986; Gould & Letven, 1987; Godfroy, 1989; en recenter Huberman, 1995; Slaats, 1995; Lieberman, 1996). Bovendien voelen netwerkdeelnemers zich vaak kwetsbaarder tegenover collega's van de eigen school en zijn daardoor minder geneigd om problematisch liggende zaken met 'eigen-school-collega's' te bespreken. Naar onze mening kan de communicatie tijdens netwerkbijeenkomsten opener

zijn juist door de geringere kwetsbaarheid ten opzichte van collega's van andere scholen in vergelijking met collega's van de eigen school.

Het gaat bij dit handelingsaspect om de openheid die nodig is om als ervaren volwassene nieuwe zaken te willen leren en 'ingeslepen' preconcepties te durven verlaten. De daarvoor onder meer gewenste voorwaarden worden ons inziens beschreven binnen het vierde en het vijfde principe van Thijssen, 'het feedback-principe' respectievelijk 'de aangepaste sociale context'.

Vorbereiden op vertellen over eigen schoolpraktijk

Uit de interviews en de participerende observaties bleek dat aan het systematisch vertellen over een bepaald aspect uit de eigen beroepspraktijk reflectie op die eigen praktijk vooraf gaat. Deze rol van reflectie wordt door ervaren docenten als duidelijke winst benoemd, onder meer omdat eigen expertise gebruikt wordt om de praktijk te bevragen en zo te verbeteren.

Volgens deelnemers leidt deze reflectie tot leereffecten op twee niveaus. In de eerste plaats blijkt uit opmerkingen van deelnemers dat het bewust stilstaan bij de eigen school – en klaspraktijk, teneinde deze aan andere collega's te kunnen overbrengen, wordt gezien als een belangrijk – en vaak niet verwacht – leereffect. Het systematisch stilstaan bij en voor jezelf verwoorden van concrete werkwijzen uit eigen praktijk blijkt volgens hen te kunnen leiden tot verbetering ervan. Wij veronderstellen dat door deze reflectie op aspecten van de eigen praktijk gewoonten meer bewust en meer doorzichtig worden (vergelijk Schön, 1987; Kilbourn, 1991; Korthagen, 1992; Meyer, 1992). Ook kunnen eventuele veranderingen in denken en doen zo op een duidelijker fundament van overwegingen steunen. In de tweede plaats wordt de praktische kennis van de netwerkdeelnemer ook beter communiceerbaar met collega's, waardoor wederzijdse leereffecten blijken toe te nemen (vergelijk Jansen & Van der Vegt, 1990). Bewust stilstaan bij de – vaak niet beseft – vanzelfsprekendheden in het eigen beroepshandelen, systematisch reflecteren op de eigen beroepsactiviteiten, is in onze schoolpraktijk echter geen gewoonte. Deelnemers aan de onderzochte scholennetwerken geven aan dat samenwerking in een scholen-

netwerk mogelijkheden biedt om op natuurlijke wijze, samen met collega's, te (gaan) reflecteren op de eigen onderwijspraktijk. Deelnemers werken immers aan voor henzelf belangrijke, door henzelf geformuleerde, gemeenschappelijke doelen, brengen vanuit eigen praktijk kennis en ervaring in en formuleren deze in gewone schooltaal (zie ook Galesloot & Bollen, 1995).

Citaat van een netwerkdeelnemer:

"Het is nuttig om als school na te denken over wat je aan het doen bent en zo'n netwerkbijeenkomst maakt dat noodzakelijk."

Bij dit handelingsaspect gaat het erom dat de lerende volwassene actief reflecteert op onderdelen van de eigen beroepspraktijk en daarmee tot (her)structurering ervan kan komen, aspecten van het derde principe van Thijssen, 'actief-structurende verwerking'.

Vertellen over eigen schoolpraktijk

Eerlijk en nuchter vertellen wat goed en wat niet goed gaat in eigen school- en klaspraktijk, werd door de deelnemers als stimulerend ervaren voor hun leerproces. Zij geven aan dat de valkuil bij het vertellen over eigen praktijk 'de successtory' is. Het naar voren brengen van alléén positieve ervaringen blijkt volgens deelnemers remmend te werken op het leren met en van elkaar als collega's.

In de netwerken bleek dat door collega's getoonde aanpakken van onderwijs, waarbij zowel sterke als zwakke kanten waren besproken, eerder tot concrete inspiratie en navolging in eigen schoolpraktijk leiden.

Citaat van een netwerkdeelnemer:

"Durven starten met je eigen onvolkomenheden. De successtory ligt bijna bij iedereen voor op zijn tong. Achterop liggen meer de dingen die niet zo goed gaan, waar je elkaar zou kunnen steunen."

Wij veronderstellen dat het mooier maken van de eigen werkelijkheid – soms onbewust – een normaal menselijk verschijnsel is (vergelijk Petri, 1985). Het met collega's bespreken van minder geslaagde aspecten van een bepaalde werkwijze of gebruikt onderwijsmateriaal wordt daardoor niet zo snel gedaan. Als collega's in een netwerk echter van elkaar willen leren, moeten zij niet alleen de mooie dingen vertellen, maar ook bereid zijn de zwakke kan-

ten van de eigen praktijk aanpakken te bespreken. Volgens deelnemers is hiervoor een goede, open sfeer in de netwerkbijeenkomsten noodzakelijk.

Het gaat bij dit handelingsaspect om het ook willen vertellen van zwakke kanten en daarmee om 'psychische veiligheid' voor de deelnemers. Derhalve vinden we dit handelingsaspect karakteristiek voor zowel het tweede als het vijfde principe van Thijssen, 'het toepasbaarheidsprincipe' respectievelijk 'de aangepaste sociale context'.

Actief luisteren en doorvragen

Uit de interviews en de participerende observaties bleek dat 'actief luisteren en doorvragen' bijdraagt tot het leren van de netwerkdeelnemers. Doordat zo informatie van concretere aard beschikbaar komt, blijkt er eerder in de eigen schoolpraktijk mee geëxperimenteerd te worden. Actief luisteren wil immers zeggen dat de luisteraar vragen stelt aan de verteller in de trant van: 'Kun je me precies vertellen hoe je dat doet?', 'Waarom heb je eigenlijk deze keuze gemaakt?' of 'Wat zijn eigenlijk voor- en nadelen van jouw aanpak?'. De bedoeling is 'door te vragen' over de praktijk van de verteller en juist daardoor meer aan de weet te komen. Ook de eerder genoemde door de eigen schoolcontext veroorzaakte 'blinde vlekken' kunnen door het actief luisteren naar en doorvragen van collega's van *andere* scholen sneller verhelderd worden. 'Actief luisteren en doorvragen' blijkt als handeling echter weinig voor te komen in bijeenkomsten van scholen-netwerken.

Wij veronderstellen dat een verklaring voor dit verschijnsel zou kunnen zijn dat docenten door het geïsoleerde karakter van het beroep in het algemeen niet gewend zijn om binnen de school opmerkingen over hun aanpak te krijgen of te geven. Bovendien hebben ervaren docenten op basis van langdurige ervaring een eigen – vaak succesvolle – stijl van werken ontwikkeld, waardoor het actief luisteren naar elkaars aanpakken ook niet wordt bevorderd. Doorvragen is volgens ons nauwelijks een deel van de bestaande schoolcultuur. Doorvragen ('Hoe doe je het precies? Waarom doe je het zo?') lijkt soms te wijzen op een gebrek aan respect voor wat de ervaren collega van de andere school over zijn beroepsuitoefening te vertel-

len heeft en wordt daarmee niet altijd positief gewaardeerd. Een vruchtbare communicatie ontstaat er dan niet (zie meer recent Huberman, 1995; Greenleaf, 1996).

Bij dit handelingsaspect gaat het om actief luisteren en doorvragen, om daardoor de door collega's aangeboden ervaringen zelf actief te verwerken en na te gaan in hoeverre deze in de eigen schoolpraktijk mogelijk bruikbaar zijn. Twee principes van Thijssen en wel 'het toepasbaarheidsprincipe' en 'actief-structurende verwerking', geven elementen van dit handelingsaspect weer.

Afspraken maken over het concreet uitproberen in eigen schoolpraktijk

Afspraken maken over het concreet uitproberen in eigen praktijksituatie, bleek volgens deelnemers stimulerend voor het leren met en van elkaar (zie ook Huberman, 1995).

Deelnemers gaven aan dat eerst bij het uitproberen in de eigen schoolsituatie blijkt, in hoeverre suggesties van collega's bruikbaar zijn en er dan eigenlijk pas sprake kan zijn van een daadwerkelijke stimulans voor het eigen leren. Uit het onderzoek bleek echter dat het daadwerkelijk uitproberen in de eigen school van aangedragen ideeën en suggesties van collega's uit netwerkbijeenkomsten, door de ervaren docenten lang niet altijd plaatsvindt.

Wij veronderstellen dat de keuze voor experimenteren met een nieuwe werkwijze door ervaren docenten minder snel gemaakt wordt dan door beginnende docenten, die immers nog een professioneel handelingsrepertoire moeten opbouwen (zie meer recent Huberman, 1995). De grote ervaring van veel netwerkdeelnemers bleek hier een schaduwzijde te hebben en remmend te kunnen werken op het eigen leerproces. Immers een eigen stijl van werken – juist bij ervaren docenten aanwezig – is veelal ontstaan op basis van jarenlange ervaringen. Het vertalen van zo op het eerste gezicht bruikbare ideeën van anderen naar eigen school en klas, kan daardoor een 'persoonlijke worsteling' zijn voor ervaren docenten. Wijzigingen in de eigen stijl van werken gaan vaak zeer langzaam. Het is begrijpelijk dat juist ervaren docenten, die al jaren over een adequaat professioneel handelingsrepertoire beschikken, snel resultaten van iets nieuws verwachten en bij uitblijven daarvan de nieuwe benadering verworpen.

Uitbreiding van het professionele repertoire lijkt ook bij ervaren docenten te lopen via de tijd en energie vergende 'probeerweg', dat wil zeggen een nieuwe aanpak een aantal keren in de klas uitproberen, bijstellen, verbeteren en uiteindelijk al of niet inbouwen in het bestaande repertoire van professioneel handelen. Het accepteren van deze leerweg lijkt vaak pas na geruime tijd te ontstaan. In dit verband moeten we ons realiseren dat veranderingen in de onderwijspraktijk een zaak is van lange adem (Lagerweij, 1994a; en recenter Galesloot & Bollen, 1995; Van Schijndel & Visser, 1995).

Bij dit handelingsaspect gaat het erom dat eerst bij het uitproberen in de eigen schoolsituatie de bruikbaarheid van suggesties van collega's blijkt: 'het toepasbaarheidsprincipe' van Thijssen.

Vertellen over nieuwe ervaringen

Rapporteren in netwerkbijeenkomsten over effecten van uitgeprobeerde suggesties in de eigen school bleek professionele ontwikkeling te stimuleren. Door de feedback van collega's komt men volgens deelnemers namelijk tot verbetering van aanpakken van onderwijs, die in eigen klas uitgeprobeerd zijn. Dit 'vertellen over nieuwe ervaringen' gebeurt echter slechts onregelmatig en niet vanzelfsprekend.

Wij veronderstellen dat hiervoor een aantal redenen kan zijn. In de eerste plaats zijn tijdens de bijeenkomsten geuite ideeën – volgens deelnemers aan de onderzochte netwerken – lang niet altijd zo ver geconcretiseerd dat afspraken over het uitproberen ervan gemaakt kunnen worden. Bovendien willen scholen vrij zijn om zelf te bepalen of zij binnen hun eigen cultuur iets hebben aan een idee uit een netwerk.

Maar ook als er in een school wel met ideeën uit de netwerkbijeenkomsten is geëxperimenteerd, is er zelden sprake van afspraken over het rapporteren van de resultaten daarvan. De verklaring hiervoor is wellicht dat men als deelnemer aan een netwerk in eerste instantie meedoet om er zelf als docent iets aan te hebben. Het nut van een systematische rapportage voor jezelf/de eigen praktijksituatie lijkt niet aanwezig. Men is blijkbaar niet gewend om dergelijke eerste ervaringen al aan elkaar te vertellen; het is eerder gebruikelijk pas te presenteren wanneer een en ander in redelijke mate is afgewogen en uitgekristalliseerd. Mo-

gelijk speelt ook nog mee dat men het uitgeprobeerde idee onbewust liever ziet als een eigen bedenkfel – zeker als het succes heeft – dan als een opbrengst van de samenwerking in een netwerk (Schön, 1987).

Evenals bij het eerder besproken handelingsaspect 'vertellen over eigen schoolpraktijk', gaat het bij het 'vertellen over nieuwe ervaringen' om psychische veiligheid voor de deelnemers, opdat zowel successen als mislukkingen verteld en geanalyseerd kunnen worden en de sociale situatie veilig genoeg is om feedback te geven en te ontvangen. Daarmee valt dit handelingsprincipe binnen drie van de principes van Thijssen, namelijk 'het toepasbaarheidsprincipe', 'aangepaste sociale context' en 'het feedbackprincipe'.

5 Handelingsaspecten en de principes van Thijssen

Het lijkt mogelijk de besproken 'het leren stimulerende handelingsaspecten' in te kaderen in de onderwijskundige principes van Thijssen voor het leren en opleiden van (oudere) volwassenen (1988, 1992, 1996). Daarmee wordt het inzicht groter in (de inrichting van) situaties waarin ervaren docenten wederzijds van elkaar kunnen leren, bijvoorbeeld in netwerken van scholen.

Zo lijken in principe 2 – het toepasbaarheidsprincipe – de volgende handelingsaspecten te plaatsen: 'vertrouwen in de bruikbaarheid van eigen ervaring', 'vertellen over eigen schoolpraktijk', 'actief luisteren en doorvragen', 'afspraken maken over concreet uitproberen in eigen schoolpraktijk' en 'vertellen over nieuwe ervaringen'.

In principe 3 – de actief-structurende verwerking – plaatsen we de handelingsaspecten: 'voorbereiden op vertellen over eigen schoolpraktijk' en 'actief luisteren en doorvragen'; in principe 4 – het feedbackprincipe – de handelingsaspecten: 'open opstelling', en 'vertellen over nieuwe ervaringen'. Ten slotte kunnen binnen principe 5 – aangepaste sociale context – de volgende handelingsaspecten geplaatst worden: 'open opstelling', 'vertellen over eigen schoolpraktijk', en 'vertellen over nieuwe ervaringen'.

6 Nabeschouwing

In dit artikel hebben we een aantal handelingsaspecten beschreven, die van belang lijken voor deelnemers aan netwerken van scholen die zich richten op professionele ontwikkeling. De besproken handelingsaspecten zijn een resultaat van een exploratief ontwikkelingsonderzoek, vergeleken met het bestaande wetenschappelijke kennisbestand over scholennetwerken en nascholing van docenten. Naast de besproken resultaten van het ontwikkelingsonderzoek geven indrukken uit eigen ervaringen in recenter ontstane scholennetwerken – in het kader van de Tweede Fase Voortgezet Onderwijs – aan dat genoemde handelingsaspecten ook van belang zijn in niet door deelnemers zelf gestuurde netwerken (Galesloot, 1995; Galesloot & Bollen, 1995; Galesloot & Koetsier, 1996). Ook recentere literatuurgegevens lijken dit vermoeden te steunen (Engel, 1995; Huberman, 1995; Van Schijndel & Visser, 1995; Slaats, 1995; Bergen, 1996; Greenleaf, 1996).

Tevens zijn de besproken handelingsaspecten geplaatst binnen een theoretisch kader voor het leren en opleiden van (oudere) volwassenen, zie Thijssen (1988, 1992, 1996), waardoor het vertrouwen in de geldigheid van de besproken handelingsaspecten toeneemt.

De gepresenteerde onderzoeksresultaten geven overigens niet aan hoe de gewenste handelingsaspecten bij netwerkdeelnemers bevorderd zouden kunnen worden. Hiertoe is vervolgonderzoek noodzakelijk. Bij dergelijk onderzoek lijkt het onder meer van belang dat men zich richt op het bij deelnemers ontwikkelen van een groter vertrouwen in de waarde en bruikbaarheid van eigen ervaring en deskundigheid voor andere ervaren collega's. Voorts dat deelnemers vaker en systematischer gaan reflecteren op aspecten van eigen beroepspraktijk en actiever gaan luisteren naar collega's tijdens netwerkbijeenkomsten. Ten slotte zou het onderzoek zich moeten richten op het vergroten van de bereidheid van deelnemers om met nieuwe ideeën in eigen klaspraktijk te experimenteren en erover op netwerkniveau te rapporteren.

Netwerken van scholen bieden naar onze mening 'natuurlijke' mogelijkheden voor effectief én plezierig leren met en van elkaar als

collega's, doordat deelnemers onder meer kunnen werken aan zelf geformuleerde doelen. Bovendien brengen zij vanuit eigen schoolpraktijk kennis en ervaring in, die ook nog in verstaanbare schooltaal wordt geformuleerd. Weerstand tegen veranderingen en vernieuwingen, die ook – en soms juist – bij ervaren docenten aanwezig zijn, kunnen door deze wijze van horizontaal leren worden verkleind. Er lijken echter nogal wat voorwaarden vervuld te moeten worden om horizontaal leren in een netwerk – als een betrekkelijk nieuwe vorm van nascholing – goed te laten functioneren, zowel voorwaardenscheppende maatregelen als ook handelingsaspecten. In dit artikel hebben we enkele handelingsaspecten besproken, die vanzelfsprekend lijken, maar dit in de praktijk van een scholennetwerk niet altijd blijken te zijn.

Literatuur

- Berg, R. van den, & Vandenberghe, R. (1988). *Onderwijsvernieuwing op een keerpunt*. Tilburg: Zwijzen.
- Berg, R. van den, & Vandenberghe, R. (1984). *Groot-schaligheid in de onderwijs-vernieuwing*. Tilburg: Zwijzen.
- Berg, R. van den, & Vandenberghe, R. (1981). *Onderwijsinnovatie in verschuivend perspectief*. Tilburg: Zwijzen.
- Bergen, Th. (1996). *Docenten scholen docenten*. Oratie 13-09-96. Katholieke Universiteit Nijmegen.
- Bolhuis, S.M. (1988). Het leren van volwassenen. In B. van Gent & A.L.T. Notten (Red.), *Inleiding tot volwasseneneducatie* (pp. 75-86). Meppel: Boom.
- Bollen, R. (1983). *Beleid in school*. Voordracht gehouden op de conferentie voor schoolleiders, 's Hertogenbosch, 15 november 1983.
- Bollen, R. (1991). De onderkant van het systeem II. *NGL Weekblad*, 23(18), 4-6.
- Brinke, J.S. ten, & Galesloot, L.J. (1991). Ervaringen met een Europees en een Nederlands scholennetwerk: toekomstperspectieven. *Verslagboek VELON congres 1990 Veldhoven* (pp. 317-323). Tilburg: VELON/Quasar.
- Brinke, J.S. ten, Galesloot, L.J., & Schilperoord, P. (1991). Extern gecoördineerde netwerken. In F.A.J. Korthagen, O. de Jong & Th. Rensman (Red.), *Van anders naar beter. Doelgericht inspe-*
- len op veranderingen in secundair en tertiair onderwijs* (pp. 133-135). Utrecht: WCC/IVLOS, Universiteit Utrecht.
- Brinke, J.S. ten, & Galesloot, L.J. (1992). Ervaringen met AVO-scholen netwerken. In P. Leenheer, S. Stawski & A. Jansen (Red.), *Nascholing in de markt. MESO-bibliotheek* (pp. 127-130). Culemborg: Educaboek/VELON.
- Brunschot, R.A.M. van (1990). *Begeleiding van beginnende docenten in het technisch onderwijs*. Academisch proefschrift, Technische Universiteit Eindhoven.
- Corporaal, A.H. (1984). *Begeleidingsmodellen. Een literatuurstudie ten behoeve van de praktische vorming van (aanstaande) onderwijsgeevenden*. Academisch proefschrift. Lisse: Swets & Zeitlinger.
- Dalin, P. (1989). *Organisatie-ontwikkeling in school en onderwijs*. Alphen aan den Rijn: Samsom.
- Eisner, E. (1984). *Educational research and educational practice*. Voordracht gehouden op het symposium Onderwijswetenschap en Onderwijspraktijk, Utrecht, 28 september 1984.
- Engel, P.G.H. (1995). *Facilitating Innovation: An action-oriented approach and participatory methodology to improve innovative social practice in agriculture*. Academisch proefschrift, Landbouwuniversiteit Wageningen.
- Galesloot, L.J. (1994). *Collegiale netwerken van ervaren docenten en schoolleiders: een veldonderzoek gecombineerd met een literatuurstudie over beroepsvervolmaking*. Academisch proefschrift, Universiteit Utrecht. De Lier: Academisch Boeken Centrum.
- Galesloot, L.J. (1995). Werken in een netwerk: mogelijkheden en grenzen. In Th. Bergen, M. Kamp & H. Peters (Red.), *Leren zelfstandig leren in de tweede fase*. (pp. 33-45). Nijmegen: UNILU, Universiteit Nijmegen.
- Galesloot, L.J., Jansen, H., Schildmeijer, J., & Treffers, K. (1995). *Samenwerken in netwerken: netwerkvarianten, succesfactoren en voorbeelden uit de WSNS-praktijk*. Utrecht: Algemeen Pedagogisch Studiecentrum.
- Galesloot, L.J., & Bollen, R. (1995). *Scholennetwerken in de 2e Fase: scharnier tussen wens en werkelijkheid*. Utrecht: IVLOS/APS.
- Galesloot, L. & Koetsier, C. (1996). Leren in scholennetwerken. In D. van Veen & W. Veugelers (Red.) *Vernieuwing van leraarschap en lerarenopleiding* (pp. 145-157). Apeldoorn/Leuven:Garant.
- Godfroij, A.J.A. (1989). Netwerken van organisaties.

- In P.A.E. van de Bunt & K. Nijkerk (Red.), *Handboek organisatie: aflevering 23* (BC 25.100 pp. 1-34). Alphen aan den Rijn: Samsom.
- Gould, S., & Letven, E. (1987). A center for interactive professional development. *Educational Leadership*, 45(8), 49-52.
- Greenleaf, R.K. (1996). *De dienaar als leider: over leiden om te dienen versus leiden om de macht*. Schoonhoven: Academic Service.
- Holly, M.L.H. (1989). Teacher professional development: Perceptions and practices in the USA and England. In M.L. Holly & C.S. McLoughlin (Red.), *Perspectives on teacher development* (pp. 173-203). New York/Philadelphia/London: The Falmer Press.
- Hoop, A. de (1992). Uitzicht op intervisie. *Tijdschrift voor leerlingbegeleiding*, 15(1), 23-26.
- Huberman, M. (1995). Networks that alter teaching: Conceptualizations, exchanges and experiments. *Teachers and Teaching, Theory and Practice*, 1(2), 193-211.
- Jansen, Th., & Vegt, R. van der (1990). De institutionalisering voorbij: over de duurzaamheid van vernieuwingen in de school. *Pedagogisch Tijdschrift*, 15(3), 140-151.
- Jong, J.A. de (1988). 'Je wordt aan het denken gezet.' *Ontwikkeling en evaluatie van een model voor de schoolstage van de PABO*. Academisch proefschrift, Universiteit Utrecht. Utrecht: ISOR.
- Kilbourn, B. (1991). Self-monitoring in teaching. *American Educational Research Journal*, 28(4), 721-736.
- Koetsier, C.P. (1991). *Een brug tussen opleiding en praktijk. Een studie over de individuele eindstage als brug tussen lerarenopleiding en praktijk*. Academisch proefschrift, Universiteit Utrecht. Utrecht: WCC.
- Koot, J. (1986). Wie niet verder kijkt, reikt ook niet verder. Internationaal netwerk van scholen: IADAS. *Vernieuwing van onderwijs, opvoeding en maatschappij*, 44(4), 9-14.
- Korthagen, F.A.J. (1992). Reflectie en de professionele ontwikkeling van leraren. *Pedagogische Studiën*, 69(2), 112-123.
- Lagerweij, N. (1994). *De lange adem van onderwijsvernieuwing*. Leuven/Apeldoorn: Garant.
- Lieberman, A. (1996). Creating intentional learning communities. *Educational Leadership* 54(3), 51-55.
- Louis, K.S. (1986). Permanent innovation: Theoretical and administrative issues in institutionalizing educational change. In P.N. Appelfhof e.a. (Red.), *Naar beter onderwijs* (pp. 128-143). Tilburg: Zwijsen.
- Louis, K.S., & Rosenblum, S. (1981). *Linking R & D with schools: Designing and managing inter-organizational networks*. Washington, DC: The National Institute of Education.
- Matthijssen, H.A.J.M. (1973). Bijscholing en innovatiebeleid. In W.H.A. Renders (Red.), *Tijd voor bijscholing: een bundel opstellen over bijscholing aan onderwijsgeevenden* (p.138). Den Bosch: Commissie Bijscholing HNO-MSPO.
- Meyer, S. (1992). Cultivating reflection-in-action in trainer development. *Adult Learning*, 3(4), 16-18.
- Pelkmans, A., & Kuijk, J. van (1986). *Samenwerking tussen scholen voor voortgezet onderwijs: een studie naar interorganisatorische relaties*. Nijmegen: ITS.
- Pepels, C.G.N., Bastiaans, W.A.J., & Vegt, R. van der (1989). Netwerk voor schoolverbetering. In R. Vandenberghé & R. van der Vegt (Red.), *Onderwijsvernieuwing* (pp. 15-29). Lisse/Amsterdam: Swets & Zeitlinger.
- Peters, J.J. (1984). Professioneel didactisch handelen. In J.H.C. Vonk, J.S. ten Brinke, P. Licht & J. Stakenborg (Red.), *Onderzoek en ontwikkeling* (pp. 178-185). 's-Gravenhage: SVO/VULON.
- Petri, M.W. (1985). *Samenwerking voor vernieuwing. Verslag van een onderzoek naar overdraagbaarheid in het innovatieproces midden-school*. Hoevelaken: VPLC/Berk.
- Schijndel, J. van, & Visser, E. (1995). *Verslag van een HAVO-HBO aansluitingsproject 1991-1995*. Amsterdam: Hogeschool van Amsterdam.
- Schön, D.A. (1987). *Educating the reflective practitioner: Towards a design for teaching and learning in the profession*. San Francisco/London: Jossey-Bass Publishers.
- Slaats, H. (1995). Het verschil delen: netwerkontwikkeling als vernieuwingscontext voor professionele instellingen. Academisch proefschrift, Universiteit Utrecht. Utrecht: ISOR.
- Smyth, J. (1989). Administrative leadership in the intellectual transformation of schooling. In M.L. Holly & C.S. McLoughlin (Eds.), *Perspectives on teacher development* (pp. 219-235). London/New York/Philadelphia: The Falmer Press.
- Thijssen, J.G.L. (1988). *Bedrijfsopleidingen als werkterrein*. Den Haag: Delwel/Roi.
- Thijssen, J.G.L. (1992). Ervaringsconcentratie: drempeel voor kwalificatievernieuwing in de tweede loopbaanhelft. *Gedrag en Organisatie. Tijdschrift voor sociale, arbeids- en organisatiepsychologie* 5(6), 428-445.

- Thijssen, J.G.L. (1996). *Leren, leeftijd en loopbaanperspectief*. Academisch proefschrift, Katholieke Universiteit Brabant, Tilburg.
- Thijssen, J.G.L., & Greef, J. de (1989). Het leren van volwassenen: theorieën en grondprincipes. In J. Kessels & C. Smit (Red.), *Handboek opleiders in organisaties: Aflevering 11* (pp. A3.6-1-22). Deventer: Van Loghum Slaterus/Kluwer.
- Tulder, M. van (1992). *Nascholing en onderwijsvernieuwing*. Academisch proefschrift, Katholieke Universiteit Nijmegen.
- Vegt, R. van der, & Pepels, C.G.M. (1991). Scholennetwerken als infrastructuur voor vernieuwingswerk: over zienswijzen en begrippen. *Nederlands tijdschrift voor opvoeding, vorming en onderwijs*, 7(2), 98-110.
- Weerman, C. (1993). Intervisie in het onderwijs – een momentopname. In F.A.J. Korthagen, L.C. Don & S. van den Heuvel (Red.), *Stilstaan bij onderwijs in beweging; opleiden binnen veranderende kaders* (pp. 315-317). Utrecht: VELON/IVLOS/WCC.
- Wester, F. (1991). *Strategieën voor kwalitatief onderzoek*. Muiderberg: Coutinho.
- Wubbels, Th. (1991). Terug naar de spil van de onderwijsverbetering: de docent. *School*, 19(5), 28-35.

Manuscript aanvaard 20-2-1997

Auteurs

L.J. Galesloot, universitair docent/onderzoeker,
C.P. Koetsier, universitair docent/onderzoeker,
Th. Wubbels, hoogleraar, zijn werkzaam bij het Interfacultair Instituut voor de Lerarenopleiding, Onderwijsontwikkeling en Studievaardigheden (IVLOS) van de Universiteit Utrecht.

Adres: IVLOS – Universiteit Utrecht, Postbus 80.127,
 3508 TC Utrecht

Abstract

Aspects of behaviour and the reciprocal learning of experienced teachers.

L.J. Galesloot, C.P. Koetsier & Th. Wubbels. *Pedagogische Studiën*, 1997, 74, 249-260.

For experienced teachers, cooperating with colleagues within the context of school networks seems to be a good way to learn from each other and can also be seen as a form of professional development. In this article, we discuss some aspects of mental and material behaviour that promote this 'learning from each other'. For their description we will use the results of an exploratory developmental investigation on self-steering networks (Galesloot, 1994). We have also included more recent findings taken from relevant literature and the experiences of existing networks. Moreover, because of the recent 'explosion of school networks' in the Netherlands, it seems helpful to report on these 'learning stimulating aspects of behaviour'. We focus on these aspects because in our opinion these are also important in other networks not run by the participants. Besides, in our opinion these aspects of behaviour – that seem so obvious – remain underexposed in current network practice. We place our findings on these aspects within the framework of the educational guidelines furthering the professional development of (older) adults (Thijssen, 1988, 1992, 1996).

De lineariteitsillusie bij het oplossen van meetkunde- problemen door leerlingen van het secundair onderwijs

D. De Bock, L. Verschaffel en D. Janssens*

Samenvatting

Lineair proportionele modellen zijn ongetwijfeld één van de meest voorkomende modellen bij het representeren en oplossen van zowel zuivere als toepassingsgerichte wiskundige problemen. Volgens diverse auteurs leidt de centrale plaats die deze lineaire modellen innemen in het huidig reken/wiskundeonderwijs ertoe dat vele leerlingen de neiging ontwikkelen om deze modellen ook te gebruiken in (probleem)situaties waarin dit niet gepast is. In onderhavig artikel wordt verslag uitgebracht van twee studies over deze lineariteitsillusie bij 12-13- en bij 15-16-jarigen die geconfronteerd worden met vraagstukken over lengte en oppervlakte van gelijkvormige vlakke figuren; tevens gaan deze studies de invloed na van zelfgemaakte en aangeboden tekeningen op het optreden van deze illusie. Globaal genomen bieden de resultaten overtuigende steun voor de sterkte van de lineariteitsillusie bij deze groepen van studenten.

Inleiding

Lineair proportionele (of recht evenredige) verbanden tussen grootheden krijgen in het wiskundeonderwijs op elementair en secundair niveau bijzonder veel aandacht. Men zegt dat twee grootheden recht evenredig zijn als het enige malen groter of kleiner worden van de ene tot gevolg heeft, dat de andere evenveel malen groter of kleiner wordt. Dat dergelijke verbanden zoveel aandacht krijgen in het elementair en secundair (wiskunde)onderwijs, komt doordat voor heel wat problemen uit de zuivere en de toegepaste wiskunde lineaire

functies het onderliggend wiskundig model vormen. Voorbeelden van lineair proportionele verbanden uit de (zuivere) wiskunde zijn legio: de omtrek van een cirkel is proportioneel met de straal, de grootte van een hoek is proportioneel met de bijbehorende cirkelboog. Ook tal van levensechte probleemsituaties kunnen adequaat gemodelleerd worden in termen van lineaire proportie: het aantal liter verf nodig om een egaal oppervlak te schilderen is proportioneel met de oppervlakte ervan; het volume en het gewicht van een bepaalde vloeistof zijn proportioneel.

De toenemende vertrouwdheid van leerlingen en studenten met deze lineaire modellen heeft echter ook een keerzijde: dit kan er namelijk toe leiden dat leerlingen de misvatting ontwikkelen dat deze lineaire modellen een universele toepasbaarheid hebben, waardoor zij geneigd zijn deze modellen ook te gaan gebruiken in (probleem)situaties waarvoor ze in het geheel niet gelden. Dit fenomeen wordt aangeduid als *de lineariteitsillusie*. In onderhavig artikel rapporteren we twee studies over deze illusie in één bepaald toepassingsgebied, nl. vraagstukken over lengte en oppervlakte van gelijkvormige vlakke figuren.

Het principe dat aan de basis ligt van dit soort toepassingsopgaven is bekend: een vergroting of verkleining met factor r vermenigvuldigt lengtes met factor r , oppervlaktes met factor r^2 en inhouds met factor r^3 . Bijvoorbeeld: wanneer men de straal van een bol verdubbelt, verdubbelt de omtrek van elke cirkel op het boloppervlak, verviervoudigt de boloppervlakte en verachtvoudigt het bolvolume. Bij halvering van de straal, halveert de omtrek van elke cirkel, wordt de boloppervlakte door vier en het bolvolume door acht gedeeld. Essentieel hierbij is dat deze factoren niet afhankelijk zijn van de specifieke figuur waarover sprake is in de opgave (d.w.z. of het nu om een vierkant, een cirkel of een onregelmatige

* De auteurs danken R. Stroobants, E. Ratincx en G. Van Vaerenbergh voor hulp bij de (statistische) verwerking van de onderzoeksgegevens.

figuur gaat) maar enkel van de grootheden in kwestie (d.w.z. lengte, oppervlakte of volume). Hoewel in de literatuur regelmatig melding gemaakt wordt van het ongeoorloofd toepassen van een lineair proportioneel model bij het vergroten en verkleinen van figuren, werd – voorzover ons bekend – hierover nauwelijks systematisch empirisch onderzoek verricht. Mede daardoor lopen de meningen over de ernst en de hardnekkigheid van deze misvatting sterk uiteen (National Council of Teachers of Mathematics, 1989; Streefland, 1984; Treffers, 1987).

Studie 1

Onderzoeksopzet en -materiaal

Het eerste onderzoek werd uitgevoerd in het begin van het schooljaar 1995-1996 in het eerste leerjaar algemeen secundair onderwijs van een Vlaams lyceum dat haar leerlingen rekruteert uit een brede waaier van nabijgelegen basisscholen.

Aan het onderzoek participeerden 120 leerlingen, verdeeld in drie groepen (Groep I, II en III met resp. 40, 42 en 38 leerlingen). Elke groep bestond uit twee intacte klassen, één uit de 'Moderne' afdeling (met zes wekelijkse lestijden wiskunde) en één uit de 'Latijnse' afdeling (met vier wekelijkse lestijden wiskunde).

Het onderzoek bestond uit twee fasen. In een eerste fase lieten we alle 120 leerlingen – zonder speciale voorafgaande aanwijzingen of instructies – dezelfde schriftelijke toets oplossen bestaande uit 12 experimentele items (Toets 1). Deze 12 experimentele items hadden alle betrekking op schaalvergroting van vlakke figuren en werden geconstrueerd rond drie verschillende soorten figuren: 4 items hadden betrekking op vierkante (V), 4 op cirkelvormige (C) en 4 op grillige figuren (G). Per categorie waren er telkens 2 proportionele items en 2 niet-proportionele items. In Tabel 1 geven we een voorbeeld van één proportioneel en één niet-proportioneel item voor elk van de drie soorten figuren uit Toets 1.

Tenslotte werd bij elk van de drie catego-

Tabel 1

Voorbeelden van experimentele items

Schaalvergroting van een vierkantige figuur

Proportioneel item:

Om een gracht te graven rond een vierkant stuk weiland met zijde 100 m, heeft boer Gust ongeveer 4 dagen nodig. Hoeveel dagen zal hij ongeveer nodig hebben om een gracht te graven rond een vierkant stuk weiland met zijde 300 m? (Antwoord: 12 dagen)

Niet-proportioneel item:

Om een vierkant stuk grond met zijde 200 m te bemesten, heeft boer Karel ongeveer 8 uur nodig. Hoeveel uur zal hij ongeveer nodig hebben om een vierkant stuk grond met zijde 600 m te bemesten? (Antwoord: 72 uur)

Schaalvergroting van een cirkelvormige figuur

Proportioneel item:

Om met een zeilschip rond een cirkelvormig eiland met een diameter van 70 km te varen, heb je ongeveer 6 uur nodig. Hoeveel uur heb je ongeveer nodig om rond een cirkelvormig eiland met een diameter van 140 km te varen? (Antwoord: 12 uur)

Niet-proportioneel item:

Om een cirkelvormig bloemenperkje met een diameter van 10 m aan te leggen, heb je ongeveer 400 gram bloemenzaad nodig. Hoeveel gram bloemenzaad heb je ongeveer nodig om een cirkelvormig bloemenperkje met een diameter van 20 m aan te leggen? (Antwoord: 1600 gram)

Schaalvergroting van een grillige figuur

Proportioneel item:

Op een landkaart van België in een atlas meet de afstand Genk-Leuven ongeveer 5 cm en de afstand Genk-Gent ongeveer 11 cm. Op een landkaart van België die vooraan in de klas hangt, meet de afstand Genk-Leuven ongeveer 20 cm. Hoeveel meet de afstand Genk-Gent op deze landkaart? (Antwoord: 44 cm)

Niet-proportioneel item:

Op een landkaart van België in een atlas meet de afstand Genk-Tongeren ongeveer 2 cm en is de oppervlakte van België ongeveer 250 cm². Op een landkaart van België die vooraan in de klas hangt, meet de afstand Genk-Tongeren ongeveer 6 cm. Hoe groot is de oppervlakte van België op deze landkaart? (Antwoord: 2250 cm²)

riën van figuren (V, C en G) ook nog één bufferitem (rond een andere bewerking) geconstrueerd. Om spieken te voorkomen en om volgorde-effecten te kunnen neutraliseren, werden de items in twee (inverse) volgorden aangeboden. Op de antwoordbladen was er niet enkel plaats voorzien voor het antwoord, maar ook voor het maken van bewerkingen, bedenkingen, tekeningen, enz.

Twee weken na de eerste toetsafname werden de drie groepen van leerlingen geconfronteerd met een tweede toets (Toets 2), die een parallelversie van Toets 1 was. Groep I, die fungeerde als controlegroep, kreeg – net zoals tijdens de eerste toetsafname – geen verdere instructies of hulp. De leerlingen van Groep II kregen de instructie om een schets of tekening te maken alvorens de oplossing neer te schrijven. Deze aanbeveling werd bij de aanvang van de toets gedaan en verduidelijkt aan de hand van een voorbeeld-item, – uiteraard *geen* niet-proportionele opgave. De leerlingen van Groep III tenslotte kregen bij elk item een correcte tekening aangeboden. Bij wijze van voorbeeld geven wij in Figuur 1 de tekening die hoort bij het niet-proportioneel item rond 'vierkanten' uit Tabel 1.

Hypothesen

Ten eerste verwachtten we, op basis van de beschikbare literatuur (Berté, 1993; Freudenthal, 1983; NCTM, 1989; Rogalski, 1982; Rouche, 1992), dat de overgrote meerderheid van de leerlingen sterk zou 'lijden' onder de lineariteitsillusie. Dit impliceert dat we voor de

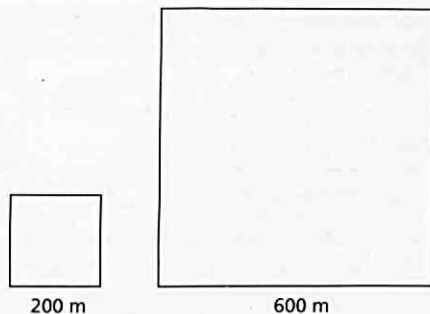
6 niet-proportionele items zeer weinig, en voor de 6 proportionele items zeer veel correcte antwoorden voorspelden.

Ten tweede veronderstelden we dat een schets of tekening van de probleemsituatie een positieve invloed zou uitoefenen op de prestaties van de leerlingen, in het bijzonder voor wat betreft de niet-proportionele items. De argumentatie daarvoor is dat de kans op het doorbreken van de lineariteitsillusie – en dus op het geven van het juiste antwoord – door zo'n schets of tekening toeneemt. Immers, door aan leerlingen te vragen *zelf een tekening te maken* worden zij a.h.w. gedwongen tot het construeren van een eigen (mentale) representatie van de essentiële elementen en relaties in de probleemsituatie (Pólya, 1945; Schoenfeld, 1992). Bij de niet-proportionele items zal dit ertoe leiden dat zij de onjuistheid van een stereotiepe, lineair proportionele redenering inzien en tevens de precieze aard van het niet-lineaire verband tussen de gegeven en de gevraagde grootheid ontdekken. Tevens gingen we ervan uit dat *het aanreiken van een juiste tekening* de kans op het ontdekken van de niet-lineariteit van het verband – en dus op het vinden van het juiste antwoord – nog sterker zal verhogen dan enkel de suggestie tot het maken van een bijpassende tekening. De reden hiervoor is dat niet alle leerlingen die als instructie krijgen om een tekening te maken, erin zullen slagen om een *correcte, bruikbare* tekening te produceren. Uitgaande van deze hypothesen werden de volgende voorspellingen gemaakt: in Groep I zullen de resultaten van beide toetsafnamen

Opgave:

Om een vierkant stuk grond met zijde 200 m te bemesten, heeft boer Karel ongeveer 8 uur nodig. Hoeveel uur zal hij ongeveer nodig hebben om een vierkant stuk grond met zijde 600 m te bemesten?

Tekening:



Figuur 1. Voorbeeld van een item zoals aangeboden in Groep III

(Toets 1 en Toets 2) nagenoeg identiek zijn; in Groep II zal er beter gepresteerd worden op Toets 2 dan op Toets 1, en in Groep III tenslotte zal de toename van het aantal correcte antwoorden nog sterker zijn dan in Groep II. Bovendien voorspelden we dat de verwachte stijgingen in Groep II en III hoofdzakelijk het gevolg zouden zijn van een daling van het aantal fouten gebaseerd op een lineair proportionele redenering bij de niet-lineaire items.

Ten derde verwachtten we een effect van de aard van de figuur waarover in het vraagstuk sprake is, op het aantal correcte antwoorden – met name op de niet-proportionele items – en wel in die zin dat de meeste correcte oplossingen gegeven zouden worden op de niet-proportionele opgaven rond vierkante figuren (V-items) en het minst bij de items met grillige figuren (G-items). Onze argumentatie voor de derde hypothese luidt als volgt. Voor de niet-proportionele items rond vierkanten (de V-items) bestaan er drie adequate oplossingsstrategieën: (1) het grote vierkant a.h.w. 'betegelen' met kleine vierkantjes, (2) van beide vierkanten de oppervlakte berekenen met behulp van een meetkundige formule ('oppervlakte vierkant = zijde maal zijde'), (3) het algemeen principe 'als de lengte (zijde) maal r , dan de oppervlakte maal r^2 ' toepassen. Voor cirkels (de C-variant) is de eerste strategie (nl. een exacte betegeling) onmogelijk en wordt ook de tweede strategie moeilijker (omwille van de grotere complexiteit van en de geringe vertrouwelijkheid van de leerlingen met de betreffende formule). Voor grillige figuren (de G-variant) tenslotte valt ook de tweede oplossingsmethode weg (er bestaat immers geen formule) en kan men dus enkel nog terugvallen op de derde methode, nl. het algemeen principe toepassen.

Analyse

De antwoorden van de leerlingen op de proportionele en op de niet-proportionele items werden geanalyseerd aan de hand van een analyse-schema bestaande uit meerdere categorieën. Maar omdat vrijwel alle fouten toe te schrijven zijn aan een verkeerdelijk toegepaste lineair proportionele redenering, zal hierna enkel gebruik gemaakt worden van het onderscheid tussen 'correcte' en 'incorrecte' antwoorden (zie De Bock, Verschaffel & Janssens, 1996).

De toetsing van de hypothesen geschiedde aan de hand van een variantie-analyse met als onafhankelijke variabelen de between-variabele 'Groep' (Groep I, II en III) en de within-variabelen 'Toets' (Toets 1 versus Toets 2), 'Proportionaliteit' (proportionele versus niet-proportionele items) en 'Figuur' (vierkante, cirkelvormige en grillige figuren). Het aantal 'Correcte antwoorden' fungeerde in de analyse als afhankelijke variabele. Significante hoofd- en interactie-effecten werden achteraf verder onderzocht met behulp van a posteriori LSD-tests.

Resultaten

Tabel 2 geeft een overzicht van het percentage correcte antwoorden voor de drie groepen van leerlingen (I, II en III) voor de proportionele en de niet-proportionele items over vierkanten (V), cirkels (C) en grillige figuren (G) in Toets 1 en Toets 2.

De eerste hypothese werd door deze onderzoeksresultaten op overtuigende wijze bevestigd. De variantie-analyse bracht een zeer sterk hoofdeffect van de taakvariabele 'Proportionaliteit' aan het licht ($F(1,117) = 4994.915$; $MSE = .233446$; $p < .01$): voor de drie groepen en de twee toetsen samen, bedroegen de percentages correcte antwoorden op de proportionele en

Tabel 2

Percentage correcte antwoorden in de drie groepen 12-13-jarigen op de verschillende soorten proportionele en niet-proportionele items tijdens Toets 1 en 2

	Toets 1						Toets 2					
	Proportionele items			Niet-proportionele items			Proportionele items			Niet-proportionele items		
	V	C	G	V	C	G	V	C	G	V	C	G
I	96	98	89	5	0	1	99	96	85	3	0	0
II	93	95	89	6	1	0	93	95	95	4	2	0
III	91	91	87	4	3	0	93	89	89	8	5	1

niet-proportionele items respectievelijk 92% en 2%. Dit betekent dat de proportionele items door deze 12-13-jarigen bijna altijd juist opgelost werden, terwijl de niet-proportionele items zelden of nooit correct beantwoord werden.

De tweede hypothese in verband met de verwachte positieve invloed van tekeningen op de prestaties van de leerlingen, werd niet bevestigd. De variantie-analyse bracht noch een significant 'Groep × Toets' interactie-effect aan het licht, noch een significant 'Groep × Toets × Proportionaliteit' interactie-effect: in geen van de drie groepen nam het percentage correcte antwoorden tussen beide toetsafnamen op significante wijze toe, noch in het algemeen, noch in het bijzonder voor wat betreft de niet-proportionele items. Blijkbaar was noch de instructie om een tekening te maken (Groep II), noch het aanbieden van een tekening (Groep III) sterk genoeg om tegenwicht te bieden aan het overweldigend effect van de lineariteits-

illusie bij deze 12-13-jarigen.

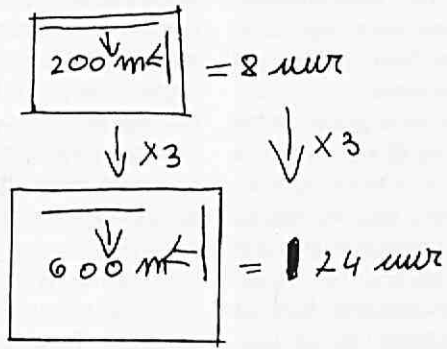
Zoals voorspeld op basis van de derde hypothese, had de aard van de figuur een significante invloed op het percentage juiste antwoorden. Uit de variantie-analyse kwam een significant hoofdeffect voor deze taakvariabele naar voren ($F(2,234) = 12.963$; $MSE = 0.086106$; $p < .01$). De totale percentages correcte antwoorden voor de V-, C- en G-items lagen in de verwachte richting, nl. respectievelijk 49%, 48% en 45%, maar aanvullende LSD-tests wezen uit dat enkel het verschil tussen de V- en de G-items en tussen de C- en de G-items significant was (telkens op het 1% niveau). Er werd ook geen interactie-effect gevonden tussen de taakvariabelen 'Proportionaliteit' en 'Figuur'. Dit betekent dat het hoofdeffect van de aard van de figuur zich in dezelfde mate manifesteerde bij de proportionele en bij de niet-proportionele items.

Aanvullend bij de resultaten met betrekking

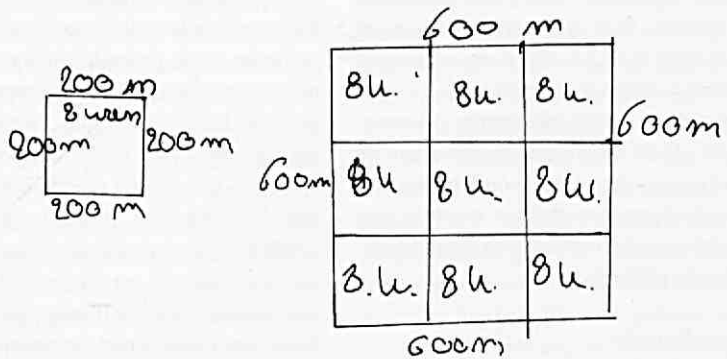
Opgave:

Om een vierkant stuk grond met zijde 200 m te bemesten, heeft boer Karel ongeveer 8 uur nodig. Hoeveel uur zal hij ongeveer nodig hebben om een vierkant stuk grond met zijde 600 m te bemesten? (Antwoord: 72 uur)

Tekening 1:



Tekening 2:



Figuur 2. Twee voorbeelden van zelfgemaakte tekeningen bij het niet-proportioneel item rond vierkanten uit Tabel 1

tot de drie hypothesen, vermelden we ook nog enkele markante bevindingen die gebaseerd zijn op een kwalitatieve analyse van de leerlingnotities.

Ten eerste analyseerden we alle notities van de leerlingen van Groep I, II en III bij de niet-proportionele items uit Toets 1. Deze analyse bracht aan het licht dat – voor de drie groepen samen – slechts bij 2% van de niet-proportionele items uit Toets 1 een tekening werd aangetroffen. De leerlingen maakten dus zelden of nooit *spontaan* een schets of tekening als onderdeel van hun oplossingsproces van de niet-proportionele items.

Uit de analyse van de notities van de leerlingen van Groep II bij Toets 2 bleek verder dat de expliciete instructie om een tekening te maken slechts bij 46% van de niet-proportionele items was opgevolgd. Daarenboven produceerden de leerlingen die de instructie wel opvolgden meestal een tekening van gebrekkige kwaliteit (voor een voorbeeld van een zelfgemaakte tekening zonder en met een goede voorstellingswaarde, zie resp. tekening 1 en 2 uit Figuur 2). Vermoedelijk beschouwden de leerlingen het maken van een (goede) tekening als puur tijdverlies omdat zij het 'problematisch' karakter van de niet-proportionele opgaven in het geheel niet doorhadden. Deels verklaart dit het geringe effect van de instructie.

Ten derde gingen we in de antwoordbladen van de leerlingen van Groep III bij Toets 2 na of zij de gegeven tekeningen op de een of andere manier hadden 'bewerkt'; elke aanduiding op de gegeven tekening – hoe minimaal of hoe rudimentair ook – werd als een 'bewerking' gescoord. Deze analyse bracht aan het licht dat slechts 6% van de tekeningen bij de niet-proportionele opgaven door de leerlingen 'bewerkt' werden. Ook dit gegeven suggereert dat de leerlingen over het algemeen weinig of geen aandacht besteedden aan of gebruik maakten van de gegeven tekeningen (vermoedelijk omdat zij de moeilijkheidsgraad van de niet-proportionele items sterk onderschatten). Dit zou mede kunnen verklaren waarom ook het verwachte positieve effect van de aangeboden tekeningen uitbleef.

Conclusie en discussie

De eerste studie leverde overtuigende steun voor de basishypothese dat de lineariteitsillusie

een heel sterke rol speelt bij 12-13-jarige leerlingen die geconfronteerd worden met toepassingsopgaven over lengte en oppervlakte van gelijkvormige figuren. Het uitermate geringe aantal correcte antwoorden van deze 12-13-jarigen op de niet-proportionele items roept evenwel onmiddellijk de vraag op hoe sterk deze illusie nog zou zijn bij leerlingen die ouder zijn en over een veel grotere 'vakinhoudelijke bagage' beschikken om deze illusie te overwinnen. Daarom werd het eerste onderzoek hernomen bij een groep 15-16-jarigen.

Studie 2

Onderzoeksopzet en -materiaal

Het tweede onderzoek werd uitgevoerd op het einde van het schooljaar 1995-1996 in hetzelfde lyceum van de eerste studie. Alle 222 leerlingen van het vierde leerjaar van deze secundaire school namen eraan deel. Deze leerlingen volgden een brede waaier van studierichtingen van het algemeen secundair onderwijs.

Op basis van een aantal beschikbare gegevens over de leerlingen werden zij direct in drie gelijkwaardige groepen ingedeeld: Groep I (waarin de leerlingen geen speciale aanvullingen of hulp kregen), Groep II (waarin de leerlingen de instructie kregen om telkens eerst een tekening te maken) en Groep III (waarin bij elk item een correcte tekening werd aangeboden). Voor een gedetailleerde beschrijving van de drie groepen zie De Bock, Verschaffel en Janssens, 1996.

Voor het vervolgonderzoek werden dezelfde 12 experimentele items gebruikt als in studie 1. (zie Tabel 1). Ook de wijze waarop de toets werd afgenomen, de vormgeving van de antwoordbladen, de scoringswijze van de antwoorden en de procedure voor de data-analyse bleven onveranderd. De voorspellingen die aan de basis lagen van studie 2 waren dezelfde als die van de eerste studie. Alleen veronderstelden we – in verband met de eerste hypothese – dat 15-16-jarigen globaal genomen minder zouden 'lijden' onder de lineariteitsillusie dan de 12-13-jarigen, en dus betere resultaten zouden behalen op de niet-proportionele items. Men mag immers niet vergeten dat het leerplan wiskunde voor het derde en vierde jaar van het secundair onderwijs ruimschoots aandacht

besteedt aan kennis en vaardigheden behulpzaam voor het oplossen van dit soort toepassingsopgaven (verkenning van de meetkunde van het vlak, studie van niet-lineaire verbanden, ...).

Resultaten

Tabel 3 geeft een overzicht van het percentage correcte antwoorden voor de drie groepen van 15-16-jarigen (I, II en III) voor de proportionele en de niet-proportionele items over vierkanten (V), cirkels (C) en grillige figuren (G).

Tabel 3
Percentage correcte antwoorden in de drie groepen 15-16-jarigen op de verschillende soorten proportionele en niet-proportionele items

	Proportionele items			Niet-proportionele items		
	V	C	G	V	C	G
I	91	97	97	26	11	1
II	89	91	97	26	20	5
III	89	93	97	39	21	7

De eerste hypothese werd bevestigd. Een variantie-analyse bracht wederom een zeer sterk hoofdeffect van de taakvariabele 'Proportionaliteit' aan het licht ($F(1,219) = 1591.635$; $MSE = .484985$; $p < .01$): voor de drie groepen samen bedroegen de percentages van correcte antwoorden op de proportionele en de niet-proportionele items resp. 93% en 17% (te vergelijken met resp. 92% en 2% in studie 1). Een variantie-analyse waarin de resultaten van studie 1 en 2 te zamen betrokken werden, bracht – zoals verwacht – een significant interactie-effect aan het licht tussen de factoren 'Leeftijd' en 'Proportionaliteit' ($F(1,340) = 25.334$; $MSE = .362244$; $p < .01$): enkel de stijging van het aantal correcte antwoorden op de niet-proportionele items van het eerste naar het vierde jaar was significant (op het 1% niveau).

De hypothese over de verwachte positieve invloed van de tekeningen werd wederom niet bevestigd. De globale percentages correcte antwoorden voor de groepen I, II, en III lagen weliswaar in de verwachte richting – nl. respectievelijk 54%, 55% en 58% –, maar de onderlinge verschillen waren opnieuw te klein om een significant hoofdeffect op te leveren van de variabele 'Groep'.

Zoals voorspeld in de derde hypothese en

overeenkomstig de resultaten van de eerste studie, was de aard van de figuur wel van belang. De variantie-analyse bracht immers een significant hoofdeffect voor deze taakvariabele aan het licht ($F(2,438) = 24.667$; $MSE = .157318$; $p < .01$). De globale percentages correcte antwoorden voor de V-, C- en G-items lagen in de verwachte richting (resp. 60%, 56% en 51%), en uit aanvullende LSD-tests bleek dat alle onderlinge verschillen significant waren (op het 1% niveau). Bovendien werd er een interactie-effect gevonden tussen de taakvariabelen 'Proportionaliteit' en 'Figuur' ($F(2,438) = 69.574$; $MSE = .179697$; $p < .01$). De percentages correcte antwoorden op de niet-proportionele items lagen in de verwachte richting (resp. 30%, 17% en 4% voor de V-, C- en G-items), terwijl we – enigszins tot onze verrassing – vaststelden dat de percentages van correcte antwoorden op de proportionele items helemaal in de omgekeerde richting gingen (resp. 90%, 94% en 97% voor de V-, C- en G-items). Aanvullende LSD-tests brachten aan het licht dat voor de niet-proportionele items alle onderlinge verschillen significant waren (op het 1% niveau), terwijl bij de proportionele items enkel het verschil tussen de V- en de G-items en tussen de V- en de C-items significant was (op het 1%, resp. 5% niveau). Het vastgestelde hoofdeffect van de aard van de figuur was dus uitsluitend te wijten aan de geconstateerde verschillen bij de niet-proportionele items.

Aanvullend beschrijven we – zoals bij studie 1 – enkele bijkomende onderzoeksgegevens in verband met het gebruik van tekeningen in het oplossingsproces van de niet-proportionele items. Door het groter aantal zelfgemaakte en/of bewerkte tekeningen bij en correcte antwoorden op de niet-proportionele items in studie 2 was het mogelijk om na te gaan in hoeverre er een rechtstreeks verband was tussen tekening en antwoord en toetsten we de (on)afhankelijkheid van beide variabelen met behulp van een χ^2 -toets. Uit de χ^2 -toetsen bleek dat de nulhypothese in de drie groepen kon worden verworpen (op een significantieniveau van 1%). Hieruit blijkt dus dat zowel het (al dan niet spontaan) maken van een tekening als het effectief gebruik maken van een gegeven tekening de kans verhogen op het ontdekken van de onjuistheid van een stereotiepe, lineair proportionele redenering bij een

niet-proportioneel item en bijgevolg ook op het vinden van de correcte oplossing. Anderzijds blijkt uit deze analyse ook dat men het effect daarvan ook niet moet overschatten: het maken of gebruiken van een tekening zorgt er in geen geval voor dat men de juiste oplossing *gegarandeerd* vindt.

Doordat de 15-16-jarigen al wat meer correcte antwoorden gaven op de niet-proportionele items, was het ook mogelijk om via de analyse van de leerlingnotities te onderzoeken welke van de drie eerder beschreven oplossingswegen – de methode van het ‘betegelen’, de weg van het gebruiken en uitrekenen van de meetkundige formule of het toepassen van het algemene principe – door de leerlingen het meest gevolgd werd. Vastgesteld werd dat het toepassen van de relevante meetkundige formule veruit de meest gebruikte aanpakstrategie was. Voor het niet-proportioneel item rond een vierkante figuur uit Tabel 1 betekent dit de oppervlakte van beide vierkanten effectief berekenen ($200 \times 200 = 40\,000\text{ m}^2$ en $600 \times 600 = 360\,000\text{ m}^2$) en vervolgens de uitkomst bepalen van de deling $360\,000 : 40\,000$, nl. 9. Negentig procent van de leerlingen die dit item correct beantwoordden, pasten deze tweede oplossingsvariant toe (soms in combinatie met één van de andere methoden); ‘zuivere’ toepassingen van de eerste en de derde oplossingsmethode troffen we slechts in respectievelijk 7% en 3% van de gevallen aan. Vooral het geringe aandeel van de – hier toch voor de hand liggende – betegelingsstrategie was verrassend. Eigenlijk is ‘betegelen’ een zeer eenvoudige, context-gebonden strategie, waarvoor in feite weinig of geen ‘schoolse’ wiskundekennis vereist is. Dat zo weinig leerlingen van deze voor de hand liggende informele strategie gebruik maakten, houdt wellicht verband met het in tal van onderzoeken vastgestelde denkbeeld (‘belief’) bij vele leerlingen en studenten dat het oplossen van een wiskundig probleem vooral een kwestie is van het vinden en toepassen van de juiste formule (Schoenfeld, 1992; Verschaffel & De Corte, in press).

Discussie

De enorm sterke negatieve invloed van de lineariteitsillusie op het oplossen van toepas-

singsopgaven over lengte en oppervlakte van gelijkvormige vlakke figuren, komt als belangrijkste conclusie uit dit onderzoek naar voren. Bij de 12-13-jarigen was het effect van deze illusie ronduit overweldigend, maar ook 15-16-jarigen bleken er nog in zeer sterke mate door te worden beïnvloed. Bij geen van beide groepen studenten was er een effect van zelfgemaakte of aangeboden tekeningen, al gaven de resultaten bij de 15-16-jarigen wel aan dat het *effectief* (gebruik) maken van een tekening de kans op het vinden van de juiste oplossing bij dit soort vraagstukken verhoogt. De vraag of de gevonden sterkte van de lineariteitsillusie typisch is voor vraagstukken rond gelijkvormige meetkundige figuren dan wel bij andersoortige toepassingsopgaven even sterk is, kan op basis van de twee gerapporteerde studies uiteraard niet beantwoord worden. Daarvoor is verder onderzoek vereist waarin de relatieve sterkte van de lineariteitsillusie nagegaan wordt bij verschillende categorieën van probleemsituaties.

Het foutief denkproces dat bij de leerlingen tot de foutieve antwoorden op de niet-proportionele items heeft geleid, is eenvoudig te omschrijven: deze leerlingen volgden een lineair proportionele redenering in probleemsituaties waarvoor dit ongepast was. Dit brengt ons echter bij de vraag welke aspecten van of elementen uit de kennisbasis van deze leerlingen leidden tot deze verkeerde denkweg. Waarschijnlijk gaat het om meerdere, sterk aan elkaar gerelateerde elementen zoals: (1) tekorten in de vakinhoudelijke conceptuele kennisbasis van de leerlingen (zoals hun gebrekkige kennis van het begrip ‘recht evenredigheid’), (2) gebrekkige beheersing van waardevolle heuristieken (zoals het maken van een anschouwelijke voorstelling van de probleemsituatie in de vorm van een schets of tekening), (3) ontoereikende metacognitieve vaardigheden (zoals de vaardigheid in het vooraf plannen en het achteraf controleren van de ‘uitvoerende handelingen’) en (4) allerlei inadequate denkbeelden (‘beliefs’) in verband met het oplossen van wiskundige problemen (zoals het denkbeeld dat het oplossen van een vraagstuk nooit meer dan een paar minuten in beslag neemt, of de opvatting dat het oplossen van een vraagstuk altijd neerkomt op het toepassen van een gekende formule) (zie Verschaffel & De Corte,

in press). In de toekomst zullen we trachten om – o.a. via individuele interviews – een duidelijk beeld te krijgen van de bijdrage van en samenwerking tussen al deze kennisaspecten in het tot stand komen van het beschreven foutieve denkproces van de leerlingen bij niet-proportionele opgaven.

Een volgende vraag voor verder onderzoek is welke aspecten van het huidige wiskundeonderwijs mede verantwoordelijk zijn voor het ontstaan en de ontwikkeling van de lineariteitsillusie bij leerlingen. Een plausibele verklaring hiervoor is dat het wiskundeonderwijs te eenzijdig de aandacht richt op rekentechnische aspecten en stereotiepe voorstellingswijzen van proportionele relaties, terwijl de geschiktheid van een lineair model voor een gegeven probleemsituatie (bijv. in concurrentie met andere wiskundige modellen) maar zelden het voorwerp van reflectie en discussie uitmaakt (Greer, 1993; Verschaffel & De Corte, 1996). Doch ook dit dient nader onderzocht te worden in constaterend onderzoek waarin de prestaties en moeilijkheden van leerlingen uitdrukkelijker gerelateerd worden aan kenmerken van het wiskundeonderwijs dat ze tot dan toe gevolgd hebben.

Tenslotte rijst de zeer belangrijke vraag hoe men leerlingen beter dan thans het geval is kan behoeden voor en/of wapenen tegen de valstrik van de lineariteitsillusie, zowel in het algemeen als in het bijzonder bij problemen over lengte en oppervlakte van gelijkvormige vlakke figuren. Met het oog op het beantwoorden van deze vraag zal er gewerkt moeten worden aan de ontwikkeling en evaluatie van nieuw instructiemateriaal. In dit materiaal zal in elk geval ruimschoots aandacht moeten worden besteed aan het leren wiskundig modelleren, d.w.z. aan het opstellen en het in vraag stellen van een wiskundig model voor een gegeven situatie; ook de interpretatie van de uitkomsten in het licht van een toegepast model zal daarbij een belangrijk aandachtspunt moeten vormen (Greer, 1993; Verschaffel & De Corte, 1996). Waardevolle bouwstenen daarvoor zijn te vinden in een aantal oudere en meer recente vakdidactische publicaties (zie bijv. Feys, 1995; Groupe d'Enseignement Mathématique, 1994; NCTM, 1994; Streefland, 1984; Treffers, 1987).

Literatuur

- Berté, A. (1993). *Mathématique dynamique*. Paris: Nathan.
- De Bock, D., Verschaffel, L., & Janssens, D. (1996). *De lineariteitsillusie: een exploratief onderzoek*. (Intern rapport). Leuven: Centrum voor Instructiepsychologie en -Technologie/Departement Didactiek, K.U.Leuven.
- Feys, R. (1995). Meten en metend rekenen. In L. Verschaffel & E. De Corte (Eds.), *Naar een nieuwe reken/wiskundendidactiek voor de basisschool en de baseducatie. Deel 3: Verder bouwen aan gecijferdheid* (pp. 99-135). Brussel/Leuven: Studiecentrum Open Hoger Onderwijs (StOHO)/Acco.
- Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Dordrecht: Reidel.
- Greer, B. (1993). The modeling perspective on word problems. *Journal of Mathematical Behavior*, 12, 239-250.
- Groupe d'Enseignement Mathématique (1994). *De question en question. Mathématiques 2*. Brussel: Didier Hatier.
- National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (1994). *Curriculum and evaluation standards for school mathematics addenda series Grades 5-8. Understanding rational numbers and proportions*. Reston, VA: NCTM.
- Pólya, G. (1945; 2nd edition, 1957). *How to solve it*. Princeton: Princeton University Press.
- Rogalski, J. (1982). Acquisition de notions relatives à la dimensionalité des mesures spatiales (longueur, surface). *Recherches en Didactique des Mathématiques*, 3(3), 343-396.
- Rouche, N. (1992). *Le sens de la mesure*. Brussel: Didier Hatier.
- Schoenfeld, A. (1992). Learning to think mathematically: problem solving, metacognition, and sense making in mathematics. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334-370). New York: Macmillan.
- Streefland, L. (1984). The design of a mathematics course. A theoretical reflection. *Educational Studies in Mathematics*, 15, 109-135.
- Treffers, A. (1987). *Three dimensions. A model of goal and theory description in mathematics instruction. The Wiskobas project*. Dordrecht: Reidel.

Verschaffel, L., & Corte, E. De (1996). Lerend modelleren en interpreteren van vraagstukken. Een exploratief onderwijsexperiment bij leerlingen van de bovenbouw van de basisschool. *Pedagogische Studiën*, 73, 322-337.

Verschaffel, L., & Corte, E. De (in press). Word problems. A vehicle for promoting authentic mathematical understanding and problem solving in the primary school. In T. Nunes & P. Bryant (Eds.), *How do children learn mathematics?* Hillsdale, NJ: Erlbaum.

Manuscript aanvaard 22-5-1997

Auteurs

D. De Bock doceert wiskunde aan de Economische Hogeschool Sint-Aloysius (EHSAL) te Brussel en is Wetenschappelijk Medewerker aan de Faculteit Psychologie en Pedagogische Wetenschappen van de Katholieke Universiteit Leuven.

L. Verschaffel is Onderzoeksleider bij het Nationaal Fonds voor Wetenschappelijk Onderzoek en Hoogleraar (deeltijds) aan de Faculteit Psychologie en Pedagogische Wetenschappen van de Katholieke Universiteit Leuven.

D. Janssens is Docent aan de Faculteit Wetenschappen en verantwoordelijke voor de Lerarenopleiding Wiskunde van de Katholieke Universiteit Leuven.

Adres: K.U. Leuven, Faculteit Psychologie en Pedagogische Wetenschappen, Afdeling Didactiek, Vesaliusstraat 2, B-3000 Leuven

Abstract

The linearity illusion in secondary school students' solutions of geometry problems

D. De Bock, L. Verschaffel & D. Janssens. *Pedagogische Studiën*, 1997, 74, 261-270.

Linear (proportional) functions are undoubtedly one of the most common models for representing and solving both pure and applied problems in elementary mathematics education. But according to several authors, several aspects of the current culture and practice of school mathematics develop in students a tendency to use these linear models also in situations wherein they are not applicable. This article reports two ascertaining studies about the occurrence of this 'illusion of linearity' in 12-13- and 15-16-years-old students' working on word problems involving lengths and areas of similar plane figures of different kinds of shapes, as well as about the influence of drawings in breaking the linearity illusion for that kind of geometrical problems. Generally speaking, the results provide a convincing demonstration of the omnipresence and the strength of the illusion of linearity among these students.

Systemtheoretische benaderingen bij de behandeling van klachten van jongeren

A. Lange en J.-P. van de Ven

Samenvatting

Dit artikel gaat in op de vraag welke systeemtheoretische concepten het meest hebben bijgedragen aan de behandeling van stoornissen van jongeren, met name depressie, angst en gedragsstoornissen. Om deze vraag te beantwoorden zijn eerst de belangrijkste systeemtheoretische stromingen op een rij gezet. Vervolgens wordt een overzicht gegeven van het onderzoek naar de samenhang tussen gezinsvariabelen en de genoemde stoornissen bij jongeren. Dit wordt gevolgd door een overzicht van het onderzoek naar de effecten van gezinsbehandelingen voor deze stoornissen. Uit dit onderzoek kan worden geconcludeerd dat vooral het ingrijpen in de rechtstreekse interactie tussen ouders en kind, de manier waarop ouders opvoeden, van belang lijkt te zijn voor de ontwikkeling en het verloop van deze stoornissen. Van de systeemtheoretische stromingen zijn met name de concepten uit de leertheoretische, de structurele, de communicatieve en de psycho-educatieve gezinstherapie van belang, en in iets mindere mate die van de intergenerationale en cybernetische gezinstherapie. Tot slot wordt ingegaan op de discrepantie tussen klinische bevindingen en de gegevens uit gecontroleerd onderzoek, en wordt een pleidooi gedaan voor systematisch onderzoek naar de effectiviteit van multi-dimensionele behandelingspakketten.

1 Inleiding

Vanaf het begin van deze eeuw zijn er publicaties verschenen waarin een verband wordt gelegd tussen psychiatrische stoornissen en de manier waarop leden van het gezin met elkaar omgaan (bijv. Flügel, 1921). In de jaren vijftig nam de op deze notie gebaseerde gezinstherapie een grote vlucht. In deze periode werd de

gezinstherapie ook steeds meer toegepast wanneer het ging om problemen met kinderen. De zogenaamde 'Palo Alto'-groep kwam in die tijd met haar concept van het 'schizofrene gezin': het gezin dat door haar manier van met elkaar omgaan (met name de communicatie tussen moeder en kinderen) een grote kans liep om schizofrene kinderen voort te brengen. De behandeling moest op die communicatie worden gericht (Bateson, Jackson, Haley & Weakland, 1956; 1963). Tot in de jaren tachtig hebben belangrijke systeemtheoretische auteurs als Haley (1989) vastgehouden aan dit concept. Minder controversieel en meer praktisch was in die tijd het werk van Minuchin, de structurele gezinstherapeut. In de klassieker *The families of the slums* (Minuchin, Montalvo, Guerney, Rosman & Schumer, 1967) beschrijven hij en zijn medewerkers hoe men jonge delinquenten uit een kansarm gezin het beste met behulp van dat gezin kan behandelen. De ideeën in dat boek staan, zoals we zullen zien, nog ten dele overeind.

In dit artikel wordt ingegaan op de vraag in hoeverre onderzoek de laatste decennia de meerwaarde van gezinstherapie bij de behandeling van kinderen heeft bevestigd. Alvorens ons op deze vraag te richten wordt eerst een korte karakteristiek gegeven van de belangrijkste gezinstherapeutische stromingen. Daarna geven we weer wat er uit onderzoek bekend is over de relatie tussen gezinsinteracties en stoornissen bij kinderen, waarbij we ons beperken tot depressie, angst en gedragsstoornissen. Vervolgens bespreken we het onderzoek naar de effecten van gezinsbehandeling. Hierbij nemen we enkele korte gevalbeschrijvingen op als illustratie en komen we tenslotte toe aan de vraag wat de hierboven genoemde stromingen betekenen voor het behandelen van stoornissen bij kinderen.

2 De gezinstherapeutische stromingen

2.1 Transgenerationale gezinstherapie

Interacties tussen gezinsleden worden volgens de visie van deze stroming gedomineerd door projecties uit het verleden. De relaties in het huidige gezin worden beïnvloed door de relaties die de ouders hadden in hun gezin van herkomst. Het begrip 'loyaliteit' speelt een belangrijke rol. Niet goed verlopende processen van losmaking ten opzichte van het gezin van herkomst hebben negatieve gevolgen voor de relaties in het huidige gezin. Onbewuste loyaliteiten ten opzichte van ouders kunnen intimiteit en een volwassen relatie met een partner in de weg staan. Het intergenerationale gezinstherapeutisch model is er dan ook vooral op gericht deze onbewuste loyaliteiten op te sporen, en het losmakingsproces alsnog op positieve wijze af te ronden, of het mislukken daarvan te verwerken (Boszormenyi-Nagy, Grunebaum & Ulrich, 1991; Goldenthal, 1991).

2.2 De structurele systeembenadering

De vertegenwoordigers van deze stroming gaan ervan uit dat er pathogene gezinsstructuren zijn. Er wordt veel waarde gehecht aan de aard van de coalities binnen het gezin en de grenzen tussen de subsystemen. Ouders behoren samen een hecht, goed afgegrensd, subsysteem te vormen en de hiërarchie in het gezin te bewaken. Coalities tussen een ouder en één of meer kinderen, waarvan de andere ouder wordt buitengesloten, worden beschouwd als een potentiële bron van psychische stoornissen (Colapinto, 1991; Haley, 1980; Minuchin, 1974; Minuchin & Fishman, 1981).

2.3 De communicatie-theoretische systeembenadering

Het accent in deze stroming ligt op het blootleggen van wetmatigheden in de manier waarop gezinsleden met elkaar omgaan. In duurzame systemen speelt, al of niet bewust, het willen uitoefenen van invloed een grote rol. Conflicten, die schijnbaar een inhoudelijk meningsverschil betreffen, blijken vaak gebaseerd op de behoefte een stempel te drukken op de onderlinge verhouding (Haley, 1963; Lederer & Jackson, 1968; Segal, 1991; Watzlawick,

Beavin & Jackson, 1967). Het analyseren en doorbreken van dit soort processen met behulp van gerichte opdrachten, wordt binnen deze stroming als de belangrijkste taak van de systeemtherapeut gezien.

2.4 De experiëntiële systeembenadering

In deze stroming gaat men ervan uit dat een tekort aan 'awareness' (het dichtbij de eigen gevoelens staan) de belangrijkste bron is van psychische stoornissen. Om die 'awareness' te vergroten, worden in gezinssessies concepten en technieken uit de individuele gestalttherapie toegepast. Stevige confrontaties worden hierbij niet geschuwd. Anders dan bij de eerder genoemde stromingen, zijn er bij deze benadering geen duidelijke systeemtheoretische concepten (Keith, Connell & Whitaker, 1991; Satir, 1964).

2.5 De leertheoretische en cognitieve systeembenadering

Aanvankelijk zag men binnen deze stroming psychiatrische symptomen voornamelijk als een uitvloeisel van het systematisch bekrachtigen van ongewenst gedrag binnen het gezin of in de relatie tussen partners. De behandeling werd dan ook vooral gericht op het opsporen van dergelijke bekrachtigingspatronen, en op het aanbrengen van veranderingen in deze patronen (Holtzworth-Munroe & Jacobson, 1991; Stuart, 1980). De gedragstherapeutische gezinstherapie is de laatste jaren veelzijdiger geworden. Onder invloed van cognitieve theorieën is men zich ook gaan richten op het beïnvloeden van irrationele gedachten die de interacties tussen gezinsleden schade berokkenen (Dattilio, 1990; Epstein, Baucom & Rankin, 1993; Munson, 1994; Schmaling, Fruzzetti & Jacobson, 1989).

2.6 De cybernetische systeembenadering

Deze benadering gaat terug naar de basis van de systeemtheorie. Het 'orthodoxe' uitgangspunt, dat psychische stoornissen meestal een homeostatische functie hebben en het systeem behoeden voor desintegratie, gaat gepaard aan een onorthodox therapiemodel. De therapeut dient het rigide homeostatische gezinssysteem te doorbreken door de symptomen van 'de patiënt' positief te etiketteren en het gezin te adviseren voorlopig niet te veranderen. Hierbij

wordt verondersteld, dat een dergelijk advies juist de krachten oproept om wél te veranderen (Boscolo, Cecchin, Hoffman & Penn, 1987; Campbell, Draper & Crutchley, 1991; Selvini Palazzoli, 1979).

2.7 De psycho-educatieve gezinstherapie

Hierbij gaat men er van uit dat diverse stoornissen hun grondslag hebben in een neurologische kwetsbaarheid, die zo mogelijk met medicatie behandeld kan worden. Ouders van kinderen dienen over de ziekte, het verloop en de medische behandeling te worden voorgelicht. Zij behoren niet opgezadeld te worden met de gedachte dat zij schuld hebben aan de depressie, de angsten, de gedragsproblemen van hun kind. De rol van het gezin wordt echter niet over het hoofd gezien. Het gezin kan bijdragen aan de noodzakelijke rust, het vermijden van conflicten en negatieve emoties. Deze stroming heeft zich in eerste instantie ontwikkeld bij de behandeling en preventie van terugval bij schizofrene patiënten (Anderson, Hogarty & Reiss, 1980; Falloon, Boyd & McGill, 1984) en is later ook toegepast bij andere stoornissen.

3 Gezinsinteractie en stoornissen bij kinderen en adolescenten

3.1 Angst en depressie

Retrospectief onderzoek bij volwassen angstige patiënten geeft aanwijzingen dat hun gezin van herkomst onderhevig was aan dysfunctionele interactiepatronen: overprotectie, ambivalentie en negatieve houding van de kant van de ouders (Gerlsma, Emmelkamp & Arrindell, 1990). Aan retrospectief onderzoek kleven uiteraard methodologische bezwaren; de herinneringen kunnen systematisch gekleurd zijn. Kashani et al. (1990) onderzochten derhalve de huidige interactiepatronen in gezinnen met angstige kinderen. De moeders van de meest angstige kinderen scoorden het negatiefst op de door hen gebruikte gezinsrelatiemaat, wat volgens de auteurs wijst op een relatief slechte ouderlijke relatie en weinig harmonie bij de ouders van sterk angstige kinderen in vergelijking met het gezinsklimaat bij minder angstige en normale kinderen. In een studie van Stark, Humphrey, Crook en Lewis (1990) bleek ech-

ter het omgekeerde: moeders van kinderen met een angststoornis zagen het gezinsklimaat positiever dan moeders uit de controlegroep. De angstige kinderen zelf wezen wel op een negatiever gezinsklimaat. Ook Messer en Beidel (1994) vonden bij de ouders geen verschillen in gezinsklimaat tussen gezinnen met angstige en niet-angstige kinderen, en ook hier waren er aanwijzingen dat de angstige kinderen zelf wel problemen hadden met de sfeer in het gezin. Op grond van een analyse van de gegevens van de vaders, moeders en de kinderen gezamenlijk komen de auteurs tot de voorzichtige hypothese dat de sfeer in het gezin slechts een indirecte invloed uitoefent op het angstige kind. Psychopathologie van de ouders zou de werkelijke factor zijn waar het om gaat. We komen later nog op dit aspect terug.

Ginsburg, Silverman en Kurtines (1995) vatten de gegevens van een door hen uitgevoerd literatuuronderzoek op dit terrein samen. Zij komen tot de conclusie dat er een forse kans is om in gezinnen van angstige kinderen onaanpaste gezinsrelaties aan te treffen, met veel conflicten, overmatige controle, weinig support en cohesie en weinig gemeenschappelijke activiteiten.

Er zijn diverse studies die aanwijzingen geven dat er een verband is tussen psychopathologie van de ouders, met name de moeder en unipolaire depressie van kinderen (Beardslee, Keller & Klerman, 1985)). Volgens Gelfand en Teti (1990) en Keller et al. (1986) wordt dit verband behalve door een genetische factor ook verklaard door de geringe opvoedkundige vaardigheid van de moeder en het geringe vermogen om een positief affectieve relatie aan te gaan. Er is flink wat evidentie dat depressieve kinderen het gezinsklimaat als negatiever (minder cohesief en minder steunend, meer dwingend en meer conflictueus) ervaren dan normale kinderen (Kaslow, Deering & Racusin, 1994). De bevindingen inzake correlaties tussen gezinsmetingen en depressie van kinderen zijn echter niet eenduidig. Feldman, Rubenstein & Rubin (1988) en Garrison et al. (1990) vonden geen verband. Reinherz et al. (1989) vonden wel een verband tussen een cohesieve gezinsstructuur en bescherming van adolescenten tegen depressie na stressfull life events. Cole en McPherson (1993) preciseren dit: het blijkt niet zozeer te gaan om de relatie

tussen de ouders en ook niet om de band tussen moeder en de adolescent, maar meest bepalend is de mate van cohesie tussen vader en de adolescent.

Er zijn ook onderzoekers die aanwijzingen vinden voor effecten van systemische variabelen (variabelen die betrekking hebben op de interactie tussen leden van het sociale systeem, met name het gezin). Zo beschrijft Coyne (1984) een interpersoonlijk depressiemodel met cognitieve factoren. Hierin gaat hij er van uit dat depressieve personen met hun symptomen schuldgevoelens oproepen bij andere leden van het systeem, juist doordat zij hun negatieve klachten niet rechtstreeks uiten. Uitgaande van dit model beschrijft Kaslow et al. (1994) haar observatie dat depressieve kinderen vaak hulpeloos acteren om zich van steun en geruststelling te verzekeren. Dit leidt bij sommige ouders tot negatief vijandig gedrag, en bij andere ouders kan het een gevoel van onnipotentie geven, dat zij in stand houden door hun kind op subtiele wijze te bekrachtigen in de negatieve zelfopvattingen. De empirische steun voor deze theorie is echter gering.

Meer steun is er voor de rol van EE (Expressed Emotion). Hoge niveaus van EE blijken net als bij schizofrenie een rol te spelen bij depressie bij kinderen. Depressieve kinderen uit een gezin waar kritiek, vijandigheid en overbetrokkenheid een grote rol spelen, vallen na behandeling en terugkeer in het gezin vaker terug dan kinderen die uit lage EE gezinnen voortkomen (Asarnow, Goldstein, Tompson & Guthrie, 1993).

Op grond van literatuuroverzichten concluderen zowel Kaslow et al. (1994) als Chiariello en Orvaschel (1995) dat gezinsproblemen dan wel geen causale oorzaak zijn van depressie, maar wel degelijk de depressie kunnen 'triggere' bij kinderen die biologisch (genetisch of door 'opgelopen' defecten) kwetsbaar zijn.

3.2 Gedragsstoornissen

Het *Journal of Consulting and Clinical Psychology* heeft in 1991 en 1992 een aantal artikelen opgenomen dat representatief is voor de discussie rond de rol die gezinsvariabelen spelen in de ontwikkeling en handhaving van gedragsstoornissen bij kinderen. De discussie start met een overzichtsartikel van Fauber en Long (1991). Zij bespreken een aantal studies

(o.a. Fauber, Forehand, Thomas & Wierson, 1990) waarin met behulp van multiële regressie-analyses is onderzocht welke variabelen de grootste invloed hebben op het gedrag van gedragsgestoorde kinderen. Uit deze studies blijkt dat het ouderlijk functioneren en het gezinsklimaat geen rechtstreekse voorspellende waarde hebben voor het functioneren van het kind, maar dat dit verloopt via de mediërende factor 'het opvoedkundige gedrag'. Op grond daarvan pleiten de auteurs ervoor niet automatisch alle voorhanden liggende gezinsproblemen in een behandeling te betrekken, maar te onderzoeken op welke wijze (bijv. via het gedrag ten opzichte van het kind) deze van invloed zijn, en de behandeling daarop te richten.

Emery, Fincham en Cummings (1992) zijn het niet eens met de conclusies van Fauber en Long (1991). De statistische analyses zouden de lineaire conclusies van Fauber en Long niet wettigen. Zij baseren zich op experimenteel onderzoek van o.a. Gottman en Katz (1989) en van O'Brien, Margolin, John en Krueger (1991) waaruit blijkt dat gemanipuleerde gezinsconflicten tot cognitieve, affectieve en psychofysiologische reacties bij kinderen leiden zonder dat er sprake was van rechtstreekse interacties tussen een ouder en het kind. Zij pleiten dan ook voor een multidimensionele benadering. Dit houdt in dat zowel het gezinsfunctioneren van de ouders als hun ouderlijk gedrag onderzocht en behandeld worden. Niettemin worden de conclusies van Fauber en Long (1991) ook door anderen bevestigd. Zo hebben Frick, Lahey, Hartdagen en Hynd (1989) en Lahey et al. (1988) aangetoond dat huwelijksproblemen van ouders als factor bij de gedragsstoornis van kinderen verwaarloosbaar worden wanneer er in de multivariate analyses rekening wordt gehouden met het antisociale gedrag van de ouders. Laub en Sampson (1988) vonden zelfs dat het verband tussen criminaliteit van ouders en van de criminaliteit van hun kinderen grotendeels verklaard kon worden uit een ineffektieve hardvochtige opvoeding. Frick et al. (1992) concluderen op grond van een eigen vergelijkend onderzoek bij gezinnen van gedragsgestoorde kinderen en op grond van een openvolging van vergelijkbare studies dat kinderen met gedragsproblemen opvallend vaak ouders hebben met antisociale

persoonlijkheden, die negatieve, agressieve en inconsistente opvoedingsstijlen hanteren.

4 Gezinsinterventies bij de behandeling van stoornissen bij kinderen

4.1 Depressie en angststoornissen

Er is niet veel gecontroleerd onderzoek naar de effecten van gezinsbehandeling bij depressies van kinderen. In een aantal meta-analyses en reviews die de laatste jaren zijn uitgevoerd over de effecten van gezinstherapie komen dergelijke studies niet voor (Markus, Lange & Pettigrew, 1990; Pinosof & Wynne, 1995; Shadish, Ragsdale, Glaser & Montgomery, 1995). Diamond, Serrano, Dickey en Sonis (1996) noemen maar één onderzoek (van Brent, Holder & Kolko, 1993) waarin de gezinsverhoudingen het onderwerp waren van interventie en enkele studies waarin het gaat om Parent Management Training. In een literatuuronderzoek naar gezins- en relationele behandeling van depressie bij volwassenen vinden Prince en Jacobson (1995) weinig aanwijzingen dat relatie- of gezinstherapie bij stemmingsstoornissen superieur is aan individuele cognitieve behandeling. De bevindingen zijn echter moeilijk te interpreteren aangezien de meeste studies gaan over relatietherapie waarbij niet duidelijk is in hoeverre er ook aandacht is geweest voor de individueel pathologische kant van de depressie; en dat blijkt juist van belang te zijn, zoals consistent naar voren komt in de 'reader' van Clarkin, Haas en Glick (1988). Hierbij gaat het vaak om farmacologische interventies (Epstein, Keitner, Bishop & Miller, 1988) en om cognitieve interventies met de nadruk op het uitdagen van negatieve verwachtingen (Dobson, Jacobson & Victor, 1988). Dergelijke individuele benaderingen zijn zeer wel te combineren met gezinstherapie.

Gebaseerd op wat zij noemen 'generations of outcome studies' breken Epstein et al. (1988, p. 159) een lans voor 'problem-centered systems therapy of the family' voor wat betreft de combinatie tussen individuele en gezinsinterventies. Dit therapiemodel voorziet in het algemeen in het nauwkeurig in kaart brengen

van de mate van probleem-oplossend vermogen, de aard van de affectieve responsiviteit en gedragsbeïnvloeding binnen het gezin ten aanzien van verschillende rollen in verschillende levensstadia. Aan de behandeling gaan een uitgebreide taxatie en schriftelijke contracten met de betrokkenen vooraf. Dan verloopt de behandeling volgens vaste stappen: prioriteiten vastleggen, doelen stellen, taken formuleren en evalueren. Wanneer er gezinsleden (ook adolescenten en kinderen) met depressie in het gezin zijn, dan dient de aard en de zwaarte van de depressie te worden vastgesteld en zo nodig eerst of parallel met medicatie te worden behandeld.

Epstein et al. (1988) beschrijven als voorbeeld een 17-jarige scholier die opgenomen was na een episode van een half jaar met depressie, veel alcoholgebruik, drugsgebruik, verbale agressiviteit in het gezin en tenslotte een suïcidepoging. Tot een jaar voorafgaande aan de opname had het gezin goed gefunctioneerd. Toen kreeg de vader een ernstig ongeluk, moest ophouden met werken en kwam als invalide thuis te zitten. Bovendien verloor de jongen in hetzelfde jaar zijn grootvader op wie hij zeer gesteld was geweest. Vader werd depressief en trok zich terug uit de warme, goede relatie met de jongeman. In deze behandeling hoefde geen medicatie te worden verstrekt. Met behulp van gezins- en individuele gesprekken werden bovenstaande verliezen besproken en werden de sterke kanten in het gezinsfunctioneren weer hersteld. Dit leidde tot dramatische verbetering in stemming van de patiënt die vrij snel kon worden ontslagen. Er volgden enkele follow-up gesprekken waaruit kon worden opgemaakt dat het thuis ook goed ging. Helaas verliepen de gebeurtenissen daarna minder voorspoedig: de gezinsgesprekken met de oorspronkelijke behandelaars moesten al vrij snel worden gestopt (het gezin woonde te ver weg) en er was geen vergelijkbare follow-up behandeling in hun woonplaats (al hadden ze zich wel aangemeld bij een instelling voor ambulante geestelijke gezondheidszorg, vergelijkbaar met een Nederlandse Riagg). Zes maanden later schoot de jongen zich in de klas dood. De auteurs concluderen dat het gezin, door gebrek aan continuïteit in de behandeling, in haar depressieve patroon was teruggevallen.

Een ander voorbeeld uit dezelfde studie betreft een 16-jarige zwaar depressieve jongen, die opgenomen was na een aantal suicide-pogingen. Zijn vader was overleden toen hij drie jaar was. Zijn moeder hertrouwde enkele jaren later en er ontstond een warme en positieve relatie met zijn stiefvader. Een jaar voorafgaande aan de opname stierf deze man. Sinds die tijd woonde hij met zijn moeder alleen. Zijn moeder was een sterke vrouw, die vooral uitblonk in haar verzorgende rol. Andere aspecten van ouderlijk gedrag waren niet sterk bij haar vertegenwoordigd. Toen haar tweede man ziek werd stortte zij zich op diens verzorging en voelde de jongen zich van haar vervreemd. Hij werd geïsoleerd en eenzaam. In het begin van de behandeling kreeg de jongen anti-depressiva. Toen deze eenmaal aansloegen werd het mogelijk om een gezinsbehandeling te richten op de problemen tussen moeder en zoon. De jongen floreerde in deze behandeling. Hij bleek flexibeler dan zijn moeder en was sneller dan zij in staat tot verandering in de onderlinge verhouding. Maar ook zij reageerde uiteindelijk goed op dit aspect van de behandeling. Een jaar na ontslag functioneerde de jongen goed en waren de veranderingen in de relatie tussen moeder en zoon in stand gebleven.

Over gezinsbehandeling bij kinderen met angststoornissen zijn nog minder onderzoeksgegevens dan ten aanzien van depressie. Diamond et al. (1996) noemen in hun overzicht maar één vergelijkende studie (Barrett, Dadds, Rapee & Ryan, 1993), waarin individuele cognitieve therapie wordt gecombineerd met een gezinsinterventie. Bij 88% van de kinderen leidt dit ertoe dat zij niet meer voldoen aan de diagnostische criteria voor angststoornis. Individuele cognitieve therapie zonder gezinsinterventie heeft een lager succespercentage (61%). Uit de beschrijving wordt de aard van de gezinsinterventie echter niet duidelijk.

Een literatuuronderzoek van Estrada en Pinsof (1995) geeft ook een magere oogst. De onderzoeken die gevonden werden, hadden betrekking op angst in het donker, simpele fobieën en schoolfobie. In alle studies ging het om 'parents aided', dus door ouders ondersteunde behandeling. De focus lag op rechtreekse behandeling van de angst van het kind, met behulp van verandering van operante ketens, cognitieve therapie en coping strate-

gieën, waarbij de ouders waren betrokken. De relaties tussen de ouders onderling en gezinsrelaties stonden in de gerapporteerde studies niet ter discussie. Alleen Ginsburg et al. (1995) beschrijven een project waarin ouders van angstige kinderen zelf het doelwit van de behandeling kunnen zijn, daar waar zijzelf (wat vaak het geval is) last hebben van angsten. Ginsburg et al. (1995) maken hierbij gebruik van protocollen waarin bijvoorbeeld het angstige kind en de angstige ouder elkaar helpen bij het uitvoeren van exposure-taken (waarbij men zich blootstelt aan een angstige situatie).

Dat het betrekken van de ouders bij de behandeling toch wel een extra bijdrage kan leveren zien we geïllustreerd in een gevalbeschrijving van Lange (1994, hfdst. 3). Hierin wordt een 11-jarig meisje beschreven dat al meer dan een jaar leed aan een schoolfobie. Sinds een maand ging ze helemaal niet meer naar school. Elke ochtend had ze last van buikpijn, wat haar belette om op te staan. Grondig medisch onderzoek bracht geen lichamelijke oorzaak aan het licht. Als haar ouders druk uitoefenden om toch op te staan en naar school te gaan werd de pijn alleen maar erger. Om een uur of elf begon ze zich meestal beter te voelen. Ze stond dan op, hielp haar moeder een beetje in het huishouden en ging met haar boodschappen doen. 's Avonds, voor het naar bed gaan, begonnen haar lichamelijke klachten terug te komen. Het meisje had twee oudere zusters die geen problemen vertoonden. Haar vader was een hard werkende vertegenwoordiger die meestal pas laat in de avond thuis kwam. Haar moeder was huisvrouw. Aangezien zij geen werk buitenshuis had, besteedde zij al haar tijd en aandacht aan het huis en de kinderen.

Het meisje was intelligent genoeg, maar had last van een combinatie van een extreem hoog aspiratieniveau en faalangst. De problemen waren begonnen toen zij naar een nieuwe klas was gegaan met kinderen die ouder waren dan zij en een nieuwe leraar die zij niet aardig vond. Hoewel haar ouders niet veel gemeenschappelijks leken te hebben waren er geen zichtbare huwelijksconflicten. Haar moeder was betrekkelijk eenzaam, maar dit leidde niet tot conflicten met haar man. Het had waarschijnlijk wel invloed op de verhouding tussen dochter en moeder. Deze verhouding was nogal symbiotisch. Door thuis te blijven leverde het meisje

aan moeder een 'bestaansrecht' en moeder bekrachtigde het niet naar school gaan door allerlei dingen met haar dochter te gaan ondernemen.

De therapeute besprak met het gezin de vicieuze cirkel van bekrachtiging waarin de dochter was terecht gekomen. Er werd afgesproken dat moeder ervoor ging zorgen dat de leraar dagelijks zoveel huiswerk bij het meisje zou laten bezorgen dat ze de klas weer zou kunnen inhalen. Ze zou zelfs vooruit werken. Dit zou haar terugkomst in de klas kunnen faciliteren. Als zij niet naar school ging, zou zij op haar kamer blijven en huiswerk maken gedurende de uren dat zij anders op school zou zijn. Het was haar verboden om moeder te helpen, mee te gaan boodschappen doen, of te spelen.

Binnen twee weken ging het meisje weer naar school. De crisis was over, maar er waren onderliggende structuren waaraan iets moest worden gedaan. De dochter vertoonde diverse symptomen van sociale angst, die tot gevolg zouden kunnen hebben dat de schoolfobie opnieuw zou toeslaan. Dit werd behandeld met behulp van individuele cognitieve gedrags therapie, die uit twee componenten bestond. Enerzijds stond het uitdagen van dysfunctionele gedachten centraal: door middel van het stellen van vragen over voor het meisje vanzelfsprekend geworden gedachtengangen en door het blootleggen van strijdigheden in de onderliggende redeneringen werd zij aan het twijfelen gebracht over contra-productieve cognities (Arntz, 1991). Aan de andere kant werd volop aandacht besteed aan het vergroten van haar sociale vaardigheden (door middel van rollenspelen). De patiënte had plezier in deze training en veranderde aanmerkelijk.

Nadat de behandeling al een zekere mate van succes had gehad, werd alsnog geïnterveerd op het niveau van het gezinssysteem. Moeder was immers nogal eenzaam. Stilletjes had dit patroon ertoe bijgedragen dat het moeder wel goed uitkwam als haar dochter thuis bleef. Dit aan de kaak te stellen leek echter niet de juiste aanpak. De ouders waren niet voor dit soort confrontaties gekomen. In plaats daarvan besprak de therapeut met de ouders de levensfasen waarin zij (speciaal moeder) nu verkeerden. Aangezien de kinderen overdag niet meer thuis waren zou het misschien goed zijn als ook moeder enige bezigheden buitenshuis zou creë-

ren. Moeder was het hiermee eens en werd daarin gesteund door vader.

In een follow-up sessie een half jaar later bleken de veranderingen meer dan stand te hebben gehouden; moeder was lid geworden van een fitness club en deed vrijwilligerswerk. Het beviel haar uitstekend. Bovendien had haar toegenomen onafhankelijkheid haar kennelijk aantrekkelijker voor haar echtgenoot gemaakt. Hij was vaker thuis en ze gingen regelmatig samen uit. Hun dochter vertoonde geen terugval.

4.2 Gedragsstoornissen

In tegenstelling tot bij depressie en angst is er flink wat gecontroleerd onderzoek naar de effectiviteit van gezinsbehandelingen van gedragsstoornissen bij kinderen en adolescenten. Zo doen Henggeler, Melton en Smith (1992) verslag van een studie bij 84 jeugdige delinquenten waarbij een 'multi-systemic therapy' groep (MST) vergeleken werd met een controlegroep. De MST bestond uit een eclectische gezinstherapie die op maat van het kind en het gezin was gesneden, en waarbij gezins- en gedragstherapie werden geïntegreerd. Diamond et al. (1996) brengen enkele studies vanuit de 'Functional Family Therapy' (FFT) van Alexander (1988) in herinnering. Deze FFT kan beschouwd worden als een op cybernetische leest geschoeide eclectische gezinsbehandeling, waarin getracht wordt het tot zondebok maken van de geïdentificeerde patiënt te voorkomen in combinatie met Parent Management Training. Deze behandeling en de eerder beschreven MST van Henggeler et al. (1992) doen denken aan het hierboven door Epstein et al. (1988) beschreven model van problem-centered therapie, waarbij in dit geval extra nadruk leek te liggen op 'parenting management', de manier waarop ouders hun opvoedkundige taken waarmaken. De resultaten waren positief, met significante effecten met betrekking tot crimineel gedrag, en ook met betrekking tot gezinsvariabelen als 'cohesie binnen het gezin'. Het is jammer dat in dit onderzoek niet is nagegaan in hoeverre de gezinsvariabelen een mediërende rol spelen bij het voorkomen van nieuwe criminaliteit.

Sayger, Horne en Glaser (1993) voerden bij een populatie van niet al te zwaar gedragsgestoorde kinderen een uitgebreide Parent Management Training (PMT) uit die was

gebaseerd op de sociale leerprincipes van Patterson, Reid, Jones en Conger (1975). De gedragsproblemen thuis verminderden, bovendien zetten de veranderingen zich door in de klas. Opvallend was dat bij de groep van gezinnen waar de ouders het samen niet zo best hadden, de PMT ertoe leidde dat de tevredenheid van de ouders met hun onderlinge relatie significant toenam. Ook de gevoelens van cohesie in het gezin namen toe. Deze bevindingen suggereren dat de causaliteit bij gedragsstoornissen niet begint bij de ouderlijke relatie, maar dat – omgekeerd – slecht opvoedingsgedrag van de ouders leidt tot moeilijkheden en spanningen in het gezin. Dit komt overeen met bevindingen die we in de sectie over gezinsinteractie hebben besproken.

Estrada en Pinsop (1995) voerden een literatuuronderzoek uit. Op grond hiervan komen zij tot de volgende conclusies: van PMT is weliswaar aangetoond dat het positieve effecten heeft voor kinderen met gedragsstoornissen, maar er zijn ook flink wat aanwijzingen dat aanvullende behandelingscomponenten van belang zijn, zoals training van zelf-controle van ouders en het veranderen van de cognities van de ouders met betrekking tot het betreffende kind. Verder vonden zij studies die erop wezen dat als de relatie tussen de ouders niet goed is, het aanpakken daarvan de kans op een positieve verandering bij het kind vergroot. Er zijn geen studies gevonden die traditionele gezinstherapie sec (waarbij de aandacht alleen ligt op de structuur en de gezinsinteracties) vergelijkt met behandelingen waarin gezinsinterventies worden gecombineerd met PMT.

Bovenstaande wordt geïllustreerd door een momenteel 'lopende' behandeling in de eigen praktijk. Het betreft een jongen van 12 jaar. Bij diens aanmelding vertelt de moeder dat er ook gezinsproblemen zijn, met name een slechte relatie tussen de ouders en tussen de kinderen onderling. Er volgt een gezinsintake en een gesprek met de ouders. De slechte verhouding tussen de ouders blijkt er inderdaad te zijn. Moeder is een emotionele, ietwat chaotische vrouw; vader een wat sombere, enigszins cynische man. De verhouding met zijn vrouw vindt hij hopeloos. Ze maakt overal toestanden over. In tegenstelling tot zijn vrouw is hij niet bereid om veel aan hun onderlinge problemen te doen. Hij wil wel bij de behandeling betrokken blij-

ven om te zien wat ze aan hun zoon kunnen doen. Het behandelingscontract wordt derhalve daarop gericht: de ouders leren om anders met hun zoon (eventueel ook met hun tienjarige dochter die minder problemen geeft) om te gaan: Parent Management Training derhalve. Alvorens tot gezinszittingen over te gaan vraagt de therapeut elk der ouders om bij wijze van huiswerk een aantal concrete doelen op te stellen voor verandering in het gedrag van de jongen, en daarbij aan te geven in welke volgorde zij die doelen hopen te bereiken. In de eerste gezinszitting wordt het eerste doel geselecteerd: stoppen met voortdurend zingen en lawaai maken (de jongen blijkt in huis, in welke ruimte dan ook een voor de anderen enorme hoeveelheid storend lawaai te maken, niet om te pesten maar omdat hij dat leuk vindt). De therapeut helpt de ouders om met hun zoon te onderhandelen. Er komt een precieze afspraak uit de bus: de ouders zullen de momenten registreren waarop hun zoon het 'target' gedrag vertoont op een manier die overlast veroorzaakt. Zij zullen hem dan eenmaal vragen om op te houden. Als hij die waarschuwing opvolgt levert dit hem 50 cent op met een maximum van f 5,- per week. Elke dag waarop geen waarschuwing nodig is levert hem een gulden op. Het principe dat belonen beter is dan straffen blijkt goed te werken. Het 'target' gedrag verdwijnt. Iedereen, inclusief de zoon is tevreden. De afspraak wordt verlengd, en na enige tijd wordt de beloning geleidelijk teruggeschroefd. Inmiddels blijkt een tweede doelstelling belangrijk te zijn: pesten en uitschelden van zijn jongere zusje. De therapeut helpt zijn ouders om met hem en zijn zusje hierover te onderhandelen. Hier komt eveneens een afspraak uit de bus, die hem bij positief gedrag een beloning oplevert. Ook zijn zusje kan een beloning verdienen als zij nalaat hem te treiteren. Ook deze afspraken worden goed en naar tevredenheid nagekomen. Het gezin is verbaasd dat dergelijke op het oog simpele afspraken zoveel veranderingen bewerkstelligen. De behandeling loopt nog, maar het laat zich aanzien dat de veranderingen standhouden, temeer daar de ouders en de kinderen er zich zeer door aangesproken voelen. Net als in de hierboven beschreven studie van Sayger et al. (1993) blijkt het samen bezig zijn met positiever en effectiever opvoedingsgedrag dan voorheen

een gunstige invloed op de onderlinge verhouding van de ouders te hebben.

Er zijn minder studies over het effect van gezinsbehandeling bij ADHD dan bij gedragsstoornissen. De literatuur samenvattend komen Estrada en Pinsof (1995) tot de conclusie dat ook hier PMT het meest effectief lijkt, maar dat de effecten vooral betrekking hebben op de gedragsproblemen (ongehoorzaamheid en agressie) van de ADHD-kinderen. Met betrekking tot de typische ADHD-kant (de impulsiviteit, de concentratieproblemen en overactiviteit) is van PMT en andere gezinsinterventies sèc nog geen overtuigend effect aangetoond (Diamond et al., 1996). Wel zijn er duidelijke aanwijzingen dat intensieve multidimensionele gezinstherapie in combinatie met psychofarmaca (bijv. methylfenidaat) positieve effecten oplevert: beter functioneren thuis en op school, en, wat betreft delinquente jongeren, minder arrestaties (Satterfield, Satterfield & Schell, 1987).

5 Conclusies

Het meeste onderzoek naar de interacties in gezinnen bevat aanwijzingen dat het niet de interactiepatronen of structurele variabelen in het gezin zijn die vooral samenhangen met angst, depressie en gedragsstoornissen bij kinderen. Van groter belang lijken de gezinsvariabelen die ingrijpen in de rechtstreekse interactie tussen de ouders en hun kinderen: psychopathologie van de ouder, een al te kritische houding, inconsistent en agressief opvoeden en (bij angstige kinderen) ook het bekrachtigen van de angsten. In het schaarse Nederlandse onderzoek vinden we aanwijzingen dat dit ook geldt voor somatiserende kinderen. Osterhaus, Lange en Richard (in voorbereiding) vonden dat onverklaarbare somatische klachten van kinderen vooral samenhangen met het niet kunnen uiten van gevoelens in het gezin en met een hoog aspiratieniveau en negatief gedrag van de vaders. De moeders bleken een minder grote rol te spelen.

Opvallend is dat er in nogal wat onderzoek wordt geconcludeerd dat niet de interacties de verklarende factor zijn, maar de psychopathologie van de ouder (De Haan, 1991). Het is echter aannemelijk dat deze psychopathologie

toch van grote invloed is op de interactie van die ouder met de rest van het gezin (in casu het kind), hetgeen gevolgen heeft voor de ontwikkeling en handhaving van klachten bij het kind. Deze redenering klopt met het klassiek geworden onderzoek van Patterson (1982) waarin de interactie tussen ouders en gedragsgestoorde kinderen minutieus in kaart werd gebracht. Inconsistente patronen van aandacht en straf leidden tot toenemende opstandigheid van de kinderen, wat weer leidde tot toename van inconsistentie en negatieve aandacht door de ouders. Ook de gegevens van Frick et al. (1989) wijzen op het belang van de ouderlijke opvoedingsstijl. Een antisociale persoonlijkheidsstijl van ouders blijkt de opvoeding in negatieve zin te beïnvloeden en tot gedragsstoornissen te leiden. Het klopt ook met de bevindingen van Epstein et al. (1988) en Gelfand en Teti (1990) wat betreft depressie.

In veel van de onderzoeken worden differentiële effecten gevonden voor de verhouding tussen het kind en de vader versus de verhouding met moeder. Dit vraagt om een meer precieze diagnostiek dan in het algemeen gebruikelijk is. In dat verband rapporteren Lange, Blonk, Wiers en Van der Schaar (in druk) over een diagnostisch instrument (de Ouder-Kind Interactie Vragenlijst, OKIV) waarmee de opvoedkundige relatie tussen beide ouders apart met elk kind in het gezin kan worden vastgelegd. Het instrument is gebaseerd op de Family Assessment Measure van Skinner, Steinhauer en Santa-Barbara (1983). Met behulp van dit instrument toonden Blonk, Wolbert, Hansen, Heijdra en Snippert (1993) aan dat de effecten van een vaardigheidstraining voor sociaal angstige kinderen vooral werden gemedieerd door de verhouding van het betreffende kind met de moeder. In het eerder benoemde onderzoek van Cole en McPherson (1993) bleek daarentegen dat bij de ontwikkeling van depressie bij adolescenten de verhouding met de vader een speciale rol speelt. In toekomstig onderzoek naar de ontwikkeling en het beloop van stoornissen bij kinderen zal het van belang zijn de gedragsverhoudingen tussen elk der ouders en de kinderen in kaart te brengen en te toetsen op de predictieve waarde ervan. Daarbij dient ook het ouderlijk gedrag ten opzichte van de niet-gestoorde kinderen te worden betrokken.

De resultaten van het onderzoek naar behandelingseffecten zijn in lijn met de bevindingen in het onderzoek naar de gezinsinteracties. Bij alle drie de onderzochte stoornissen blijken gezinsinterventies in ieder geval van belang te zijn voor zover ze ingrijpen in het ouderlijk gedrag ten opzichte van het kind (Parent Management Training). Uit de literatuur wordt niet altijd duidelijk hoe de omgeving waarin de PMT is uitgevoerd, er uit heeft gezien. Vaak gaat het om 'mediatietherapie' waarbij het kind eenmaal is gezien en de behandeling verder alleen met de ouders wordt uitgevoerd. Soms zijn ouders en het 'target' kind, of zelfs meerdere kinderen, gezamenlijk bij de PMT betrokken. Klinisch lijken er weinig nadelen en veel voordelen aan de laatstgenoemde vorm te kleven (vgl. Lange, 1994, hfdst.13). Van gezinsinterventies *sec.*, die betrekking hebben op bijvoorbeeld de ouderlijke relatie of de algehele structuur van het gezin, is het effect minder duidelijk aangetoond.

Wat betekenen deze onderzoeksgegevens voor de verschillende systeemtheoretische scholen? De experiëntiële gezinstherapie lijkt geen bijdrage te kunnen leveren in de behandeling van stoornissen van kinderen. Ook voor de toepassing van de cybernetische gezinstherapie op zich vinden we weinig ondersteuning – wèl in combinatie met PMT, wanneer er duidelijke aanwijzingen zijn voor een symbiotische relatie tussen het kind en één of meer gezinsleden. De concepten uit de intergenerationele gezinstherapie zijn alleen van belang als een ouder geblokkeerd is in noodzakelijk opvoedingsgedrag door de manier waarop hij of zij zelf is opgevoed. De claims van de structurele gezinstherapie uit de jaren zestig zijn wat betreft de gedragsstoornissen voor een deel gerechtvaardigd gebleken. Met name het accent op de noodzakelijke 'hiërarchie' (het nemen van ouderlijke verantwoordelijkheid) is een aspect dat direct verweven is met de opvoedkundige taken die van belang zijn gebleken. De concepten uit de communicatietheoretische stroming kunnen van belang zijn als ouders er door onderlinge machtsstrijd niet goed in slagen één duidelijke lijn in de opvoeding uit te stippelen. Het ontwarren van de machtsstrijd kan hen helpen om hun opvoedkundige taken beter uit te voeren. De communicatie-concepten kunnen bovendien rechtstreeks gebruikt worden in de

opvoedkundige interactie tussen ouders en kinderen, in die gevallen waarin de conflicten met kinderen door machtsstrijd tussen ouders en kinderen geïnstigeerd en gehandhaafd blijken te worden. De nadruk op Parent Management bij vooral gedragsstoornissen vindt uiteraard een belangrijk aanknopingspunt in de leertheoretische gezinstherapie met de nadruk op exacte gedragsanalyses, dat wil zeggen het in kaart brengen en veranderen van bekrachtigingspatronen (vgl. Lange, 1994 hfdst. 13). De psycho-educatieve vorm van gezinstherapie blijkt eveneens van grote waarde. In nagenoeg al de besproken onderzoeken komt naar voren dat het van belang is om schuldgevoelens bij ouders over de stoornis van het kind weg te nemen, en om voorlichting te geven over de individuele kant van de stoornis en de manieren om er mee om te gaan.

Tot slot: Als we de onderzoeksliteratuur op ons laten inwerken dan wordt het duidelijk dat bij stoornissen bij kinderen zowel de individuele aanleg als de aangeleerde component een grote rol kan spelen en dat deze rechtstreeks, al of niet met behulp van de ouders, behandeld moet worden. Het is echter ook duidelijk dat er bij de ouders zelf, in hun relatie met het kind, of in hun onderlinge relatie, factoren *kunnen* spelen die een eenzijdig op het kind gerichte behandeling bemoeilijken, zelfs ineffectief maken. Dit vraagt om een op maat gesneden multi-dimensionele aanpak waarin zowel de op het kind gerichte interventie als de voor dat gezin specifieke nodige gezinsinterventies mogelijk zijn. Diamond et al. (1996) vraagt zich af hoe het komt dat er zoveel aanwijzingen zijn voor de invloed van gezinsrelaties op het ontstaan en handhaven van depressie en angst bij kinderen, terwijl er toch maar zeer weinig systematisch onderzoek naar is gedaan. De oorzaak ligt er waarschijnlijk in dat dergelijke behandelingsvormen moeilijker in gecontroleerde experimenten zijn te onderzoeken dan strak geprotocolleerde onderzoeken. Markus et al. (1990) stelden op grond van een meta-analyse dan ook vast, dat er geen effectstudies waren met betrekking tot een multidimensionele gezinstherapie. In de hier onderzochte onderzoeksliteratuur komen we zulke behandelingsvormen ook nauwelijks tegen. Ginsburg et al. (1995) beschrijven wel een gelaagde behande-

lingsopzet voor gezinnen met angstige kinderen, maar rapporteren geen systematisch onderzoek daarnaar. Een uitzondering vormt het onderzoek van Henggeler et al. (1992) bij delinquente kinderen, waar inderdaad een meerwaarde wordt gevonden boven behandeling in een forensische kliniek. Het lijkt van belang om in de toekomst onderzoek te verrichten naar gezins- en ouderfactoren die predictief zijn voor het al of niet slagen van individuele behandelingen bij kinderen. Vervolgens kunnen de hieruit voortkomende multi-dimensionele gezinsbehandelingen op hun effectiviteit worden onderzocht.

Literatuur

- Alexander, J.F. (1988). Phases of family therapy process: A framework for clinicians and researchers. In L.C. Wynne (Eds.), *The state of the art of family therapy research* (pp. 175-188). New York: Family Process Press.
- Anderson, C.M., Hogarty, G.E., & Reiss, D.J. (1980). Family treatment of adult schizophrenic patients: A psycho-educational approach. *Schizophrenia Bulletin*, 6, 490-505.
- Arntz, A. (1991). Principes en technieken van de cognitie therapie. *Directieve Therapie*, 11, 252-268.
- Asarnow, J.R., Goldstein, M.J., Tompson, M., & Guthrie, D. (1993). One-year outcomes of depressive disorders in child psychiatric in-patients: Evaluation of the prognostic power of a brief measure of expressed emotion. *Journal of Child Psychology and Psychiatry*, 34, 129-137.
- Barrett, P.M., Dadds, R.M., Rapee, R.M., & Ryan, S. (1993). *Cognitive-behavioral and family therapy for childhood anxiety disorders: A controlled trial*. Presentation at the annual convention of the Association for the Advancement of Behavioral Therapy, Atlanta.
- Bateson, G., Jackson, D.D., Haley, J., & Weakland, J. (1956). Toward a theory of schizophrenia. *Behavioral Sciences*, 1 (4), 251-264. Ook In D.D. Jackson (Eds.), *Communication, family and marriage. Human communication, Vol. 1*. Palo Alto, CA.: Science and Behavior Books, 1968. En in G. Bateson. *Steps to an ecology of mind*. New York: Bantam Books, 1972.
- Bateson, G., Jackson, D.D., Haley, J., & Weakland, J. (1963). A note on the double bind. *Family Process*, 2, 154-161.
- Beardslee, W.R., Keller, M.B., & Klerman, G.D. (1985). Children of parents with affective disorder. *International Journal of Family Psychiatry*, 1985, 6(3), 283-299.
- Blonk, R., Wolbert, A., Hansen, H., Heijdra, E., & Snippert, M. (1993). *Samenhang tussen ouder-kind interactie en sociale (in)competentie bij kinderen. Betrouwbaarheid en validiteit van de Ouder-Kind Interactie Vragenlijst (okiv)*. Amsterdam: Universiteit van Amsterdam, Vakgroep Klinische Psychologie.
- Boscolo, L., Cecchin, G., Hoffman, L., & Penn, P. (1987). *Milan systemic family therapy: Conversations in theory and practice*. New York: Basic Books.
- Boszormenyi-Nagy, I., Grunebaum, J., & Ulrich, D. (1991). Contextual therapy. In A.S. Gurman & D.P. Kniskern (Eds.), *Handbook of family therapy, Vol. 2* (pp. 200-238). New York: Brunner/Mazel.
- Brent, D., Holder, D., & Kolko, D. (1993). *A psychotherapy treatment for depressed adolescents: Models and treatment integrity*. Panel Presentation at the Society for Psychotherapy Research. Pittsburgh.
- Campbell, D., Draper, R., & Crutchley, E. (1991). The Milan systemic approach to family therapy. In A.S. Gurman & D.P. Kniskern (Eds.), *Handbook of family therapy, Vol. 2* (pp. 325-362). New York: Brunner/Mazel.
- Chiariello, M.A., & Orvaschel, H. (1995). Patterns of parent-child communication: relationship to depression. *Clinical Psychology Review*, 15, 395-407.
- Clarkin, J.F., Haas, G.L., & Glick, I.D. (Eds.) (1988). *Affective disorders and the family: assessment and treatment*. New York: Guilford Press.
- Colapinto, J. (1991). Structural family therapy. In A.S. Gurman & D.P. Kniskern (Eds.), *Handbook of family therapy, Vol. 2* (pp. 417-443). New York: Brunner/Mazel.
- Cole, D.A., & McPherson, A.E. (1993). Relation of family subsystems to adolescent depression: Implementing a new family assessment strategy. *Journal of Family Psychology*, 7, 119-133.
- Coyne, J.C. (1984). Strategic therapy with married depressed persons: Initial agenda, themes, and interventions. *Journal of Marital and Family Therapy*, 10, 53-62.
- Dattilio, F.M. (1990). Cognitive marital therapy: A case report. *Journal of Family Psychotherapy*, 1, 15-31.

- Diamond, G.S., Serrano, A.C., Dickey, M., & Sonis, W.A. (1996). Current status of family-based outcome and process research. *Journal American Academy of Child & Adolescent Psychiatry, 35*, 1-11
- Dobson, K.S., Jacobson, N.S., & Victor, J. (1988). Integration of cognitive therapy and behavioral marital therapy. In J.F. Clarkin, G.L. Haas & I.D. Glick, (Eds.), *Affective disorders and the family: assessment and treatment* (pp. 53-88). New York: Guilford Press.
- Emery, R.E., Fincham, F.D., & Cummings, E.M. (1992). Parenting in context: systemic thinking about parental conflict and its influence on children. *Journal of Consulting and Clinical Psychology, 60*, 909-912.
- Epstein, N., Baucom, D.H., & Rankin, L.A. (1993). Treatment of marital conflict: A cognitive-behavioral approach. *Clinical Psychology Review, 13*, 45-57.
- Epstein, N.B., Keitner, G.I., Bishop, D.S., & Miller, I.W. (1988). Combined use of pharmacological and family therapy. In Clarkin, J.F., Haas, G.L., & Glick, I.D. (Eds.), *Affective disorders and the family: assessment and treatment* (pp. 153-172). New York: Guilford Press.
- Estrada, A.U., & Pinsof, W.M. (1995). The effectiveness of family therapies for selected behavioral disorders of childhood. *Journal of Marital and Family Therapy, 21*, 403-440.
- Falloon, I.R.H., Boyd, J.L., & McGill, Ch.W. (1984). *Family care of schizophrenia*. New York: Guilford Press.
- Fauber, R.L., & Long, N. (1991). Children in context: The role of the family in child psychotherapy. *Journal of Consulting and Clinical Psychology, 59*, 813-820.
- Fauber, R.L., Forehand, R., Thomas, A.M., & Wierson, M. (1990). A mediational model of the impact of marital conflict on adolescent adjustment in intact and divorced families: The role of disrupted parenting. *Child Development, 61*, 1112-1123.
- Feldman, S.S., Rubenstein, J.L., & Rubin, C. (1988). Depressive affect and restraint in early adolescents: Relationships with family structure, family process and friendship. *Journal of Early Adolescence, 8*, 279-296.
- Flügel, J.C. (1921). *The psychoanalytic study of the family*. London: Hogarth Press.
- Frick, P.J., Lahey, B.B., Hartdagen, S.E., & Hynd, G.W. (1989). Conduct problems in boys: relations to maternal personality, marital satisfaction, and socio-economic status. *Journal of Clinical Child Psychology, 18*, 114-120.
- Frick, P.J., Lahey, B.B., Loeber, R., Stouthamer-Loeber, M., Christ, M.A.G., & Hanson, K. (1992). Familial risk factors to oppositional defiant disorder and conduct disorder: Parental psychopathology and maternal parenting. *Journal of Consulting and Clinical Psychology, 60*, 49-55.
- Garrison, C.Z., Jackson, K.L., Marsteller, F., McKeown, R.E., & Addy, C. (1990). A longitudinal study of depressive symptomatology in young adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*, 581-585.
- Gelfand, D.M., & Teti, D.M. (1990). The effects of maternal depression on children. *Clinical Psychology Review, 10*, 329-353.
- Gerlsma, C., Emmelkamp, P.M.G., & Arrindell, W.A. (1990). Anxiety, depression, and perceptions of early parenting: A meta-analysis. *Clinical Psychology Review, 10*, 251-278.
- Ginsburg, G.S., Silverman, W.K., & Kurtines, W.K. (1995). Family involvement in treating children with phobic and anxiety disorders: A look ahead. *Clinical Psychology Review, 15*, 457-473.
- Goldenthal, P. (1991). Contextual therapy with children and families. In P.A. Keller & S.R. Heyman (Eds.), *Innovations in clinical practice: A source book* (pp. 85-97). Sarasota: Professional Resource Press.
- Gottman, J.M., & Katz, L.F. (1989). Effects of marital discord on young children's peer interaction and health. *Developmental Psychology, 25*, 373-381.
- Haan, E. de (1991). Angstige ouders, angstige kinderen? *Kind en Adolescent, 12*, 167-175.
- Haley, J. (1963). *Strategies of psychotherapy*. New York: Grune & Stratton. Nederlandse uitgave: *Strategieën in de psychotherapie*. Utrecht: Bijleveld.
- Haley, J. (1980). *Leaving home*. New York: McGraw-Hill.
- Haley, J. (1989). The effect of long-term outcome studies on the therapy of schizophrenia. *Journal of Marital and Family Therapy, 15*, 127-132.
- Henggeler, S.W., Melton, G.B., & Smith, L.A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal Consulting and Clinical Psychology, 60*, 953-961.
- Holtzworth-Munroe, A., & Jacobson, N.S. (1991). Behavioral marital therapy. In A.S. Gurman & D.P. Kniskern (Eds.), *Handbook of family ther-*

- py, Vol. 2 (pp. 96-133). New York: Brunner/Mazel.
- Kashani, J.H., Vaidya, A.F., Soltys, S.M., Dandoy, A.C., Katz, L.M., & Reid, J.C. (1990). Correlates of anxiety in psychiatrically hospitalized children and their parents. *American Journal of Psychiatry*, 147, 319-323.
- Kaslow, N.J., Deering, C.G., & Racusin, G.R. (1994). Depressed children and their families. *Clinical Psychology Review*, 14, 39-55.
- Keith, D.V., Connell, G.M., & Whitaker, C.A. (1991). The symbolic-experiential approach to the resolution of therapeutic obstacles in family therapy. *Journal of Family Psychotherapy*, 2, 41-56.
- Keller, M.B., Beardsley, W.R., Dorer, D.J., Lavori, P.W., Samuelson, H., & Klerman, G.R. (1986). Impact of severity and chronicity of parental affective illness on adaptive functioning and psychopathology of children. *Archives of General Psychiatry*, 43, 930-937.
- Lahey, B.B., Hartdagen, S.E., Frick, P.J., McBurnett, K., Connor, R., & Hynd, G.W. (1988). Conduct disorder: parsing the confounded relation to parental divorce and antisocial personality. *Journal of Abnormal Psychology*, 97, 334-337.
- Lange, A. (1994). *Gedragsverandering in gezinnen*, 6e compleet herziene druk. Groningen: Wolters-Noordhoff.
- Lange, A., Blonk, R., Wiers, R., & Schaar, D. van der (in druk). De Ouder-Kind Interactie Vragenlijst (OKIV). *Kind en Adolescent*.
- Laub, J.H., & Sampson, R.J. (1988). Unraveling families and delinquency: a re-analysis of the Gluecks' data. *Criminology*, 26, 355-380.
- Lederer, W.J., & Jackson, D.D. (1968). *The mirages of marriage*. New York: Norton.
- Markus, E., Lange, A., & Pettigrew, T. (1990). The effectiveness of family therapy: A meta-analysis. *Journal of Family Therapy*, 12, 205-222.
- Messer, S.C., & Beidel, D.C. (1994). Psychosocial correlates of childhood anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 33, 975-983.
- Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press. Nederlandse uitgave: *Gezinstherapie*. Utrecht: Het Spectrum, 1973.
- Minuchin, S., & Fishman, H.Ch. (1981). *Family therapy techniques*. Cambridge, MA: Harvard University Press.
- Minuchin, S., Montalvo, B., Guerney Jr., B.G., Rosman, B.L., & Schumer, F. (1967). *Families of the slums. An exploration of their structure and treatment*. (pp. 19) New York: Basic Books.
- Munson, C.E. (1994). Cognitive Family Therapy. In D.K. Granvold (Eds.), *Cognitive and behavioral treatment; Methods and applications* (pp 19). Pacific Grove, CA: Brooks/Cole.
- O'Brien, M., Margolin, G., John, R.S., & Krueger, L. (1991). Mothers' and sons' cognitive and emotional reactions to simulated marital and family conflict. *Journal of Consulting and Clinical Psychology*, 59, 692-703.
- Osterhaus, S.O.L., Lange, A., & Richard, R. (in voorbereiding). The influence of family factors on the reporting of secondary symptoms in young headache patients.
- Patterson, G.R. (1982). *Coercive family process*. Eugene, OR: Castalia Publishing Co.
- Patterson, G.R., Reid, J.B., Jones, R.R., & Conger, R.E. (1975). *A social learning approach to family intervention: Vol. 1. Families with aggressive children*. Eugene, OR: Castalia.
- Pinsof, W.M., & Wynne, L.C. (1995). The efficacy of marital and family therapy: an empirical overview, conclusions, and recommendations. *Journal of Marital and Family Therapy*, 21, 585-613.
- Prince, S.E., & Jacobson, N.S. (1995). A review and evaluation of marital and family therapies for affective disorders. *Journal of Marital and Family Therapy*, 21, 377-401.
- Reinherz, H.Z., Stewart-Barghauer, G., Pakiz, B., Frost, A.K., Moeykens, B.A., & Holmes, W.M. (1989). The relationship of early risk and current mediators to depressive symptomatology in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, 28, 942-947.
- Satir, V. (1964). *Conjoint family therapy*. San Francisco, CA: Science and Behavior Books.
- Satterfield, J.H., Satterfield, B., & Schell, A.M. (1987). Therapeutic interventions to prevent delinquency in hyperactive boys. *Journal American Academy of Child & Adolescent Psychiatry*, 26, 56-64.
- Sayger, T.V., Horne, A.M., & Glaser, B.A. (1993). Marital satisfaction and social learning family therapy for child conduct problems: Generalization of treatment effects. *Journal of Marital and Family Therapy*, 19, 393-402.
- Schmaling, K.B., Fruzzetti, A.E. & Jacobson, N. (1989). Marital problems. In K. Hawton, P.M. Salkovskis, J. Kirk & D.M. Clark (Eds.), *Cognitive behaviour therapy for psychiatric problems; A practical guide* (pp. 339-369). Oxford: Oxford University Press.

- Segal, L. (1991). Brief therapy: The MRI approach. In A.S. Gurman & D.P. Kniskern (Eds.), *Handbook of family therapy, Vol. 2* (pp. 171-199). New York: Brunner/Mazel.
- Selvini Palazzoli, M., e.a. (1979). *Paradox en tegenparadox*. Alphen aan den Rijn: Samsom.
- Shadish, W.R., Ragsdale, K., Glaser, R.R., & Montgomery, L.M. (1995). The efficacy and effectiveness of marital and family therapy: A perspective from meta-analysis. *Journal of Marital and Family Therapy, 21*, 345-360.
- Skinner, H.A., Steinhauer, P.D., & Santa-Barbara, J. (1983). The Family Assessment Measure. *Canadian Journal of Community Mental Health, 2*, 91-105.
- Stark, K.D., Humphrey, L.L., Crook, K., & Lewis, K. (1990). Perceived family environments of depressed and anxious children; Child's and maternal figure's perspectives. *Journal of Abnormal Child Psychology, 18*, 527-547.
- Stuart, R.B. (1980). *Helping couples change*. New York: The Guilford Press.
- Watzlawick, P., Beavin, J., & Jackson, D.D. (1967). *Pragmatics of human communication*. New York: Norton. Nederlandse uitgave: *De pragmatische aspecten van de menselijke communicatie*. Deventer: Van Loghum Slaterus, 1970.

Manuscript aanvaard 6-3-1997

Auteurs

A. Lange is als bijzonder hoogleraar verbonden aan de vakgroep Klinische Psychologie van de Universiteit van Amsterdam.

J.-P. van de Ven is verbonden aan het ambulatorium van de Faculteit der Psychologie, vakgroep Klinische Psychologie van de Universiteit van Amsterdam.

Correspondentie-adres: Universiteit van Amsterdam, Roetersstraat 15, 1018 WB Amsterdam

Abstract

Systemic approaches in the treatment of children and adolescents

A. Lange & J.-P. van de Ven. *Pedagogische Studiën*, 1997, 74, 271-284.

This article discusses the system theoretical approaches that have contributed to the treatment of disorders in children and adolescents, such as depression, anxiety and conduct disorders. Subsequently, an overview is given of the research into the association of family variables and psychiatric disorders. This is followed by an overview of studies into the outcome of family treatment with regard to these disorders. The research emphasizes the importance of intervening in the direct interaction between parent and child (the child rearing pattern). Among the system theoretical approaches, mainly the behavioral, the structural, the communicational and the psycho-educational approach have been shown to be important. Intergenerational and cybernetic system approaches appear to be less effective. Finally, discrepancies between clinical findings and the outcome of controlled experimental studies are discussed, and the need for systematic studies into the effects of multi-modal treatments is expressed.

Inleiding (N. Verloop, ICLON, Rijksuniversiteit Leiden)

De goed geoliede conferentie-machine van de AERA draaide dit jaar in Chicago. De voorbereiding en de organisatie van deze conferentie heeft in de afgelopen tientallen jaren een graad van perfectie bereikt waar Europese organisaties als de EARLI en de EERA inmiddels met enige jaloezie naar kijken. Het grote verschil is natuurlijk dat de AERA het zich, met haar meer dan 20.000 leden, kan veroorloven er een flinke vaste staf op na te houden, terwijl kleinere organisaties het voornamelijk moeten hebben van, veelal per conferentie-locatie wisselende, groepen vrijwilligers. Het ziet er niet naar uit dat hierin voorlopig verandering zal komen, zeker nu het zelfs onhaalbaar is gebleken in 1999 een gezamenlijke conferentie van EARLI en EERA te organiseren. Naar goed Europees gebruik bleek bij het overleg hierover vooral de nadruk gelegd te worden op de onderlinge verschillen.

Een gunstige ontwikkeling binnen de AERA is dat het aantal specifieke groepen (SIG's en dergelijke) zich nog steeds uitbreidt. Er was dit jaar zelfs een nieuwe divisie in het leven geroepen, "Education Policy and Politics". Het voordeel hiervan is dat het voor de deelnemers steeds gemakkelijker wordt om collega-onderzoekers te treffen die met (vrijwel) dezelfde onderzoeksthematiek bezig zijn en niet te verdwalen in het gigantische aanbod.

Tijdens de vorige AERA was men gestart met het als zodanig markeren van bijeenkomsten die sterk op de praktijk gericht waren. Dit was niet alleen voor geïnteresseerde docenten functioneel maar ook voor onderzoekers. In het verleden kwam het immers vaak voor dat onderzoekers die geïnteresseerd waren in de onderzoeksmatige aanpak van een bepaalde thematiek, verzeild raakten in sessies waarin uitsluitend over praktische toepassingsproblemen werd gesproken. Dit jaar ging men nog

een stap verder door de sessies die 'of special interest to teachers' waren te concentreren op een bepaalde dag, gekoppeld aan de mogelijkheid zich alleen voor deze dag in te schrijven. Een ongetwijfeld zinnige poging tot het dichterbij elkaar brengen van theorie en praktijk.

Van de invloed van het conferentie-thema, *Talking together in Educational Research and Practice*, was verder weinig terug te vinden in het programma, althans niets dat afweek van hetgeen ook in andere jaren over deze problematiek aan bod kwam. Het definiëren van een 'thema' schijnt een gewoonte te zijn waar de organisatoren maar moeilijk vanaf te brengen zijn, ondanks het feit dat de divisies en SIG's zich, begrijpelijkerwijs, niets van een dergelijk thema aantrekken. Zelfs de constatering van de 'program chair' voor 1998, Christine Sleeter, dat zelfs onderzoekers die al tien jaar de AERA bezoeken niet eens weten dat er zoiets als een conferentie-thema bestaat, brengt haar er niet vanaf vol goede moed een nieuw thema te kiezen (Sleeter, 1997).

In het verslag is ook dit jaar weer gekozen voor een divisie-gewijze aanpak. Gerapporteerd wordt over 7 van de inmiddels 12 bestaande divisies van de AERA. Alle grote divisies zijn echter vertegenwoordigd, zodat van het totale aanbod hier ongeveer driekwart wordt bestreken. Dit laat uiteraard onverlet dat elke auteur uit het grote aanbod binnen een bepaalde divisie een selectie moest maken. Aan de auteurs is gevraagd vooral in te gaan op algemene trends die binnen hun divisie zichtbaar waren.

Literatuur

Sleeter, C. (1997). Annual meeting 1998. *Educational Researcher*, 26(5), 34.

Divisie A: Administration (M.L. Krüger, Universiteit van Amsterdam)

Divisie A kent vier secties: organizational contexts; school reform and leadership studies; policy perspectives in administration; teaching and learning in educational administration. Rond deze vier thema's worden ook de papers binnen de sessies op het jaarlijkse congres gegroepeerd. Er waren dit jaar 75 sessies georganiseerd: 36 papersessies naast 39 meer interactieve sessies zoals symposia en ronde-

tafelbijeenkomsten. Voor bezoekers van de AERA is het moeilijk om een goede keuze uit het grote aanbod te maken. De keuze wordt al snel bepaald op grond van namen, maar dat blijkt maar al te vaak geen waarborg voor kwaliteit. Omdat ik het meest geïnteresseerd ben in leiderschapsstudies en mijn keuze voor sessies dikwijls daardoor werd bepaald, zal de nadruk in deze bespreking onvermijdelijk op dit thema liggen. De bespreking van bijgewoonde sessies zal geconcentreerd worden rond trends die dit jaar binnen divisie A konden worden geconstateerd.

Onderzoekers binnen divisie A lijken steeds sterker op zoek te gaan naar de relatie theorie-praktijk. Wat is het nut van organisatieonderzoek voor de onderwijspraktijk? Deze trend komt ook tot uiting in het centrale thema dat voor alle divisies was vastgesteld, namelijk: 'Talking together in Educational Research and Practice: how can we increase the impact of new knowledge in education on educational practice?' De vraagstelling naar de wijze waarop door onderzoek vastgestelde effectieve schoolkenmerken geïmplementeerd kunnen worden in de schoolpraktijk, is niet nieuw binnen het onderzoek naar organisatie en management. Wel lijken steeds meer onderzoekers de vraag te stellen. Anderzijds wordt onderzoekskennis vaak om politieke redenen niet gebruikt bij beleidsvorming. Seashore Louis verwoordde dat heel mooi: "Schools in America are so big because of football".

Nauw verbonden aan de vraag naar het nut van onderzoeksmatige kennis voor de praktijk, is de vraag naar de richting waarin toekomstig onderzoek zich dient te ontwikkelen. Deze vraagstelling lijkt ingegeven door de veranderingen in het onderwijs en de veranderende rol van de schoolleider in de jaren negentig in de USA en in Europa. Die veranderingen stonden in het merendeel van de sessies centraal, zoals al uit titels van sessies blijkt: 'School change: leadership challenges'; 'Organizational change: the effects of and on leadership'; 'Organizational change and school reform'; 'Organizational leadership and school change'; 'Challenges and transitions in school leadership', enzovoorts. In de sessie 'School leadership, research methods and theoretical framework' werd expliciet ingegaan op mogelijke richtingen voor toekomstig onderzoek. Heck

en Hallinger leidden de sessie in met de vraag: "Where do we come from and where are we going?" Interessant is dat niet meer gezocht wordt naar één perspectief van waaruit onderzoek plaatsvindt, maar naar een integraal perspectief. Men kiest niet meer voor alleen het functionele perspectief of voor het perspectief van restructuring. Ook transformationeel leiderschap lijkt op zijn retour. Volgens Marsch zal er in de komende 10 jaar steeds meer nadruk komen op een 'strategic results driven and an educational improvement perspective'. Opvallend was dat de presentatoren in deze sessie in het zoeken naar nieuwe richtingen in onderzoek allen op de een of andere wijze trachtten om een constructivistisch perspectief in hun onderzoek te incorporeren. Reed vatte dit als discussiant samen als: "knowledge and the need for sense-making aspects in leadership are being focussed on".

Nu door het autonoom worden van scholen de accountability van de school groter wordt, wordt onderzoek interessant naar wijzen waarop scholen werken aan kwaliteitsbewaking. Er was zelfs een serie van vier symposia over accountability (samen met divisie L: Education Politics and Policy) onder de titel 'Negotiating standards and accountability for student performance and system change'. Schoolleiders worden steeds meer gezien als change agents, of als katalysatoren van verandering zoals een discussiant het verwoordde. Het gevaar dat daarin m.i. schuilt is dat bij het niet goed verlopen van veranderingen de beschuldiging vinger naar de schoolleider wordt uitgestoken. Dat is tegenstrijdig aan een andere trend die ook te constateren was: het besteden van aandacht aan 'distributed leadership' of 'leadership throughout the organization'. Dit denken is in gang gezet door twee samenhangende artikelen die zijn verschenen in de Educational Administration Quarterly in 1995 van Ogawa en Bossert en van Pounder, Ogawa en Adams. Zij spreken over 'leadership as an organizational quality'. Zo waren er diverse sessies over 'teacher leadership', vaak gekoppeld aan 'teacher decision making'. Er wordt in de USA al langer onderzoek verricht naar teacher leadership, maar de betekenis van het begrip verandert. Het gaat niet meer zozeer om het onderzoeken van formele rollen of functies van docenten als teacher leader, maar om

onderzoek naar de invloed van de leraar als informele leider. Nu onderzoek begint uit te wijzen dat de invloed van de schoolleider op de leerlingresultaten niet zo groot is als we misschien gehoopt hadden, wordt gekeken wie dan wel invloed hebben op het primaire proces en hoe. Lynn Zinn verbond de aandacht voor teacher leadership aan veranderingen in scholen waardoor ook de rollen van docenten veranderen. Aan scholen wordt geadviseerd om de definitie van onderwijzen te verbreden zodat leraren tevreden blijven tijdens hun carrièreverloop. Dat de aandacht voor deze vorm van teacher leadership een trend is geworden, blijkt ook uit het feit dat Leithwood, een bekende onderzoeker op het terrein van schoolleiderschap, dit jaar samen met anderen meerdere papers presenteerde over teacher leadership. Volgens Leithwood e.a. wordt in alle schoolvernieuwingsplannen gepleit voor gedeelde besluitvorming en professionalisering van docenten. "Both elements require teachers routinely to exercise more leadership outside the classroom than traditionally has been expected from them". Uit hun onderzoek blijkt dat de invloed van de schoolleider 1/3 sterker is dan die van de teacher leader. De invloed van de teacher leader is het sterkst en sterker dan de schoolleidersinvloed op de aspecten schoolplanning en schoolstructuur/schoolorganisatie. De invloed van de schoolleider op deze aspecten is trouwens ook hoog, evenals op de aspecten missie en schoolcultuur waarop de docenten veel minder invloed hebben. Opvallende resultaten die ervoor pleiten dergelijk onderzoek ook in Nederland te verrichten. Anderzijds blijkt dat schoolleiders niet onbelangrijk zijn. Hier is de presentatie van Lee in één van de andere sessies vermeldenswaard. Zij onderzocht de bijdrage van schoolleiders aan het creëren van een omgeving voor docenten die bevorderlijk is voor onderwijsinnovatie. Alhoewel docenten op de door haar onderzochte scholen actief participeerden in de besluitvorming en teacher leaders rollen vervulden, zagen zij de steun van de schoolleider als zeer belangrijk. De schoolleider bleek een sleutelrol te vervullen in het helder maken van het belang van pogingen van docenten om innovatie in de klaslokalen door te voeren. Voor het onderzoeken van leiderschap heeft de hier beschreven verschuiving van perspectief nogal wat com-

plicerende consequenties. Wie en wat moet onderzocht worden om leiderschap in beeld te brengen?

Ook in het gepresenteerde onderzoek naar de opleiding van schoolleiders werd voortdurend een relatie gelegd met de veranderende omgeving. Hoe moeten opleiders omgaan met de 'changing world in which the principal operates?' zoals Matczynski het verwoordde. Lori vroeg zich af hoe nieuwe schoolleiders op te leiden nu door de veranderende omgeving van scholen er een gebrek is aan rolmodellen. In de verschillende sessies werd onderzoek gepresenteerd naar effecten van en ervaringen met moderne versus traditionele opleidingsprogramma's. In Amerika is, in tegenstelling tot in ons land, een diploma voor schoolleiders verplicht en is permanente nascholing heel normaal. De ontwikkeling van een opleidingsaanbod met daaraan gekoppeld onderzoek naar opleidingen is dan ook belangrijk in de USA. In Nederland ontbreekt het nagenoeg aan het onderzoeksmatig ontwikkelen van schoolleidersopleidingen. De relatie tussen onderzoek en training wordt nauwelijks gelegd, terwijl dat in deze tijd van snelle veranderingen in het onderwijs toch van wezenlijk belang zou zijn.

Op de AERA wordt onderzoek naar sekse doorgaans in aparte sessies geclusterd (met name binnen de Special Interest Group 'Research on women in education') die bijna voor 100% door vrouwen worden bezocht. Dit gebeurt overigens niet alleen met het onderzoek naar sekse, maar ook met andere 'minderheden'. Een voorbeeld is de SIG 'Critical examination of Race, Ethnicity, Class and Gender'. Vorig jaar was ik voor de presentatie van mijn paper over sekse en leiderschap ingedeeld in een sessie die de naam had gekregen 'Striding towards diversification of the professional pool: Focus on Women and African Americans'. En wat te denken van de receptie 'Welcoming women and people of color into AERA' die elk jaar wordt georganiseerd? Dit jaar bleek dat sekse iets vaker was meegenomen als variabele in divers onderzoek, waardoor het thema aan de orde kwam in meerdere sessies. Bijvoorbeeld in sessies over onderzoek naar de invloed van vernieuwingen op schoolleiders (Harold), naar de opleiding van schoolleiders (Alley & MacDonald), naar rollen van schoolleiders in diverse nascholingsprogram-

ma's (Read, Brown & Royal), en naar de aspiraties van schoolleiders (Read). Het sekse-aspect in leiderschap lijkt een iets bredere belangstelling te krijgen.

Divisie A wordt van oorsprong gedomineerd door kwantitatief onderzoek. In de afgelopen jaren is men op zoek gegaan naar alternatieve wijzen van en perspectieven op het verrichten van onderzoek. Deze trend lijkt zich vooralsnog slechts te versterken. Etnografisch onderzoek, narratief onderzoek ('What do principals see and say themselves?'; 'Stories of principals count'; 'Talking heads') en onderzoek met wel zeer kleine responsaantallen, soms zelfs maar één casus ('N=1 onderzoek' zouden we dat kunnen noemen), werd veelvuldig gepresenteerd. Aan postmodern onderzoek was een afzonderlijk symposium gewijd onder de passende titel 'Deconstructing the Talk, Reconstructing the Walk: Modern promises for Teaching Theory and Practice in Educational Administration'. Toch niet de onbekendsten (Fenwick English, William Foster, Gail Furman en Peter McLaren) presenteerden hun papers in deze sessie. Centraal stond dat er meerdere antwoorden op dezelfde vraag zijn. Foster is van mening dat het niet gaat om 'truth', maar om 'historically developing knowledge'. Daarin is leiderschap volgens hem 'a contested concept'. Wat het betekent voor het onderzoek wanneer zijn criterium 'create our own definitions' zou worden toegepast, deed hij jammer genoeg niet uit de doeken. De kloof theorie-praktijk lijkt in het postmoderne onderzoek slechts te worden vergroot, zoals ook bleek uit een vraag in de discussie: "Kunt u mij zeggen waar ik als gewone schoolleider literatuur kan vinden die voor mij toegankelijk is?" De stromingen lijken zich steeds sterker te polariseren in het denken over toekomstig onderzoek. Hieruit zou weleens de interesse kunnen zijn ontstaan voor een andere manier van denken over onderzoek die men lijkt te vinden in wat ik maar even noem de Europese (Nederlandse?) onderzoeksbenadering. In de Nederlandse traditie is het onderscheid tussen kwantitatief en kwalitatief onderzoek minder sterk gepolariseerd: beide soorten van onderzoek kunnen hoogstaand van kwaliteit zijn. De discussie over toekomstig onderzoek centreert zich niet zozeer rond kwantitatief of kwalitatief, maar rond het vinden van werkbare con-

ceptualisering en valide operationalisering. In divisie A werden twee symposia gepresenteerd met veel of alleen maar Nederlandse inbreng: 'Comparative perspectives on school organization, teacher's work and social innovation' (Leiden, Leuven, Nijmegen, Minnesota en Londen) en 'Educational leadership reconsidered: Opening the black box to identify key factors in school improvement (Leiden, Twente, Nijmegen en Amsterdam (UvA) met Leithwood (chair) en Hallinger (discussant)). In beide symposia werd duidelijk dat het Amerikaanse publiek het problematiseren van thema's in plaats van het geven van kant en klare antwoorden nieuw en boeiend vond, evenals het discussiëren over conceptualisering. Ook de stellingname dat sommige onderzoeksvragen kwantitatief en andere kwalitatief dienen te worden aangepakt, leek een nieuwe focus. Dit bleek ook uit het uitnodigen van presentatoren voor publicaties in Amerikaanse tijdschriften. Er lijkt waardering te ontstaan voor wat ik maar even gemakshalve noem de Europese wijze van denken over onderzoek. Het feit dat in de USA in tegenstelling tot in Europa zo'n groot congres rond onderwijsonderzoek kan worden georganiseerd, ligt eerder aan de kwantiteit dan aan de kwaliteit van de onderzoekingen. Misschien is het met de internationaliseringskansen van het Nederlandse onderwijsonderzoek minder slecht gesteld dan wel wordt beweerd.

Divisie B: Curriculum Studies (J. Terwel, Vrije Universiteit, Amsterdam)

Eén van de charmes van het thema 'curriculum' is zijn breedte en diversiteit. Als onderzoeker hoeft men zich niet gevangen te voelen op een beperkt gebied waarvan de theoretische, praktische en maatschappelijke relevantie discutabel is. Het thema curriculum staat ook niet geïsoleerd ten opzichte van andere thema's. Curriculum als object van studie vormt eigenlijk het brandpunt waarin alle andere thema's samenkomen. Dat geldt in het bijzonder als men het curriculum niet alleen ziet als een document, maar ook als een leerweg van de leerling: curriculum vitae.

Het is dan ook niet verwonderlijk dat het domein 'curriculum' vogels van verschillende pluimage aantrekt: filosofen, historici, economen, sociologen, psychologen, antropologen,

pedagogen, vakdidactici, multi-mediatechnologen, en nog veel meer. Dat was ook dit jaar in Chicago weer het geval. Er waren verschillende invalshoeken en er was veel beweging. Om enkele thema's te noemen: curriculumtheorie, constructivisme, authentiek leren en onderwijzen, gender-vraagstukken, schoolgerichte curriculumontwikkeling, curriculum en ongelijkheid, hoogbegaafdheid, curriculum paradigma's, postmoderne filosofie, leraren, ouders, community, onderzoekers en curriculumvernieuwing, intercultureel onderwijs, wiskunde, natuurkunde, nieuwe media en technologie, economische en historische perspectieven, chaostheorie, alternatief onderwijs, onderwijs-politiek, etniciteit en cultuur, opleiding van leraren enz.

Met deze opsomming van inhoudelijke onderwerpen is nog weinig gezegd over de ontwikkelingen in curriculumresearch. Al in de jaren zeventig en vooral in het begin van de jaren tachtig, begon op de AERA het behaviouristische bolwerk te kraken. Rigide prescripties in het 'curriculum als document' werden vervangen door meer open ideeën en werkwijzen. Langzaam maar zeker kwamen nieuwe benaderingen in beeld. Andere definities van het curriculum kregen een kans. Anno 1997 heeft een aflossing van de wacht plaatsgevonden. Grote namen als Tyler, Scriven, Stufflebeam, Mager kom je niet meer tegen. Popham was er dit jaar nog wel, maar hij houdt zich nu niet meer met curriculum (gedragsdoelen) bezig, maar met statistische vraagstukken. Voor een deel voltrekt zich hier een verschuiving van generaties. Voor een ander deel zien we hier een verschuiving van paradigma's, of sterker gezegd een oorlog!

Onderzoekers met een bredere, meer 'humanistische', filosofische en/of meer praktische oriëntatie als Goodlad, Eisner, Walker, Jackson geven nu (nog steeds) de toon aan. Dit zijn ook de auteurs die bijdragen leverden aan het Handbook of Research on Curriculum, met Jackson als redacteur (1992). De Stanford University blijft op curriculumgebied interessante bijdragen leveren dankzij onderzoekers/theoretici als Eisner, Cuban, Shulman en Walker. Deze laatste was voorzitter van een boeiend symposium getiteld 'Constructivisme, Interaction, and Authentic Learning: Implications for Curriculum Theorie and Practice'. Bijdra-

gen aan dit symposium werden geleverd door onderzoekers uit de Verenigde Staten, Australië en Nederland. Speciale aandacht werd besteed aan het vraagstuk van 'top-down' versus 'bottom-up' benaderingen bijvoorbeeld bij het leren van wiskunde en talen. In bepaalde constructivistische kringen lijkt men eenzijdig te kiezen voor de leefwereld en de idiosyncratische constructies van de leerlingen. De structuur van de wetenschappen komt dan niet in beeld. Hetgeen Walker de kritische vraag ontlokte: 'Tennis is toch ook veel leuker met een net dan zonder?' In relatie hiermee koos Walker op nuchtere en evenwichtige wijze positie in centrale dilemma's op het curriculum-domein. Bijvoorbeeld door de exclusieve betekenis van theorieën als het constructivisme te relativieren. Hij verwees naar het gegeven dat curriculum eerst en vooral een praktische onderneming is (onder verwijzing naar zijn leermeester Joseph Schwab). Walker is van mening dat deze praktische onderneming gebaat is bij goed overleg (deliberatie), een uitgangspunt dat aan de basis lag van het CURVO-project (een Nederlands curriculum-project uit de jaren zeventig en tachtig).

Een alternatieve citeergemeenschap op het gebied van curriculum bestaat uit de redactie van het handboek 'Understanding Curriculum', met name Pinar, Reynolds, Slattery en Taubman (1995). Deze groep was ook nu weer goed vertegenwoordigd op de AERA in Chicago. Zij brengt thema's in als: geschiedenis van het curriculum, conflict en crisis, politiek en curriculum, gender, kunst en theater, en theologie. Men zet zich af tegen oriëntaties waarin het curriculum primair in termen van 'ontwikkeling' of 'constructie' wordt opgevat. Deze groep wil het curriculum 'begrijpen' door kritische analyse en theorievorming. Deze verschuiving is te zien als een 'reconceptualisatie' van het curriculum als veld van studie. Hun artikelen zijn vooral aan te treffen in 'Curriculum Inquiry'.

Dan is er nog een beweging die ik tot het terrein van curriculumstudies reken: de nieuwe Informatie en Communicatie Technologie (ICT) in het onderwijs. Bijna bij het scheiden van de markt, die de AERA-meeting toch ook is, waren er boeiende symposia op het gebied van ICT. Zo hield redacteur David Jonassen zijn nieuwe handboek ten doop met als titel:

'Handbook of Research for Educational Communications and Technology' (1996). De centrale boodschap was verrassend: we moeten stoppen met systemen die de leraar proberen te vervangen. Bij alle nieuwe technologieën wordt in de eerste fase geprobeerd de oude productiewijze te vervangen. Daarna ontdekt men pas het volle potentieel van de nieuwe technologie. Dat potentieel ligt niet in een voortzetting van hetzelfde met andere middelen, maar in het verkennen en openleggen van nieuwe mogelijkheden. Het gaat nu veel meer om de vraag 'hoe betrekken we de leraar hierin en wat betekent dit voor het curriculum'? Jonassen noemde het werk van de Cognition and Technology Group at Vanderbilt University (Bransford c.s.) als een excellent voorbeeld van een benadering die juist vanuit die integrale visie is opgezet. Met name het 'Jasper project' met multi-mediale toepassingen werd in dit verband genoemd. Filosofen als Foucault hebben de nieuwe ICT als 'dehumaniserend' aan de kaak gesteld. Men kan deze technologie ook positief benaderen en stellen dat het voor velen toegang tot 'resources' kan bieden die voordien voor hen gesloten bleven, bijvoorbeeld gehandicapte mensen. Geheel in lijn met deze positieve visie was er een symposium over 'Leren door Ontwerpen'. Hoe leerlingen science leren begrijpen door hen te betrekken in het oplossen van ontwerpproblemen zoals het ontwerpen van een robot of een kunstmatige long. Hier kwam naar voren dat leren door ontwerpen een veelbelovend perspectief is maar dat er nog veel werk moet worden verricht om de nieuwe media ook werkelijk te integreren in de sociale en curriculaire context van school en klas. Dit bracht de voorzitter tot de minder vleiende opmerking over leraren: 'je kunt een oude hond geen nieuwe kunstjes leren'. In de discussie ging Roth in op de vraag hoe je leerlingen kunt betrekken in de 'culturele praktijk' van een 'expert designer'. Bij alle enthousiasme voor ICT blijft er een punt dat onderzoekers op het gebied van ICT en curriculum zich moeten aantrekken: de gebieden zijn tot nu toe te sterk gescheiden. Convergentie en integratie is nodig. Zonder verbinding met het curriculum en de opleiding van leraren dreigt ICT een 'Fremdkörper' in het onderwijs te blijven.

De internationale groep rondom Westbury en Hopmann was ook dit jaar weer present. In deze groep wordt geprobeerd de Amerikaanse curriculum studies te relateren aan Europese tradities op het gebied van didactiek en curriculum, met bijdragen van o.a. Klafki, Schaller, Shulman, Doyle, Posner, Van Manen en Westbury. Dit jaar werd de scope nog verbreed in een comparatieve studie naar beleidsaspecten en implementatievraagstukken in vijf Europese landen. Deze groep heeft ook een 'eigen' handboek met als titel 'Didaktik and/or Curriculum: Basic Problems of Comparative Didaktik' onder redactie van Hopman en Riquarts (1995), uitgegeven door het Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel. Het werk van deze groep wordt ondersteund door artikelen in Journal of Curriculum Studies over 'grote Europese didactici' in de didactiek (Klafki, Herbart, Freudenthal). Dit is een buitengewoon interessant project: 'didaktik meets curriculum'. Dit project is vooral voor de Nederlandse onderwijskunde en onderwijspedagogiek zo spannend omdat hier zo'n 25 jaar geleden de Duitse didactiek werd afgezworen en een liefdesverklaring werd afgegeven aan het adres van de Amerikaanse curriculum- en instructietheorie (onder meer Glasers model dat aan de basis ligt van het model Didactische Analyse van Van Gelder). Klafki probeerde het 'Bildungstheoretische' denken te verbinden met het praktisch didactisch handelen, maar dat werd hier in Nederland als minder geslaagd beschouwd. De Bildungstheoretische didactiek werd als elitair gekwalificeerd en mede verantwoordelijk gehouden voor de scheiding tussen het algemeen vormend onderwijs en beroepsopleiding.

Dit brengt mij tot de volgende overweging. Wie de ontwikkelingen in de onderwijswetenschappen over een langere periode beschouwt kan constateren dat wetenschappelijke en pedagogische criteria van uitvoerbaarheid, effectiviteit en rechtvaardigheid van bepaalde modellen slechts een ondergeschikte rol spelen. Maatschappelijke en politieke factoren lijken dominant in het denken over onderwijs en opvoeding. Misschien is het curriculum wel één van de eerste en gevoeligste gebieden in het onderwijs waar de maatschappelijke strijd zich manifesteert. De hernieuwde aandacht voor bijvoorbeeld Klafki lijkt eerder ingegeven

door maatschappelijke en politieke factoren dan door wetenschapsinterne ontwikkelingen. Het categoriale Gymnasium is 'in', de 'comprehensive' school is 'uit'. Wie maakt zich nog druk om het educatief en sociaal isolement van leerlingen in het beroepsonderwijs en van werkende jongeren in de 'bouw' of in de 'metaal'? 'Parental Choice' waart rond als een losgebroken tijger. 'De school aan de ouders', lijkt het motto voor deze tijd te zijn. In de Angelsaksische wereld kijkt men soms jaloers naar Nederland als het Mekka van 'Parental Choice'!

Gelukkig is er ook een beweging die iets probeert te doen aan de toenemende sociale kloof in het onderwijs, gegeven de dominantie van de 'markt' en het duale schoolsysteem. Een belangrijk thema op de AERA-meeting was dan ook het nadenken over strategieën voor individuele leerlingen om uit achterstandssituaties te ontsnappen. Hoe kunnen leerlingen hun eigen schoolloopbaan ter hand nemen? Het denken in termen van structuurverandering, collectief verzet en weigering is anno 1997 niet meer zo populair. Nu gaat het meer om de vraag: hoe kan men begaafde leerlingen uit lagere inkomensgroepen ondersteunen om binnen de bestaande schoolstructuur succesvol te zijn. De weg van collectieve emancipatie blijkt moeilijk begaanbaar. De macro-sociologische reproductietheorieën vormen geen inspiratiebron meer voor het gros van de onderzoekers. Nu lijkt de individuele emancipatie meer in de belangstelling te staan. Hoe kunnen deze leerlingen toegang krijgen tot de 'resources'. Welke strategieën hanteren ouders om leerlingen te beschermen tegen de gevaren in hun directe leefomgeving? Hoe spelen leerlingen het klaar om enerzijds hun achtergrond niet te verloochenen en anderzijds actief in de wereld van de school te staan. Het blijkt dat succesvolle leerlingen een duale identiteit ontwikkelen: een 'academic identity' en een 'streetwise identity'. Ze wisselen van rol als ze naar school gaan of naar huis.

Dit thema van 'multiple identities' klonk door in vele bijdragen. Penelope Peterson, in haar 'presidential address', maakte hier dan ook een belangrijk punt van. Persoonlijk heeft zij een ontwikkeling doorgemaakt. Zij stelde vast dat de traditionele psychologie haar greep op het onderwijsonderzoek begint te verliezen en dat er meer plaats wordt ingeruimd voor

gezichtspunten uit bijvoorbeeld de sociologie en antropologie. Zij ervaart zichzelf in toenemende mate als activiste. Ze illustreerde haar visie en stellingname aan de hand van haar ervaringen als moeder, maar ook als deelnemer aan een groep die opkomt voor het verbeteren van het onderwijs in Mexico. Is het geen schande dat zoveel kinderen niet kunnen lezen, zo vroeg zij zich af. Samen met Slavin wil zij proberen politici te engageren voor de minimale doelstelling dat alle kinderen moeten kunnen lezen als ze negen jaar zijn. We maken deel uit van verschillende gemeenschappen, zei Peterson. Er zijn zoveel verschillende stemmen (so many voices, different communities). Zij probeerde haar gehoor wakker te schudden met de boodschap dat onderwijsresearch zo'n lage waardering heeft bij politici, ouders en leraren, en dat stemmen van minderheidsgroepen in de samenleving niet worden gehoord.

Onderwijsonderzoekers zouden volgens Peterson meer voor deze groepen moeten opkomen. Interessant was ook te vernemen dat zelfs de kinderen van Peterson, (maar ook die van Christine Keitel, Vrije Universiteit Berlijn) worstelen met het vraagstuk van 'multiple identities'. Datgene wat thuis tot de cultuur behoort (zelf nadenken, eigen oplossingen bedenken, reflecteren, kritische vragen stellen) wordt op school niet gepikt. Daar worden deze leerlingen als het ware teruggezet: ze moeten stomme platen inkleuren en algoritmische kunstjes leren.

Concluderend kan worden gesteld dat de studie van het curriculum springlevend is: curriculum vitae. Men reageert alert op nieuwe impulsen uit de samenleving en uit de andere wetenschappelijke disciplines. De Annual Meeting van de AERA in Chicago was weer zeer inspirerend, al zou naar mijn smaak het echte degelijke, empirische onderzoek een wat sterker accent mogen krijgen. Ook de fragmentatie in curriculumstudies blijft een punt van zorg. Maar zolang de vertegenwoordigers van de verschillende paradigma's met elkaar in gesprek blijven, is er alle reden tot optimisme. Want in de uitwisseling van standpunten ligt de mogelijkheid van groei en ontwikkeling. Laat duizend bloemen bloeien!

Met haar meer dan 175 symposia, posterpresentaties en ronde-tafeldiscussies is Divisie C de grootste van de AERA. Ondanks dit enorme aanbod kan men zich toch afvragen of een samenvatting van alles wat onder deze divisie gerangschikt was, een goed beeld geeft van recente ontwikkelingen op het onderzoeksdomein van Leren en Instructie. Wie geïnteresseerd is in de kenmerken van leerprocessen die leerlingen met elkaar doormaken, deed er bijvoorbeeld goed aan om ook te letten op sessies van de Special Interest Group (SIG) 'Cooperative Learning - Theory, Research, and Practice' en op sommige symposia in divisie B (curriculum); wie geïnteresseerd is in de inbedding van leer- en instructieprocessen in sociaal-culturele activiteiten moest beslist over de grenzen van divisie C heen naar de Cultural-Historical-SIG en naar divisie G (Social Context of education); wie zich wilde verdiepen in leerprocessen bij jonge kinderen vond vooral veel van zijn of haar gading in de 'Early Education and Child Development-SIG'. Ook in verschillende andere divisies (ik noem slechts B: 'Curriculum Studies' en K: 'Teaching and Teacher Education') vonden diverse sessie plaats die voor een verrijking van ons inzicht in 'Leren en Instructie' van belang zijn. Kortom: wie naar de AERA gaat om ideeën op te doen en te toetsen rond 'Learning and Instruction' doet er goed aan zich niet al te stringent te houden aan de divisiegrenzen. Ik heb dat voor mijn beschrijving van de ontwikkelingen op het onderzoeksgebied 'Learning and Instruction' in elk geval niet gedaan.

Afgaande op de titels van de symposia (in divisie C) en de verwijzingen daarin naar specifieke onderwerpen, laten zich wel enkele trends aftekenen. Onderzoek waarin op de een of andere manier domeinspecifieke leerprocessen aan de orde zijn (mathematics, science, literacy) stond het vaakst op de agenda, in aantal gevolgd door presentaties waarin 'leren in context' of 'learning environment' aan de orde was. Daarnaast was er natuurlijk ook het nodige te doen rond het gebruik van nieuwe media (computer, virtual realities, internet) en instructietheorie en informatieverwerkingsprocessen. Op zich is dit allemaal niet vernieu-

wend, want die trends waren vorige jaren ook al zichtbaar. Wat niet wegneemt dat daaronder toch nog wel interessante ontwikkelingen te melden zijn.

Een opvallende trend in het onderzoek op het leergebied wiskunde is het onderzoek naar de rol van het taalgebruik voor het wiskundeleerproces. In het symposium 'If they're talking they're learning? Teachers interpretations of meaningful mathematical discourse' (georganiseerd door Dominic Peressini, University Colorado) werd onderzoek gepresenteerd naar het verloop van de discussies in de wiskundeles in relatie tot de leerresultaten van de leerlingen. Hoewel er vrij algemeen wordt aangenomen dat taalgebruik en discussie essentieel is voor leerprocessen waarin inzicht wordt nagestreefd, blijkt in de praktijk dat er aan nogal wat voorwaarden voldaan moet zijn (voorwaarden in termen van leerkrachtvaardigheden in het managen van een discussie, in termen van de aard van de problemen, in termen van sociale leeromgeving) wil er een positief effect vanuit gaan op de leerprocessen van leerlingen. Ondanks alle 'discussie' blijven leerlingen vaak toch wel erg gefixeerd op het vinden van de oplossing en besteden veel minder aandacht aan verklaringen en uitleg aan elkaar. In haar discussie bij dit symposium onderstreepte Hilda Borko het belang van wiskundig rijke problemen, maar zij plaatste tevens vraagtekens bij de competenties van veel leerkrachten om zulke discussies te leiden. Leerkrachten moeten dan zelf bijvoorbeeld ook goed weten wat 'wiskundige verklaring' eigenlijk is. Ditzelfde geluid was trouwens ook te horen in diverse andere symposia. Onder meer was dit het geval in een symposium van de 'Constructivist Theory, Research, and Practice' SIG getiteld 'Teaching based on a constructivist View of Learning': alleen het uitlokken van probleemgerichte discussies onder leerlingen is niet voldoende, de discussies moeten op een bepaalde manier georganiseerd zijn, moeten een beroep doen op de eigen oplossingen van de leerlingen, maar moeten ook een reflectie op onderliggende ideeën, normen en regels oproepen. De leerkracht moet hiertoe speciaal opgeleid worden, met gebruikmaking van goede voorbeelden.

In drie andere symposia, georganiseerd door Erna Yackle (Purdue), 'Learning opportunities

and classroom mathematical practices in an inquiry mathematics class', door Paul Cobb (Vanderbilt) 'Accounting for the learning and teaching mathematics in social context: participation in the classroom microculture', en een derde door Ellice Forman (Pittsburgh), 'Discourse based educational research: perspectives on the interface between discourse analysis and education' stond de problematiek van het (wiskunde) leren in relatie tot rijke problemen en discussie daarover meer in het bijzonder centraal. Bij alle specifieke opbrengsten van deze drie symposia afzonderlijk wordt er toch een gemeenschappelijke lijn zichtbaar die benadrukt, dat verklaringen interactief geconstitueerd worden, door gezamenlijke reflectie op symbolische modellen, op sociomathematische normen, op onderdelen van argumentaties (bijv. de grondslag van gebruikte argumenten). In de klas wordt zo eigenlijk een (mathematische) werkgemeenschap gecreëerd en leren leerlingen ook steeds beter aan die 'community of practice' deel te nemen. Daarbij is het onderzoek van Cobb c.s. (uitgevoerd in samenwerking met Koeno Gravemijer van het Nederlandse Freudenthal instituut) een prachtig voorbeeld van ontwikkelingsonderzoek dat zowel voor de ontwikkeling van de praktijk als de theorie over (wiskunde)onderwijs van grote waarde is.

Overigens is de notie van de klas of groep als een 'community of learners'/'community of practice' niet iets dat alleen in de wiskunde-les onderkend is. Het begrip 'Community of learners' was een van de hoofdthema's in diverse symposia. In een interessante bijdrage in het symposium 'Peer collaboration in learning and research' lieten Polman en Pea zien hoe in het kader van een science project de discussie tussen leerlingen het verloop van het hele project in hoge mate stuurt: projectformulering, formulering van vraagstelling, bijstellingen van vraagstelling, keuze en legitimering van onderzoeksmethode etc.. Deze zogeheten 'transformative communication' geeft de leerlingen dus een wezenlijke medeverantwoordelijkheid in de afwikkeling van het project; de leerkracht heeft hierbij meer de rol van coach dan van stuurman. De leerresultaten die in dit project bereikt werden, waren van een soort waar je als leraar op mag hopen: inzichtelijk en gekenmerkt door een hoge mate van betrok-

kenheid van de kant van de leerlingen.

In de diverse presentaties die zich op een of andere manier met literacy bezig hielden, was het opvallend dat vele daarvan gericht waren op 'emergent literacy'. Opmerkelijke trend daarin was dat het idee van leren lezen door het leren decoderen kennelijk zwaar onder druk staat. Vele presentaties gingen dan ook over hoe jonge kinderen kunnen worden ingeleid in dat deel van onze cultuur waarin geschreven teksten een belangrijk onderdeel vormen. In het symposium 'Theoretical perspectives on Early Literacy Learning' hield Janet Bus (Leiden) bijv. haar betoog over het belang van het voorlezen vanuit de attachmenttheorie; in dat symposium werd zij geflankeerd door diverse andere sprekers die vooral ingingen op de verschillende voorwaarden waaronder opgroeiende kinderen kunnen leren gaan deelnemen aan een 'lezersgemeenschap'. Inbedding van beginnende leesactiviteiten in voor kinderen relevante activiteiten, beschikbaarheid van middelen, communicatie, vraag-en-antwoord komen hier ook weer naar voren als belangrijke ingrediënten voor het aanzwengelen van de literacy-ontwikkeling. Wel werd door de discussiante (Lea McGee, Boston College) aan het eind van de sessie terecht opgemerkt, dat toekomstige analyses nog wel verder moeten verduidelijken wat nu precies de aard van de zich hier voltrekkende literacy-leerprocessen is: 'wát wordt er nu eigenlijk precies geleerd door deze kinderen?', was haar vraag.

Een veel terugkerende vraag bij de studie van leer- en instructieprocessen is altijd weer die naar de betrouwbaarheid, validiteit en 'fairness' van de leerresultaatmetingen. De metingen via strak gestandaardiseerde, schriftelijke (multiple choice) toetsen zijn niet altijd zo bevredigend waar het gaat om het bepalen van het inzicht of de betrokkenheid/interesse van de leerling in het leerresultaat, of voor het bepalen wat de leerling nu geleerd heeft van het leerproces zelf (strategieën? reflectie?). Bovendien toonde Barry Cooper en zijn collega's van de universiteit van Sussex (UK) aan in een symposium 'Cultural contexts of mathematics learning' dat kale toetsopgaven (zoals gebruikt in de toetsen bij het Engelse National Curriculum) in de regel een overschatting opleveren van de mathematische vaardigheden van de leerlingen, in vergelijking tot een meting van

dezelfde vaardigheden via realistisch ingebedde items. Kinderen uit arbeidersmilieus lijden daar bovendien zwaarder onder dan kinderen van hoger opgeleide ouders.

Gelet op dit soort resultaten, is het dan ook niet verwonderlijk dat verscheidene presentaties gericht waren op het verbeteren van de assessment-technieken (zoals in het door Martin Ippel mede georganiseerde symposium 'Cognitive Technologies for knowledge assessment') of op het exploreren van alternatieve methoden van assessment. Opvallend in dezen is de ruime mate van belangstelling die er nog steeds is voor portfolio's als instrument om getrouw leer- en ontwikkelingsprocessen te kunnen volgen van leerlingen in langlopende leerprocessen. Gezien de grote belangstelling voor dit soort assessment procedures in de praktijk laat het zich voorspellen dat deze trend zich nog wel even doorzet.

Het is eigenlijk onmogelijk om hier alle presentaties aan te halen die zeker de aandacht waard zijn omdat ze wellicht kiemen vormen van interessante toekomstige ontwikkelingen. Ik denk dan aan onderzoek dat op verschillende plaatsen gerapporteerd werd betreffende leerling denkbeelden over kennis en leren (de zgn 'epistemological beliefs'), onderzoek over 'authentiek leren' (o.a. Roelofs en Terwel) waarin authenticiteit niet meer uitsluitend in termen van persoonlijke interesse wordt opgevat, maar ook in termen van vakmatige kwaliteiten en buitenschoolse betekenissen. Verder is vermeldenswaard onderzoek op het terrein van de motivatie, dat breekt met de individualistische traditie van intrinsieke of extrinsiek motivatie, maar motivatie vooral ziet in termen van co-constructie van persoonlijke betekenissen, zoals in het poster-symposium georganiseerd door Eugene Matusov (Santa Cruz), 'motivation in action'. Ten slotte mogen in dit rijtje verschillende Nederlandstalige voordrachten niet onvermeld blijven. In het symposium 'Investigations of Thinking, Reasoning, and Strategy use in Mathematics' rapporteerde De Corte over onderzoek naar de ontwikkeling van schattingsstrategieën bij leerlingen. Hij kon laten zien dat de optimale strategie bestaat uit het handig kiezen van verschillende strategieën al naar gelang de kenmerken van de situatie en dat deze schattingsvaardigheid zich met

de leeftijd ontwikkelt. In een ander symposium ('Learning strategies: Person- or context-bound?'), stelde Lodewijks met een groep van zijn naaste (jonge) collega's het probleem aan de orde of leerstrategieën samenhangen met persoonlijkheidskenmerken of juist met situatiekenmerken. In vlot en goed gebrachte presentaties belichtten de sprekers dit vraagstuk – meer of minder direct – van verschillende zijden. In haar discussie betoogde Stella Vosniadou dat de zaak waarschijnlijk geen of-of is, maar veel ingewikkelder ligt, omdat bijvoorbeeld de werkzaamheid van situatiekenmerken bij de selectie van een leerstrategie misschien zelf ook wel weer door persoonskenmerken beïnvloed wordt.

In vogelvlucht is hierboven ingegaan op de diverse ontwikkelingen op het terrein van Learning and Instruction, welke zich laten uittekenen naar aanleiding van rapportages vanaf de werkvloer van het empirisch onderzoek. Maar ook op een conceptueel vlak is er een ander gaande. Ik wil hier drie opvallende gebeurtenissen memoreren.

In een grote, maar niettemin afgeladen zaal presenteerden Robert Sternberg, John Mayer, en Howard Gardner hun ideeën in het symposium 'Expanding our concept of intelligence: What's missing and What could we gain?'. Sternberg hield een gloedvol betoog, gelardeerd met tal van voorbeelden, over het belang van practical intelligence dat lang aan onze aandacht ontsnapt is, omdat er in de standaard testsituaties geen ruimte voor was. Niettemin is het een belangrijk soort intelligentie, aldus Sternberg, in allerlei praktische (levens-, werk-)situaties. Daarna kwam Mayer pleiten voor het bestaan van 'emotional intelligence' dat met name zichtbaar wordt in situaties waarin we emoties moeten herkennen, de informatie die daarmee gegeven wordt moeten achterhalen of wanneer we moeten discrimineren tussen adequate en niet-adequate emoties (gelet op de situatie). Ten slotte gaf Gardner nog eens een uiteenzetting over zijn ideeën met betrekking tot 'multiple intelligences'. Tegen deze zwaargewichten had de discussiant (Carl Bereiter) ten slotte niet zoveel steekhoudends meer aan te vullen of in te brengen.

Een eveneens zeer goed bezochte panel-discussie was gewijd aan het constructivisme,

onder de veelbelovende titel 'Midcourse critical examination of constructivist learning'. Kopstukken als Carl Bereiter, Ann Brown, Jo Campione, John Bransford, Gavriel Salomon, Marlene Scardamelia bogen zich over de vragen 'Wat hebben we met het constructivisme bereikt?' 'Wat werkt wel, wat niet?' en 'Hoe gaan we verder?' Alle sprekers wezen erop dat de constructivistische benadering zeker een rol heeft gespeeld bij de ontwikkeling van het inzicht in leerprocessen (constructief karakter, sociale inbedding), dat het heeft geleid tot het inzicht dat kinderen veel meer blijken te kunnen dan we altijd dachten, en ook behulpzaam is geweest bij het verbeteren van de professionaliteit van leerkrachten die nu beter dan vroeger onderwijs kunnen inrichten. Maar geen van de sprekers had het idee, dat men al helemaal klaar was met de vervulling van de opdracht van het constructivisme. Er is nog veel te doen: vooral het blootleggen van de structurele principes achter leerprocessen en het ontwikkelen van onderwijsstrategieën die een systematische voortgang een leergang moeten realiseren (ter voorkoming van wat Salomon het 'butterfly-effect' noemde, d.w.z. touch-and-go: even ergens aan tippen dan weer verder, voor een oppervlakkige ontmoeting met iets anders etc.). Op zich waren dit niet allemaal erg schokkende verhalen en al helemaal niet bijzonder kritisch ten aanzien van het constructivisme. Alleen Bransford wist enigszins te ontspannen aan dit soort algemene betogen: hij rapporteerde over zijn 'Schools for thought'-project en de successen en problemen die zij daarin hadden meegemaakt. Daarbij bleef Bransford niet staan bij louter het constructivisme als benadering, maar hij wist een omvattend onderwijsconcept voor het voetlicht te brengen, te zamen met enkele problemen die de implementatie daarvan met zich meegebracht had. Ondanks Bransfords interessante betoog bleef het symposium als geheel toch wat steken in voorspelbare, obligate uitspraken. Een omvattende conclusie valt uit deze 'critical examination' van het constructivisme niet te trekken. Maar wellicht is de aanwezigheid van dit symposium op zichzelf al een teken dat de reflectie op de grondslagen van het constructivisme – zij het schoorvoetend – op gang begint te komen.

Ten slotte verdient ook de prikkelende voordracht van Shirley Brice Heath voor de

Cultural-Historical SIG hier nog vermelding. Onder de titel 'Contemporary influences of Vygotsky and Bachtin: the ideology of genre and youth' ging zij in op de vraag naar het buitenschoolse leren en wilde vanuit haar kritische interpretatie van Vygotsky en Bachtin vooral helder krijgen wat er precies toe bijdraagt dat sommige jongeren in buitenschoolse situaties toch hoge prestaties blijven leveren. Volgens Brice Heath heeft dit vooral te maken met identiteitsontwikkeling: als jongeren een sterke identiteit en een eigen subcultuur kunnen ontwikkelen, zal dit een drijfveer zijn voor inspanningen zonder dat daar externe dwang voor nodig is. Echter, de ontwikkeling van deze eigen identiteit vereist weer wel een erkenning van het creatieve potentieel in jongeren en hun recht op zelfbeschikking en eigen keuzen. Voor het ontwikkelen van een eigen identiteit moeten jongeren vrij kunnen zijn en geen identiteit opgedrongen krijgen, zoals dat in traditionele opvoedingssituaties maar al te vaak gebeurt. Brice Heath heeft geprobeerd haar stellingen te onderbouwen met talrijke observaties van gedragingen en ontwikkelingen van jongeren in kunstclubs. Daar zag ze jongeren met een sterk ontwikkelde eigen identiteit, die wisten wat ze wilden en er ook zonder externe impulsen voor wilden werken. Het is moeilijk te overzien hoe algemeen deze ideeën zijn, een nadere analyse en reflectie zal zeker leerzaam zijn. En als ze inderdaad zo radicaal stand houden in de gevallen die Brice Heath beschrijft, is het nog moeilijk te overzien wat de precieze consequenties zijn van deze ideeën voor buitenschoolse leerprocessen, leerprocessen in gezinnen, leerprocessen in bedrijven etc.

Hoe het ook zij, het idee om leerlingen medezeggenschap te geven en te doen verwerven over hun eigen ontwikkelingsweg komt ook hier weer naar voren. Dit wijst in dezelfde richting als de studies waarin de mogelijkheden van communities of practice/communities of learners ten behoeve van het zinvolle en inzichtelijke leren worden onderzocht. Misschien is dit wel de algemene trend die uit al deze bijeenkomsten te distilleren valt en zelfs een trend waar we de komende jaren nog flink over na kunnen denken.

Divisie D: Measurement and Research
Methodology (R.R. Meijer, Universiteit Twente)

Binnen Divisie D en binnen de gelieerde NCME (National Council on Measurement in Education) bijeenkomst werd dit jaar veel aandacht besteed aan performance assessment en standard setting en, net als vorig jaar, adaptief testen per computer. Wanneer ik mij beperk tot het programma van de NCME dan waren er, wanneer ik goed heb geteld, 58 sessies waarvan 18 sessies waren gewijd aan performance assessment en 8 sessies te maken hadden met het afnemen van tests per computer. Een willekeurige greep uit de andere onderwerpen die werden belicht: nonparametrische item response theorie, geheimhouding van testopgaven, betrouwbaarheid response-tijden, en generaliseerbaarheidstheorie. Opvallend was verder de discussie over het gebruik van statistische toetsing bij de interpretatie van onderzoeksgegevens. Of beter gezegd de poging het gebruik van statistische toetsing af te schaffen. Een discussie die in heel sociaal wetenschappelijk Amerika woedt en waar op het ogenblik een werkgroep van de American Psychological Association (APA) mee belast is die, naar alle waarschijnlijkheid, met het advies zal komen het gebruik van statistische toetsing te beperken in APA tijdschriften. Binnen Divisie D was hieraan een speciale sessie gewijd met een algemene inleiding over de geschiedenis van statistische toetsing en een discussie door (voormalig) editors van de tijdschriften *Journal of Experimental Education* (Brice Thompson and Patricia Snyder), *Journal of Educational Measurement* (Rebecca Zwick) en *Educational and Psychological Measurement* (Larry Daniel). Onderwerpen die aan de orde kwamen waren onder andere het probleem dat statistisch significante resultaten lang niet altijd inhoudelijk betekenisvol te interpreteren zijn en het feit dat een statistisch significant resultaat weinig zegt over de herhaalbaarheid van de resultaten bij een iets gewijzigde steekproef-samenstelling. Tijdschriften als *Educational and Psychological Measurement* hebben hun beleid inzake statistische toetsing aangepast, zo werd duidelijk. Belangrijke richtlijnen zijn bijvoorbeeld dat, wanneer er statistisch wordt getoetst, een betrouwbaarheidsinterval en een effectgrootte moet worden opgegeven en dat,

wanneer enigszins mogelijk, replicatiegegevens moeten worden vermeld. Veel van deze aanbevelingen zijn niet nieuw maar zo oud als de statistische toetsing zelf. Een van de meest hilarische voorbeelden van een anti-toets houding is te vinden in een artikel van Louis Guttman uit 1982: 'Cyril Burt and the careless Star Worshipper'. Zwick beargumenteerde dat, hoewel zij een voorstander is van het vermijden van betrouwbaarheidsintervallen, maten van verklaarde variantie en effectgrootte, deze maten ook hun problemen kennen. Bijvoorbeeld, bij het opnemen van een R^2 statistiek voor de verklaarde variantie in een multiple regressie analyse moet worden bedacht dat deze coëfficiënt onzuiver is en dat de onzuiverheid toeneemt wanneer het aantal predictoren toeneemt bij gelijkblijvende steekproefgrootte. Ook betrouwbaarheidsintervallen zijn onderhevig aan dezelfde logica: net als bij hypothese toetsing worden bij het gebruik van betrouwbaarheidsintervallen willekeurige grensscores gebruikt, één van de argumenten die tegen hypothese toetsing worden gebruikt.

Wat betreft adaptief testen per computer was er onder andere een symposium georganiseerd door de Law School Admission Council, een instituut dat zich in de Verenigde Staten bezighoudt met het ontwikkelen van de tests die worden gebruikt voor de selectie van juridische studies. De volgende onderwerpen spelen op dit moment een belangrijke rol: het ontwikkelen van psychometrische maten voor het bepalen van de psychometrische kwaliteit van items en groepen van items, het ontwikkelen van testdesigns om tot zo efficiënt mogelijke manieren van iteraanbod te komen, en het onderzoek naar gelijkheid en eerlijkheid van de test voor verschillende raciale groepen. In dit symposium presenteerde zowel Wim van der Linden als Douglas Jones item selectie methoden die zijn gebaseerd op lineair programmeren. Bert Green presenteerde een alternatieve methode om tot inzichtelijke scoring van de latente vaardigheid te komen en Rebecca Zwick presenteerde een Bayesiaanse methode voor itembias-analyse.

Een onderwerp dat in opkomst lijkt is het onderzoek naar response-tijden in CAT. Psychometrici hebben van oudsher vooral veel aandacht besteed aan de nauwkeurigheid van een meting. Met de opkomst van CAT kunnen

response-tijden worden geregistreerd en dit kan wellicht additionele informatie opleveren over de vaardigheid van een persoon. Scrams en Schnipke deden verslag van onderzoek naar de relatie tussen nauwkeurigheid en snelheid in CAT, waarbij werd geopperd de snelheidsfactor mee te wegen bij het schatten van de vaardigheid. Bontempo en Julian onderzochten in hoeverre de hoeveelheid tijd die een respondent ter beschikking had voor het beantwoorden van de vragen van invloed was op de snelheid waarmee werd geantwoord. Deze bleek, niet onverwacht, van groot belang te zijn.

En dan was er veel, heel veel aandacht voor performance assessment en standard setting. Bij performance assessment gaat het om onderwijskundig evaluatie-onderzoek van vaak complexe verrichtingen en het bepalen van minimale eisen waaraan deze verrichtingen moeten voldoen (standard setting). Een voorbeeld is het in kaart brengen van de kwaliteit van leraren en het bepalen aan welke minimale eisen leraren dienen te voldoen. Een lezing van Jean Miller leerde ons dat leraren worden beoordeeld op de mate waarin zij 'leerstofgericht zijn', 'zich bewust zijn van verschillende culturen', 'kennis hebben van de inhoud van het vak' en 'pedagogisch coherent zijn'. Andere lezingen omtrent dit onderwerp leerde ons dat de Angoff methode (ter bepaling van grensscores) nooit verloren zal gaan, maar wel een jaren negentig jasje heeft gekregen: De Angoff methode voor meerdimensionele gedragsbeoordelingen. Verder was er veel aandacht voor betrouwbaarheid en validiteit van dit, in methodologisch opzicht, lastige onderwerp.

Tot slot mag niet ongenoemd blijven dat er een dag was uitgetrokken voor een cursus 'Using humor in the classroom and in professional presentations'. Ik ben er echter helaas niet geweest.

Divisie G: Social Context of Education

(H.P.J.M. Dekkers, ITS/Katholieke Universiteit Nijmegen)

Divisie G omvat onderzoek naar de relatie tussen het onderwijs en zijn sociale, politieke, culturele en economische context. De relatie tussen de sociaal-economische herkomst van leerlingen, hun etnische achtergrond en hun sekse met onderwijsloopbanen vormen tradi-

tioneel belangrijke issues binnen deze divisie. Multi-cultureel onderwijs en de relatie school-community waren dit jaar programma highlights, waarbij thema's aan de orde kwamen als tweetalig onderwijs en de ontwikkeling van lerarenopleidingen waarin aandacht is voor het omgaan met maatschappelijke verschillen tussen leerlingen. De titel van de als highlight aangekondigde sessie 'Making a difference in children's lives: Stories from Administrators, Community Activists, Researchers, and Teachers' geeft een beeld van de aard van een aanzienlijk deel van de sessies uit deze divisie. Een groot aantal papers omvatte eerder filosofische/politieke bijdragen en/of case studies van zeer beperkte omvang dan 'hardere' resultaten van empirisch onderzoek. Dat neemt niet weg dat er voldoende interessante lezingen en sessies waren voor degenen die behoefte hadden aan meer generaliseerbare resultaten van onderzoek. Bovendien was het mogelijk om over onderzoek dat qua onderwerp past in deze divisie, informatie te genereren in sessies van andere divisies (of Special Interest Groups), zoals Divisie H: School Evaluation and Program Development en de nieuwe Divisie L: Education Politics and Policy. In de eerste bieden bijvoorbeeld lezingen waarin school-effectiviteit in relatie wordt gebracht met het realiseren van niet alleen 'excellence' maar ook 'equity' soelaas. In Divisie L zijn sessies over (de)centralisatie en lokaal onderwijsbeleid interessant in het kader van de decentralisatie van het Nederlandse Onderwijsachterstandsbeleid (Lokaal Beleid).

De relatie tussen de sociaal-economische en vooral etnische achtergrond van leerlingen en hun onderwijsprestaties werd zoals gezegd vaak gepresenteerd in verslagen van kleinschalig onderzoek en soms politiek geïnspireerde lezingen. Bovendien zijn bevindingen ten aanzien van Hispanics in Philadelphia niet meteen te transformeren naar belangwekkende inzichten ten aanzien van de problematiek van culturele minderheden in Nederland. Daar staat tegenover dat het zeer zinnig is kennis te nemen van de wijze waarop de bekende 'Success for All'-inzichten van Slavin in diverse landen (Canada, Mexico, Israël, Australië) worden geïmplementeerd, en dat de sessie waarin werd gediscussieerd over de implementatie en evaluatie van dergelijke programma's,

ondersteund door metingen van leerlingresultaten, zeer leerzaam was. Opmerkelijk is dat terwijl in dergelijke sessies de evaluatie van compenserende programma's voor achterstandsleerlingen aan de orde is, in andere met grote nieuwsaarde wordt vastgesteld wat maatschappelijke risicoleerlingen zijn en hoe groot hun achterstand in het onderwijs is. Dit gebeurt momenteel vooral aan de hand van de grote landelijke databases zoals de National Education Longitudinal Study (NELS:88), High School and Beyond (HS&B) en National Assessment of Educational Progress (NAEP). Deze nationale cohorten zijn niet zoals in Nederland in eerste instantie in het leven geroepen om achterstandsbeleid te evalueren, maar door oversampling van bepaalde groepen (en soms ook staten) zijn toch analyses mogelijk die achterstanden kunnen blootleggen (Peng e.a.). De oorzaken van de onderwijsachterstanden worden in dit soort studies gezocht in de leeromgeving thuis en in ervaringen op school, alles geanalyseerd aan de hand van indicatoren uit de cohorten. Op deze bestanden zijn ook dropout studies uitgevoerd, waarin wordt geconcludeerd dat voortijdig schoolverlaten significant afneemt, terwijl het armoedenniveau niet is veranderd. De nadruk wordt gelegd op de rol die de school hierbij speelt; weliswaar zijn de resultaten door de spreiding van de grade 8 leerlingen over vele scholen van voortgezet onderwijs niet meer representatief op schoolniveau, maar outlierstudies wijzen uit dat de school ertoe doet. Opmerkelijk is de nadruk die in deze (kwantitatieve) sessies wordt gelegd op het belang van kwalitatieve studies die de context moeten leveren voor de kwantitatieve data; men onderstreept het belang van inzicht in integrale cases, zij het geruggesteund door kwantitatieve data.

Verschillen in het niveau van probleemdefiniëring doen zich ook voor in het genderonderzoek. Enerzijds wordt in bepaalde sessies nog verongelijkt vastgesteld dat jongens en meisjes in het onderwijs anders worden bejegend (How schools shortchange girls), anderzijds wordt⁶ sophisticated onderzoek gepresenteerd over genderverschillen in schoolloopbanen, met name op het gebied van exacte vakken, en buigt een Pilot Expert Panel zich over de meest veelbelovende programma's ter bevordering van Gender Equity in het

onderwijs (en over de evaluaties en verdere disseminatie daarvan). Met betrekking tot de exacte vakken problematiek wordt in sommige presentaties de nadruk gelegd op de domeinen waarin vrouwen eerder een voorsprong hebben dan een achterstand (biologie, scheikunde; o.a. Lee). Daarnaast worden nog steeds variabelen onderzocht die in het bekende model ter verklaring van seksespecifieke keuzen van Eccles passen, zoals belangstelling, doeloriëntatie en nutinschatting. Verder wordt het onderwerp steeds meer bestudeerd voor de combinatie sekse en etniciteit. Eccles en Thome pleiten ervoor de wat zij noemen te categoriale benadering (too separate, too mono, too strict, too homogenous) te vervangen door een meer integrale benadering. Eccles analyseert momenteel vooral de interacties en komt tot conclusies over de achterstanden van allochtone jongens die ook in Nederland gelden. Ook Ainley onderzoekt de seksespecifieke deelname aan mathematics in Australië in relatie tot de 'home and ethnic' achtergrond. Zeer verfrissend in dit kader was de bijdrage van een aantal jonge UCLA Graduate School of Education medewerkers (o.a. Walpole en McClafferty), die, niet gehinderd door vroegere onderzoekstradities, trachten te definiëren welke leerlingen nu precies at risk zijn. Als vanzelf leveren hun historische, methodologische en beleidsstudies een pleidooi op voor het overschrijden van de disciplinaire grenzen (psychologie, sociologie) en van de monoverklaringen. Class, gender en ethnicity moeten geïntegreerd worden bestudeerd, en Bourdieu wordt niet alleen ter verklaring van klasseverschillen, maar ook van etnische en sekseverschillen van stal gehaald. Dit maakt nieuwsgierig naar hun bijdragen op een volgende AERA.

Tenslotte waren enkele sessies over decentralisatie/lokaal beleid relevant voor de relatie context/beleid-onderwijs-ongelijkheid. Met name de discussies over de State of Chicago School Reform, o.a. in de sessie 'Has local democratic decisionmaking in Chicago led to improved schooling?' was interessant. Het nieuwe onderwijsbeleid kent vijf essentiële onderdelen. De belangrijkste vernieuwing lijken de gekozen Local School Councils te zijn, waardoor idealiter de verantwoordelijkheid voor het onderwijs in handen van de community ligt. Naast bestuurders en leerkrachten moe-

ten ouders daarin een belangrijke rol vervullen, een rol die blijkens de discussies nog niet volledig tot zijn recht komt. Verder ligt de nadruk op professionele ontwikkeling en teamwerk van leerkrachten, op betrokkenheid van (alle) ouders en de community, op het schoolleiderschap (het is moeilijk een balans te vinden tussen de autonomie van de schoolleider en de rol van het centrale bestuur), en op leerlinggericht en omgevinggericht leren. Ten aanzien van het laatste onderwerp wordt vastgesteld dat deze principes makkelijker zijn te realiseren in het basisonderwijs dan in het voortgezet onderwijs. Verder blijkt het moeilijk een evenwicht te vinden tussen wat leerlingen in ieder geval moeten weten/kunnen/kennen en de link met 'what goes on in the community'. Opmerkelijk tenslotte is het verschijnsel 'Schools on probation': scholen die minder dan 15% van de leerlingen op een bepaald niveau afleveren worden tijdelijk onder curatele gesteld. Uiteraard roept dit discussies op over het feit dat dergelijke beslissingen zijn gebaseerd op testcores en niet op ontwikkeling van andere (sociale/emotionele) vaardigheden en gedrag van de leerlingen. Al met al een interessante discussie over gedecentraliseerd (achterstands)beleid, die in een andere sessie over decentralisatie en onderwijsvormingen (Australië, Engeland en VS/Wisconsin/Madison) een zinnige vierde bijdrage had kunnen leveren. In deze laatste hier besproken sessie was het effect van de combinatie van centrale standaarden en een decentrale invulling en management van het onderwijs (inclusief concurrentie) op achterstandsléerlingen een belangrijk onderwerp. De Engelsen maakten expliciet gewag van negatieve gevolgen voor de 'disadvantaged'. In geen enkel land is decentralisatie specifiek als achterstandsbeleid geprofileerd, zoals in het Nederlandse Gemeentelijke onderwijsachterstandsbeleid. Verder werd in de sessie gewezen op het ontbreken van theorieën over geschikte lokale structuren voor gedecentraliseerd onderwijsbeleid. Tenslotte werd besproken dat het onderwijs zodanig aan het veranderen is dat het niet voldoende is alleen binnenschools effectiviteitsonderzoek te doen, maar dat het tijd is voor meer omvattende conceptualisaties, waarin de rol van het beleid en de context van het onderwijs in onderzoek worden betrokken. Dit leverde o.a. een pleidooi op voor 'value added out-

comes', dus onderwijsresultaten gecontroleerd voor input, iets wat in Nederland in het ongelijkheidsonderzoek al gemeengoed is. Dergelijke gedachten vormen een geschikte afsluiting van een verslag van Divisie G, waarin enkele sessies over schooleffectiviteit (Divisie H) en beleid (Divisie L) zijn opgenomen.

Papers

Ainley, J. *Student participation in math courses in Australian secondary schools.*

Eccles, J. *Ethnicity and identity in the classroom and on the playground.*

Lee, V. *Persistence in the science pipeline.*

Peng, S.S. *Understanding achievement gaps of disadvantaged students: Major findings from studies using national databases.*

Slavin, R. *Success for all: quality matters.*

Thorne, B. *The criss-crossing of gender, 'race', ethnicity, social class and age in daily school life.*

Walpole, M. *Risk from the sociocultural perspective: examining structural variables in the definition of risk.*

Divisie H: School evaluation and program development (B.P.M. Creemers, Rijksuniversiteit Groningen)

AERA-divisie H is een divisie waarin praktische en beleidsmakers enerzijds en onderzoeksgerichte AERA-leden anderzijds elkaar ontmoeten. Het gaat bij programma-ontwikkeling en evaluatie vaak om lokale ontwikkelingen in een groep van scholen of in het schooldistrict en soms om het niveau van de staat in programma-ontwikkeling. Bovendien zijn vaak ook enthousiaste mensen vanuit de praktijk betrokken die in AERA-presentaties naar voren worden geschoven om de stem uit de praktijk te laten horen. Van oudsher was het ook de divisie waarin onderwijs-effectiviteitsonderzoek en vooral onderwijs-effectiviteitsonderzoek in relatie tot onderwijsverbetering aan de orde kwam. Daarnaast komt onderwijs-effectiviteitsonderzoek ook aan de orde in de Special Interest Group (SIG) 'School Effectiveness and School Improvement' en aanverwante SIG's, zoals 'School Indicators and Proficiencies', 'Study of Learning Environments' en 'Restructuring Public Education' en had de International Congress of School Effectiveness and

School Improvement (ICSEI) van de AERA ruimte in het programma.

De afgelopen jaren heeft het onderwijs-effectiviteitsonderzoek op het AERA-congres het vooral moeten stellen met presentaties binnen de SIG 'School Effectiveness en School Improvement' en het ICSEI-onderdeel. Maar nu het goede nieuws: onderwijs-effectiviteit is terug op de agenda in Amerika en in het onderwijsonderzoek dat in Amerika wordt uitgevoerd. Ook in andere divisies zoals 'Teaching and Teacher Education' en nog meer in Divisie A, 'Administration', werden dit jaar presentaties op het terrein van onderwijs-effectiviteit gegeven of werd in elk geval de relatie tussen organisatievariabelen, administratieve maatregelen, leerkrachtgedrag en effectiviteit neergelegd. Verder wordt bij standaarden, onderwijskundig leiderschap en schoolverbetering meer dan in het verleden de relatie met opbrengsten en onderwijs-effectiviteit gelegd.

Ook in de programma-evaluatie was meer aandacht voor de effecten op het niveau van leerlingen. Deze onderdelen van het programma werden ook, tenminste dat is de indruk die ik kreeg, beter bezocht dan in het verleden. Nu maar hopen dat theorievorming en onderzoek op het terrein van onderwijs-effectiviteit en schoolverbetering de verwachtingen van het publiek waar maken. Reden te meer om uit te zien naar de publicatie van de twee overzichtswerken op het terrein van onderwijs-effectiviteit waarvoor met verve op de AERA door de auteurs en editors reclame werd gemaakt.

Door Bosker en Witziers werden de resultaten van een meta-analyse van vierhonderd onderwijs-effectiviteitsstudies gepresenteerd. Er blijven uit de overzichten van Lefine en Lezotte en van Sammons e.a. die tot een opsomming van elf groepen van factoren kwamen, nog genoeg factoren over die een redelijke effect-grootte hebben en vandaar van belang zijn in het kader van onderwijs-effectiviteit en schoolverbetering. Alarmerend was dat aangekondigd werd dat de benadering die volgens de 'Best Evidence Approach' uitgevoerd was tot andere conclusies kwam. In z'n commentaar probeerde Sam Stringfield duidelijk te maken dat ook al waren indertijd door Edmonds de vijf factoren min of meer uit de lucht geplukt, de meerderheid van de nieuwe indelingen uiteindelijk op deze vijf factoren is terug te voeren en klaar-

blijklijk via de meta-analyse van de vierhonderd onderzoeken toch, voorlopig althans, ondersteuning krijgt.

In een sessie over scaling-up van restructureringsprogramma's (bijna allemaal gerelateerd aan de effectieve schoolbeweging, zoals 'Success for All', 'Roots and Wings' en de New American School Design) werd een duidelijke relatie gelegd tussen effectiviteit bevorderende factoren en de verdere verspreiding van de programma's. Daarbij was de aandacht vooral gericht op factoren binnen de programma's die deze verspreiding zouden kunnen bevorderen zoals in Amerika sterk onderwijskundig leiderschap, de strikte programmastructuur, de lokale ondersteuning vanuit het beleid en het hantieren van hoge standaarden voor alle studenten en het vervolgens daarop richten van het onderwijsgedrag in klassen, secties en school. In de discussie werd er op gewezen dat de steekproeven soms wel erg klein waren. Verder is er een beleidsprobleem, nl. de discrepantie tussen deze programma's die gericht zijn op het veranderen van het onderwijs en de evaluatieprogramma's die in diverse staten worden uitgevoerd. Tenslotte blijft causaliteit bij scaling-up, zeker in relatie tot de geringe omvang van de steekproeven, een probleem. Een punt waarop gewezen werd, is dat in deze schaalvergrotingsprogramma's te weinig aandacht is voor het instructieniveau waarvan uit onderzoek blijkt dat dit niveau uiteindelijk ook de kern van onderwijsverbetering vormt. In dit verband werd dan ook gevraagd om een verdere conceptualisering van scaling-up om te vermijden dat uiteindelijk een schaalvergroting van dergelijke programma's zou leiden tot scaling around. In deze sessie was duidelijk dat het in onderwijsverbetering en restructuring gaat om onderwijs-effecten en de factoren die bijdragen tot het bereiken van deze effecten. Het kan echter ook anders, zoals bleek uit een sessie over strategieën voor het ontwikkelen van effectieve scholen in Maryland. Nadat de steekproef op basis van kwantitatieve gegevens getrokken was, kwam er verder geen feitelijk materiaal meer aan te pas. Nadat de selectie van de scholen (effectief en niet-effectief voor verschillende SES-strata) had plaats gevonden, gingen interpretatieve onderzoekers op stap en constateerden allerlei verschillen tussen effectieve en niet-effectieve scholen.

Voorstanders van kwalitatief onderzoek zijn binnen onderwijseffectiviteit niet alleen groot in aantal maar hebben ook veel invloed. Dat blijkt ook uit reacties op de kwantitatieve bijdragen en presentaties die er vele waren, zeker uit de internationale hoek. Telkens werd dan de opmerking gemaakt dat case-studies en meer kwalitatieve studies nodig waren. Presentaties zoals die vanuit Maryland maken naar mijn mening duidelijk dat er altijd sprake moet zijn van mixed methodologie waarbij aan kwalitatieve gegevens en uitspraken en beschrijvingen kwantitatieve gegevens gekoppeld worden. Dat moet niet alleen op het niveau van de uitkomsten, de resultaten van leerlingen in het onderwijs gebeuren (hetgeen ook in dit geval bij de sampling wel gebeurd was), maar ook bij de factoren, de processen die in het interpretatieve onderzoek als middelen voor effectiviteit worden aangemerkt.

Geverdt en Chow-Hoy kwamen in een multilevel-analyse van de NELS-data tot de conclusie dat er bij de reductie van de grootte van klassen een niet-lineair effect was. Het effect was het grootst bij kleine en bij grote klassen maar niet aanwezig bij middelgrote klassen met tussen de 20 en 30 leerlingen. In deze sessie werden als vervolg op een paper uit 1996 over hetzelfde onderwerp, de resultaten gepresenteerd van een studie naar blokprogrammering. Blokprogrammering houdt in het indelen van onderwijs in grotere eenheden gedurende de dag. De resultaten in high-schools waren positief. In een interessante bijdrage van Young werden de eerste resultaten van een longitudinale studie naar onderwijs-effectiviteit in landelijke gebieden in Australië gepresenteerd. Landelijk betekent in dit verband ver van steden gelegen en veelal kleine scholen. Voor zowel natuurkunde als voor wiskunde bleek er een sterk effect van de klas te zijn en een gering school-effect. Op leerling-niveau was het zelfconcept van groot belang voor het behalen van effecten (of is het een resultaat?). In die richting zal het onderzoek verder worden uitgewerkt. Een aardige bijkomstigheid van deze presentatie was dat ik leerde dat het model voor onderwijs-effectiviteit eigenlijk een 'simpel input-output model' was, hoewel onderzoekers op dit terrein zich jaren druk hebben gemaakt over condities, processen en de context.

De presentaties binnen de sessie van het International Congress of School Effectiveness and School Improvement maakten in elk geval duidelijk dat onderwijs-effectiviteitsonderzoek geen exclusieve Amerikaanse zaak is, maar in een groot aantal landen plaatsvindt. Dat effectiviteitsonderzoek en belangrijke bijdrage (toegevoegde waarde) kan hebben bij programma-evaluatie werd duidelijk in een sessie over de internationale verspreiding van 'Success for All', een programma voor lees- en taalonderwijs ontwikkeld door Slavin, later uitgebreid in 'Roots and Wings' naar andere schoolvakken en leerjaren. Temidden van de enthousiaste maar weinig empirische verhalen over de grote waarde van het programma was de presentatie van Crévola en Hill een verademing. Maar ook zij konden tot positieve resultaten komen op basis van de resultaten van de (multi-level)analyse van de gegevens. Schooleffectiviteitsonderzoek staat, zoals geconstateerd, weer in de belangstelling. Nog belangrijker is, mede op basis van de beschikbare kennis over onderwijs-effectiviteit, de doorwerking op andere terreinen van onderwijsonderzoek. Hopelijk iets blijvends.

Divisie J: Postsecondary Education

(I. Bakkenes, ICLON, Rijksuniversiteit Leiden)

In een speciaal aanbevolen divisie J-bijeenkomst keek David Breneman, financieel econoom en filosoof, 'through a clouded crystal ball' naar de toekomst van het hoger onderwijs. In deze sessie werd het hoger onderwijs door hem beschreven als een zeer complex en moeilijk te voorspellen systeem. Onderzoek van het hoger onderwijs zou daardoor niet accumulatief, en al snel niet meer relevant zijn. De meeste invloed gaat volgens Breneman dan ook uit van in het veld en tijdens conferenties gepresenteerde visies, en niet van gepubliceerd onderzoek. Of dit nu gezien moet worden als de waarheid of als een zwak excuus, feit is dat in een groot deel van de bijgewoonde sessies veeleer praktijkbeschrijvingen en op de praktijk gebaseerde toekomstvisies werden gepresenteerd dan op onderzoek gebaseerde inzichten.

Het merendeel van de divisie J-sessies kon onder drie thema's worden geplaatst. Het eer-

ste thema was de productiviteit van medewerkers van colleges en universiteiten. Het tweede thema betrof de effectiviteit van docenten en het onderwijs. De toekomst van het hoger onderwijs en het te voeren beleid vormde het derde te onderscheiden thema. Op elk van deze thema's zal kort worden ingegaan.

De productiviteit van medewerkers werd in de bijgewoonde sessies beschouwd in het kader van institutionele verantwoording, fondsenwerving en personeelsbeleid. Bij het bepalen van de productiviteit en het toekennen van fondsen wordt in Amerika niet alleen gelet op publicaties en onderwijsresultaten maar ook steeds meer op de mate waarin 'publieke diensten' worden verleend. Medewerkers van colleges en universiteiten dienen aan 'professional service' te doen. Op deze manier verantwoord worden instituten zich naar hun directe omgeving. Verantwoording (accountability) vindt ook plaats naar de federale overheid. Zo worden er landelijke monitor- en evaluatiesystemen opgezet waarmee jaarlijks diverse gestandaardiseerde gegevens worden verzameld over het functioneren van promovendi die, binnen een bepaald programma, directe subsidie ontvangen. Gegevens worden opgeslagen in grote databases en openbaar gemaakt in gestandaardiseerde rapporten. Federale managers moeten op deze manier hun investeringen aan de overheid en de belastingbetaler verantwoorden. Een promovendus wordt daarmee als een 'public good' gezien.

Een ander regelmatig terugkerend onderwerp binnen het eerste thema was de vaste aanstelling (tenure-status) van medewerkers. Voorstanders verdedigden de vaste aanstelling met het argument van de academische vrijheid. Tegenstanders waren van mening dat de vele vaste aanstellingen tot stilstand in ontwikkeling en een te lage productiviteit leiden. De discussie kent nog geen duidelijke uitkomst. Feit is dat het aantal medewerkers zonder vaste aanstelling steeds verder groeit. De arbeidsvoorwaarden en de mogelijkheden voor deze groep medewerkers blijven echter een stuk slechter dan die van mensen met een vaste aanstelling. Dit leidt volgens velen tot motivatieproblemen en het verlies van jonge, interessante mensen. Gezocht zou moeten worden naar meer directe beloningsstructuren die het werken op een tijdelijke aanstelling aantrekkelijker kunnen

maken.

Het tweede thema was de effectiviteit van docenten en het onderwijs. Onder dit thema vielen bijvoorbeeld sessies over factoren die succes en uitval van studenten beïnvloeden, over kenmerken van effectieve docenten en over de 'reflective practitioner'. De kwaliteit van de bijgewoonde sessies viel echter tegen. Potentieel interessante onderzoeksresultaten met betrekking tot bronnen voor professionele ontwikkeling werden bijvoorbeeld ontkracht doordat alleen gekeken was naar de mate waarin docenten van bronnen gebruik maken en niet naar de beschikbaarheid en kwaliteit van deze bronnen.

Boven de toekomst van het hoger onderwijs, het derde thema, hangt de donkere wolk van teruglopende financiën. De aandacht van de Amerikaanse regering richt zich nu vooral op kinderen tot 14 jaar. Het is onduidelijk of er sprake is van een doorgaande lijn op dit punt. Wanneer deze lijn zich doorzet wordt het hoger onderwijs wellicht duurder voor studenten en men maakt zich daarom zorgen over de toegankelijkheid van het onderwijs. Met betrekking tot het onderzoek bestaat de zorg dat men steeds meer gebonden zal zijn aan bepaalde thema's in verband met het oormerken van subsidies. Wellicht zal er ook sprake zijn van een steeds groter wordende tegenstelling tussen de publieke en de private sector. De vraag is hoe kosten, kwaliteit en toegankelijkheid zich tot elkaar gaan verhouden.

Rest mij nog een korte blik op de toekomst van het onderzoek naar het hoger onderwijs. Volgens D. Breneman wordt het hoger onderwijs op dit moment te veel in horizontale plakjes bestudeerd. Er zijn maar weinig onderzoeken die de verticale lijn van K-12 naar het hoger onderwijs bestuderen. Hier zou een belangrijke taak voor toekomstig onderzoek liggen.

Divisie K: Teaching and Teacher Education (D. Beijaard, ICLON, Rijksuniversiteit Leiden)

Opvallend binnen Divisie K was de grote hoeveelheid onderwerpen waaruit men kon kiezen. Naar verhouding vonden veel sessies plaats over de kennis en 'beliefs' van leraren, de betekenis van kritische reflectie door leraren en de invloed van de institutionele context op het

handelen in de onderwijspraktijk van leraren-opleiders en leraren. Daarnaast werd in veel sessies ingegaan op de beoordeling of evaluatie van leraren. Met name deze laatste sessies waren doorgaans goed georganiseerd: gelijkgezinde mensen die aansluitend op elkaar erin slaagden diepgaand op een gezamenlijk gekozen problematiek in te gaan. Vanwege de relevantie van het beoordelen of evalueren van leraren (in de Verenigde Staten) wordt daaraan in dit verslag de meeste aandacht besteed. Allereerst wordt echter kort ingegaan op de eerstgenoemde onderwerpen.

Interessant waren enkele symposia over 'beliefs' van leraren-in-opleiding. Door Marble werden 'beliefs' gedefinieerd als opvattingen van een individu die voor hem of haar actueel en waar zijn. Deze 'beliefs' kunnen irrationeel zijn wanneer zij in iemands omgeving niet empirisch worden ondersteund. Uit diverse presentaties bleek dat 'beliefs' gewenste veranderingen in het onderwijs ernstig kunnen blokkeren. Het bewerkstelligen van dissonantie bij leraren (in opleiding) en het creëren van ruimte om te kunnen werken aan 'new ways of knowing' zijn onder andere mogelijkheden om 'beliefs' van leraren te veranderen.

'Beliefs' van leraren-in-opleiding komen vaak tot uiting tijdens conflicten en spanningen die men tijdens de opleiding ervaart. Uit het onderzoek van Beach kwam onder meer naar voren dat deze conflicten en spanningen gedurende de opleiding toenemen (bijvoorbeeld schooldoelstellingen waar men niet achter staat, of aardig gevonden willen worden maar dat niet kunnen realiseren). Oorzaken voor dergelijke conflicten en spanningen schrijft men na verloop van tijd steeds meer toe aan zichzelf. In het algemeen hanteren leraren-in-opleiding drie niveaus van strategieën voor het omgaan met conflicten en spanningen: het eerste niveau komt neer op ontkenning van het conflict of de spanning, het tweede niveau omvat het zoeken naar oplossingen op de korte termijn en waarmee vooral het eigenbelang gediend wordt, het derde niveau bestrijkt reflectie op onderwijstheorieën en daarmee het zoeken naar oplossingen op lange termijn. De door Beach gepresenteerde onderzoeksresultaten dragen bij tot mogelijkheden om de relatie tussen instituutactiviteiten en in de praktijk opgedane ervaringen te verbeteren.

Kritische reflectie door leraren-in-opleiding wordt door iedereen belangrijk gevonden, maar nagenoeg niemand blijkt te weten hoe dat vorm te geven. Kritische reflectie houdt in dat iemand nadenkt over de politieke, sociale, morele en ethische implicaties van diens handelen. Uit talrijke, vooral kwalitatieve studies blijkt dat leraren-in-opleiding wel reflecteren over technische en praktische aspecten van hun handelen, maar dat geen kritische reflectie plaatsvindt. Deze vorm van reflectie komt ook niet na verloop van tijd tot stand. Tijdens diverse presentaties werd ervoor gepleit kritische reflectie expliciet als onderdeel in het onderwijsprogramma van de lerarenopleiding op te nemen.

Onder leiding van Knowles was een symposium georganiseerd over de invloed van de institutionele context van lerarenopleidingen op het handelen van de daarin werkzame opleiders. De onderzoeken die in dit symposium werden gepresenteerd waren flink empirisch onderbouwd. In de Verenigde Staten blijkt van lerarenopleidingen een grote conserverende werking uit te gaan, waardoor zij een gebrekkige impact hebben op studenten. Lerarenopleiders worden in hun instituten constant geconfronteerd met onzekerheid, terwijl ze wel leraren moeten opleiden; deze situatie is niet goed voor de opleiders en dus ook niet voor de studenten. Met name de universitaire opleidingscontext kreeg er flink van langs. In deze context hebben de noodzaak om je constant te moeten bewijzen en het huidige beloningssysteem een negatieve invloed op het functioneren van opleiders. De context van lerarenopleidingen staat soms in schril contrast met die van de scholen waarin leraren-in-opleiding hun stages volbrengen en met het belang dat lerarenopleidingen toekennen aan de context of cultuur van een school voor het leren onderwijzen. Uit diverse studies bleek dat met name scholen die kenmerken van een lerende organisatie vertonen een zeer gunstige uitwerking hebben op het leerproces van leraren-in-opleiding.

De bijgewoonde sessies over het beoordelen of evalueren van docenten waren meestal zeer de moeite waard. De algemene tendens is dat het proces dat aan de beoordeling ten grondslag ligt belangrijker is dan het resultaat ervan. Zo'n proces maakt leraren bewust van hun activiteiten waardoor zij zichzelf nieuwe doelen kun-

nen stellen. Beoordelen is dus geen kille afrekening, ook niet wanneer strikt wordt uitgegaan van een rigoureuze toepassing van strak omlijnde 'teaching standards'. Centraal staat het zoeken naar essentiële componenten van een beoordelingsproces dat evaluatie en professionele groei met elkaar verbindt. Dit accent op professionele ontwikkeling vloeit voort uit de tijdens het congres veel gehoorde kritiek op beoordelingspraktijken, waarbij docenten een score wordt toegekend zonder of slechts met minimale feedback, terwijl de onderwerping aan een beoordeling een docent ongelooflijk veel tijd (en geld!) kost.

Kritiek op een dergelijke beoordelingspraktijk vloeit ook voort uit recente opvattingen over het onderwijzen zelf: het beoordelen van leraren is zo complex en dynamisch als het domein van het onderwijzen zelf, dat zich bovendien kenmerkt door verschillende werkcontexten en beroepsinhouden (vgl. bijvoorbeeld een leraar in de basisschool die van veel vakgebieden kennis moet hebben en vooral gericht is op de ontwikkeling van de leerling met een docent in de bovenbouw van het voortgezet onderwijs die zich heeft gespecialiseerd in één vakgebied en vooral gericht is op beheersing van vakspecifieke kennis en vaardigheden door leerlingen). Kortom, veel valt niet te beoordelen omdat we niet weten op welke wijze dat zou moeten, veel is daarentegen ook wel te beoordelen. Essentieel is dat te beoordelen taken adequate 'samples' zijn die het domein vertegenwoordigen. Daarbij moet in het oog worden gehouden dat naarmate we meer van een docent beoordelen we niet automatisch ook beter beoordelen en dat per definitie de menselijk cognitie beperkt is. Duidelijk is dat we hier met een validiteitsprobleem te maken hebben.

Tegen bovenstaande achtergrond wordt in de Verenigde Staten naarstig gezocht naar alternatieve manieren van beoordeling van leraren. Bij de beoordeling of evaluatie van leraren vanuit een professioneel ontwikkelingsmodel is volgens Wolf het portfolio een zeer geschikt instrument, vooropgesteld dat daarin de eigen praktijk niet zozeer wordt beschreven maar geanalyseerd. Het gaat dan vooral om het realiseren van leerdoelen. Voor strikte evaluatiedoelinden is het raadzaam de bezwaren die aan portfolio's kleven te onder-

vangen met andere methoden. In dit verband werd ook gewezen op het belang van 'client surveys' als betrouwbare databronnen. Volgens diverse onderzoekers past het gebruik van deze databronnen in het huidige klimaat waarin het leraarsberoep zich bevindt (groter accent op leren, toegenomen klantgerichtheid, meer aandacht voor professionele ontwikkeling e.d.). Ellet, tenslotte, pleitte voor een wel zeer drastische wijziging in de huidige opvattingen over het beoordelen van docenten. Zijn standpunt komt er in het kort op neer dat bij het beoordelen van leraren niet moet worden uitgegaan van het onderwijzen maar van het leren en dat het een illusie is standaarden op klassen los te kunnen laten. Kortom, inzake de beoordeling of evaluatie van leraren is sprake van een toegenomen verscheidenheid in opvattingen. De achterliggende boodschap getuigt echter van eensgezindheid: een rigoureuze toekenning van scores aan leraren heeft geen zin, met de situatie waarin de leraar zich bevindt moet rekening worden gehouden, beoordeling moet plaatsvinden op basis van voor het domein representatieve 'samples', de beoordeling is bij voorkeur gebaseerd op gegevens die zijn ontleend aan meerdere instrumenten, en aan de feedback die op basis van de beoordeling tot stand komt moet veel zorg en aandacht worden besteed.

De noodzaak tot het definiëren van de kern van het leraarsberoep en het beoordelen van leraarsbekwaamheden werd nog eens onderstreept in een symposium naar aanleiding van het recent verschenen rapport van de National Commission on Teaching & America's Future. Dit rapport is getiteld 'What matters most: Teaching for America's future'. In het mede op basis van onderzoek tot stand gekomen rapport maakt men zich grote zorgen over de kwaliteit van veel leraren en lerarenopleidingen. Hoewel sprake is van grote verschillen tussen staten, is er in de Verenigde Staten geen sprake van een werkelijk systeem op basis waarvan leraren worden gerecruteerd en voorbereid op hun beroep en op basis waarvan zij zich verder kunnen professionaliseren. De aanbevelingen die de genoemde commissie doet liggen in het verlengde van deze constatering. Veel aandacht gaat daarbij uit naar het accrediteren van lerarenopleidingen (gebaseerd op het werk van de National Council for Accreditation of Teacher

Education, NCATE), het verstrekken van licenties aan beginnende leraren (gebaseerd op het werk van de Interstate New Teacher Assessment and Support Consortium, INTASC) en het certificeren van reeds in het onderwijs werkzame leraren (gebaseerd op het werk van de National Board for Professional Teaching Standards, NBPTS). In toenemende mate wordt het belangrijk gevonden dat de opleiding en verdere professionalisering van leraren zijn gebaseerd op gedeelde opvattingen over voor leraren essentiële kennis, vaardigheden en 'commitments'. In het algemeen blijkt hier geen weerstand tegen te bestaan, wel tegen een gedetailleerde opsomming van wat leraren moeten kennen en kunnen en de vertaling daarvan in een simpel scoringssysteem aan de hand waarvan leraren beoordeeld worden.

Aan deze kroniek werkten mee: *I. Bakkenes, D. Beijaard, B.P.M. Creemers, H.P.J.M. Dekkers, M.L. Krüger, R.R. Meijer, B. van Oers, J. Terwel en N. Verloop*

De eindredactie werd verzorgd door *N. Verloop*

Boekbesprekingen

S.P.M. Nossent

Een beweeglijke psyche. Over epigenese bij baby's

Academisch proefschrift Universiteit Utrecht. Uitgeverij Kok Agora, Kampen, 1995, 363 pagina's, f 59,50, ISBN 90-391-0644-4.

'In dit boek wordt een fundamentele vraag in de ontwikkelingspsychologie behandeld'. Zo begint de samenvatting van dit proefschrift die samen met de stellingen voor de promotie los bijgeleverd werd. Die fundamentele vraag luidt: 'Hoeveel ruimte is er in de ontwikkelingspsychologische theorievorming voor de constatering dat mensen in staat zijn nieuwe gedragspatronen, nieuwe ideeën en gevoelens, nieuwe manieren van functioneren te ontwikkelen? Hoe dynamisch of historisch zijn de theorieën en verklaringen in de ontwikkelingspsychologie en welke bijdrage kunnen consequent epigenetische verklaringen leveren?' 'Psychische epigenese' verwijst hier naar het verschijnsel dat er nieuwe ontwikkelingen zijn in gedachten, gevoelens, gedragingen van de individuele persoon, van een bepaalde groep mensen of van de mensheid als geheel. Vanuit deze vragen en haar eigen psychogenetisch standpunt bekijkt de auteur kritisch het veld van theorieën t.a.v. de kinderlijke ontwikkeling, laat zien wat er aan bruikbaar overblijft van die zo vaak heterogene theorieën en weet tegelijk in haar eigen onderzoek te laten zien dat vanuit de door haar gecreëerde ruimte dynamischer, creatiever en dichter bij het gewone leven staand onderzoek mogelijk is.

In deze recensie wordt eerst aanduidenderwijs iets over de inhoud van het boek gezegd. Vervolgens wordt aangegeven waarom het een belangrijke bijdrage is aan de Nederlandstalige ontwikkelingspsychologie. Tot slot worden er een paar kritische aantekeningen gemaakt, niet om de waardering in te perken maar veel eerder om er toe bij te dragen dat de discussie verder gaat.

Na een inleiding (hoofdstuk 1) waarin de, ook historische, context neergezet wordt, volgt deel I; hoofdstuk 2, Epigenese als een vorm van zelforganisatie, waarin Piagets bijdrage kritisch behandeld wordt. Hoofdstuk 3 gaat over de psychische ontwikkeling als internalisatie van sociale interacties, waarbij uiteraard

Vygotskij veel aandacht krijgt. In hoofdstuk 4 wordt dan een synthese voorgesteld in termen van coacties op verschillende organisatie-niveaus. Daarmee wordt bereikt dat de nieuwvorming van structuren en functies niet meer deterministisch voorgeprogrammeerd gezien kan worden, maar open naar een niet voorspelbare toekomst. Nadat aldus het theoretisch grondwerk gedaan is, wordt in deel II de algemene doelstelling nader uitgewerkt voor een deelsterrein: theorievorming over en onderzoek naar de psychische ontwikkeling van baby's en hun ouders.

In hoofdstuk 5 wordt de epigenese in de klassieke babytheorie behandeld. In een vrij compleet overzicht over bestaande theorieën komen in historisch perspectief alle grote mannen aan bod: Watson, Freud, Piaget, Bowlby. Haar conclusie: aanleg en omgeving blijven bij hen onafhankelijke grootheden, een van de dualismen die overwonnen dient te worden. Hoofdstuk 6 vat onder de treffende titel 'De competente baby' recent onderzoek samen, waarbij terecht aandacht gegeven wordt aan de activiteiten van 'de prenatale baby'. Hoofdstuk 7 bespreekt de omgang met de baby, waarbij wederzijdse communicatie en de zorgrelatie aandacht krijgen.

De epigenese in recent baby-onderzoek, hoofdstuk 8, laat zien dat een dynamischer en historisch meer verantwoorde theorie mogelijk is. Deel III: De vruchtbaarheid van epigenetische babytheorie, begint met hoofdstuk 9, dat laat zien dat in dit perspectief tijd- en contextgebonden patronen van psychische ontwikkeling beschreven kunnen worden, en dat babywetenschap en ouderschapscultuur elkaar constructief beïnvloeden. In de hoofdstukken 10 en 11 worden dan de resultaten van een eigen kleinschalig 'exploratief' onderzoek gebruikt om nieuwe epigenetische begrippen te ontwikkelen, waarbij 'patroonvormen' als psychologisch het meest kenmerkende wordt ingevoerd. Daarna komen aan de orde: het bepalen van zorghoudingen, toeschrijven en wederzijds creëren. In het afsluitende twaalfde hoofdstuk wordt dan alles nogmaals op een rijtje gezet en de balans opgemaakt. Een eruit springend voordeel van de benadering is dat reductionistische en dualistische verklaringen vermeden kunnen worden. Mogelijke nadelen

hiervan zijn theoretische onbepaaldheid, breedheid van behandelingsperspectieven bij ontwikkelingsproblemen en methodologische bezinning die tot een impasse in het onderzoek kan leiden. Heldere bijlagen over de methodologische verantwoording, de manier van interviewen en de gebruikte temperamentsvragenlijst sluiten dan af.

Deze samenvatting doet slechts ten dele recht aan de uiterst zorgvuldige werkwijze van de auteur. Duidelijke opbouw, een goede contextuele, actuele en historische plaatsbepaling van het thema en een heldere schrijfstijl maken dit boek tot een plezier voor de lezer. Het betoog is haast overal overtuigend: men bemerkt dat de auteur haar veld goed beheerst. Terwijl dat veld nogal eens aanleiding geeft tot elkaar verketterende onbillijke partijdigheid en simplistische partijdigheid, is het haar gelukt haar overtuigingskracht te ontlenen aan de kracht van haar argumenten. En dat alles in een taal die een ontwikkelde leek kan lezen. Zo is haar behandeling van de *infameuze* kwestie van 'nature and nurture' voorbeeldig. 'Organisme' en 'omgeving' zijn niet te beschouwen 'als autonome entiteiten' (p. 55); het zijn niet 'twee typen oorzaken' die 'onafhankelijk van elkaar werkzaam zijn: ze reguleren elkaars werking voortdurend' (p. 57). In haar afsluitende oordeel over Piaget en Vygotskij heeft ze haarscherp een van de gronden voor haar oordeel aangegeven: 'pas wanneer dit streven naar unicausale principes wordt opgegeven ontstaat mijns inziens uitzicht op een historische of consequent epigenetische ontwikkelingstheorie over psychische processen.'

Haar eigen onderzoek laat overtuigend zien dat haar 'dynamiserende' aanpak werkt. Zij laat ons a.h.w. zelf mee ontdekken dat baby én ouders samen met elkaar 'groeien' en in elk geval deels 'anders' groeien. Met andere woorden, de baby's maken niet een voorgeprogrammeerde lineaire ontwikkeling door, die door ouders en anderen eromheen slechts bevordert of geremd kan worden. Geen quasi-gefixeerde wereld waarbinnen baby's mens worden: ontwikkeling is open. Nossent rekent af met de in de leerpraktijk tot voor kort overheersende visie van voorspelbare, gedetermineerde ontwikkeling. Dit boek zou met recht en reden in de plaats kunnen komen van de

handboeken ontwikkelingspsychologie, die te vaak hapklare diepvriesmaaltijden geven. Hier zouden studenten door haar helder overzicht van conflicterende theorieën de discussie zelf kunnen volgen vanuit een duidelijk formulerende, actieve veldonderzoeker die weet waarover zij praat.

Natuurlijk roept zo'n brede studie ook vragen en kritiek op, uiteraard vooral bepaald door het perspectief van de recensent en – naar we hopen – het perspectief van de lezers. De schrijfster meent dat haar aanpak 'interdisciplinair' onderzoek bevordert. Terecht! Maar dat geldt dan toch niet alleen biologie, sociologie en antropologie (p. 258). Was de schrijfster opgegroeid in een meer 'pedagogisch' geïntereerde ontwikkelingspsychologische traditie, dan was haar niet ontgaan dat Stern in de kwestie van natuur en omgeving al duidelijk 'der Grundsatz der Konvergenz' aanhangt. Dan is aanleg geen predestinatie, maar speelruimte voor de toekomst en binnen die speelruimte is opvoeding en omgeving actief om de feitelijke ontwikkeling mogelijk te maken, die voor hem gestuwd wordt door de persoon die zich wil ontwikkelen (W. Stern, *Psychologie der frühen Kindheit, bis zum sechsten Lebensjahre*, Leipzig, 1921, 2 Auflage, 17-23). Daarnaast zou Langevelds kritiek op Piaget haar niet ontgaan zijn: die stemt in hoge mate met de hare overeen! Ook ontbrak in haar discussie een boek dat haar zeer zeker geholpen had, van Diane E. Eyer, *Mother-Infant Bonding, a scientific Fiction*, Yale University Press, New Haven, 1992.

Zo baanbrekend als de inhoud van Nossents boek en onderzoek is, zo *traditioneel* is de verantwoording van haar methodologie. Ze noemt haar onderzoek 'exploratief' (p. 269); dat moet dan wel in tegenstelling met 'gewoon' experimenteel of anderszins hypothese-toetsend onderzoek zijn?

Maar als men, zoals de auteur, terecht meent dat we behoefte hebben aan een historische psychologie, dan is toch elk onderzoek 'exploratief'? Hier wreekt zich dat nog steeds Popper als de filosoof wordt gezien. Maar er is nu na de tijd van Popper toch echt wel het een en ander gebeurd. Na Feyerabend, Kuhn, Lakatos (leerlingen van Popper, die uiteindelijk de meester achter zich lieten) en meer recent Margolis (J. Margolis, *The Flux of His-*

tory and the Flux of Science, University of California Press, Berkeley, 1993) is er geen reden meer om angst te hebben voor 'relativisme' (o.a. p. 266). Onze kennis is simpelweg altijd lokaal en vanuit een, noodzakelijk, beperkt perspectief. Soms lijkt de auteur dat onbekommerd ook te denken, soms kruipt ze terug naar niet meer houdbare angst voor het postmodernisme. Heeft dat er haar ook toe gebracht die vreemde temperamentsvragenlijst te gebruiken, of is dat de schuld van de promotoren? Immers als er al zoiets als temperament zou bestaan, is het slecht te meten. En als men het zou meten dan kan men toch niet toestaan dat een aantal 'van 16 tot 2' (p. 271) van 30 vragen de eerste keer niet beantwoord worden; de relatieve waarden van de cijfers lijken me dan niet meer statistisch bruikbaar. Het allervreemdste lijkt me hier dat dit niet beantwoorden dan in de tabellen niet terug te vinden is (o.a. p. 276). Nee, het allervreemdste is dat deze vragenlijsten toch niets toevoegen aan wat de vraaggesprekken opleveren. Een, althans voor de recensent, onbegrijpelijke opmerking is dat haar 'methodologische herbezinning' zou kunnen leiden tot een impasse in het onderzoek, 'bij gebrek aan geschikte methoden'; immers niet alleen in de sociologie (ethnomethodologie, enz.) maar ook in de psychologie en de pedagogiek zijn er mensen - ik weet uit een andere traditie - die al jaren dit type onderzoek doen (vgl. W. Lippitz, *Phänomenologische Studien in der Pädagogik*, Deutscher Studien Verlag, Weinheim, 1993).

En het allerbeste werk dat zich aan elk methodenkorset onttrekt is uit de dertiger jaren: James Agee en Walker Evans, *Laat ons nu vermaarde mannen prijzen* (Amsterdam, Bert Bakker, 1991, vert. F. van Dixhoorn, oorspr. 1939). Zij laten zien dat rigoreuze precisie en geëngageerde nabijheid kunnen samengaan, en dat methoden steeds opnieuw uitgevonden en verantwoord kunnen worden.

Afsluitend: de auteur schreef een zeer waardevolle studie, bruikbaar en vernieuwend; we wensen haar toe onversaagd, en nog een tijdje radicaler door te gaan.

A.J. Beekman

M. Luna Rubio

Leerlingbegeleiding: wens en werkelijkheid

Garant, Leuven/Apeldoorn 1996

327 pagina's f 68,50 ISBN 90 5350 508 3

Een uur later naar school en dus wat langer uitslapen, dat is mijn voornaamste herinnering aan een mentoruur. Aan het begin van de jaren tachtig zat ik in de brugklas. Op het rooster stond een dergelijk uur, maar dat werd bijna nooit ingevuld. Bij mijn weten had ik na de brugklas geen mentor meer, en contact met andere leerlingbegeleiders, zoals een decaan of een vertrouwenspersoon, kwam zelden voor. Uit onderzoek van Hüpscher-Post (1983) en Bakker (1986) blijkt dat de inzet van een docent voor leerlingbegeleiding destijds een persoonlijke keuze was, die niet of slechts gedeeltelijk werd gehonoreerd met taakuren. Op de overgrote meerderheid van scholen in het voortgezet onderwijs was geen sprake van een structurele aanpak van de leerlingbegeleiding.

Door de invoering van de basisvorming heeft leerlingbegeleiding, en dan met name het mentoraat, meer aandacht gekregen, omdat bijvoorbeeld leerlingen al vanaf de eerste klas moeten worden begeleid bij het maken van keuzen, bij het ontdekken van wat ze willen, wat ze kunnen en het leggen van een verband tussen de verschillende schoolvakken en de beroepenwereld waar ze ooit terecht komen. In de wet staat echter niet voorgeschreven dat scholen studielessen, mentoruren of beroepsoriëntatie moeten geven, maar de meeste scholen gebruiken tegenwoordig een deel van hun zogeheten vrije ruimte voor dergelijke lessen.

Binnèn de, inmiddels opgeheven, Sectie Leerlingbegeleiding van de Vakgroep Onderwijskunde van de Universiteit Utrecht, begon Luna Rubio in 1989 met een onderzoek naar de activiteiten van de leerlingbegeleider. Voor het onderzoek werden drie doelen geformuleerd: het ontwikkelen van een instrument voor het meten van de activiteiten van leerlingbegeleiders, het ontwikkelen van indicatoren voor het meten van de effectiviteit van leerlingbegeleiding, en het analyseren van de opleiding Leerlingbegeleiding/Schooldecaan aan de hand van de twee eerdere fasen. Dit profielschrift heeft slechts betrekking op de eer-

ste doelstelling. Aan de andere doelstellingen is, zo blijkt uit het boek, geen uitvoering gegeven.

Leerlingbegeleiding, zo definieert Luna Rubio (p. 16) haar onderwerp, is "het geheel van activiteiten op het gebied van leerprocessen, keuzeprocessen, groepsprocessen en persoonlijke ontwikkeling, die berusten op een belangstellende houding van de begeleider ten opzichte van de leerling, met als doel de problemen die de prestaties en de algemene ontwikkeling van de leerling belemmeren te signaleren, en op die manier zo mogelijk te voorkomen en te verhelpen." Vervolgens probeert zij de vraag te beantwoorden wat voor activiteiten leerlingbegeleiders verrichten op het gebied van onderwijsleerprocessen, keuzeprocessen en groepsprocessen en in welke mate ze dat doen. Aan de inzet van begeleiders voor de persoonlijke ontwikkeling besteedt Luna Rubio geen aandacht, omdat het niet lukte om inhoudelijk valide items te ontwikkelen.

Om tot een antwoord op de bovenstaande vraag te komen, heeft Luna Rubio zowel kwantitatief als kwalitatief onderzoek gedaan. Het kwantitatief onderzoek bestaat uit twee enquêtes, waarbij de steekproef in beide gevallen niet representatief is voor de populatie van leerlingbegeleiders in het voortgezet onderwijs in Nederland. Bij de eerste afname zijn de vragenlijsten namelijk verspreid onder zeshonderd schoolleiders, met aan hen de vraag om de enquête te verstrekken aan de persoon die er volgens hen het meest voor in aanmerking kwam. Bij de tweede afname zijn de vragenlijsten alleen maar verspreid onder voormalige deelnemers aan de eerder genoemde opleiding Leerlingbegeleiding/Schooldecaanaat. De respons bedroeg respectievelijk 39 en 27 procent. Als aanvulling op de schriftelijke vragenlijsten, werden in het schooljaar 1993-'94 interviews afgenomen onder negentien leerlingbegeleiders.

Het proefschrift is opgebouwd uit vier delen en zeventien hoofdstukken. In het eerste deel wordt de bestaande literatuur over leerlingbegeleiding aangehaald. Hieruit blijkt dat er op dit terrein slechts sporadisch onderzoek is gedaan. Er is dan ook nauwelijks inzicht in

wat scholen in werkelijkheid aan leerlingbegeleiding doen, wat dit van docenten vergt en welke effecten dit heeft voor leerlingen. Voor zover er buitenlands onderzoek beschikbaar is, is het steeds de vraag in hoeverre dat van toepassing is op ons onderwijssysteem.

De ontwikkeling van een Likert-instrumentarium voor het meten van de activiteiten van de leerlingbegeleider, komt in het tweede deel aan de orde. Hierbij gaat Luna Rubio in op de inhouds- en begripsvaliditeit, de betrouwbaarheid en de stabiliteit van het instrumentarium. Door factoranalyse heeft zij de items geschaald. De indeling correspondeerde niet met de indeling die door het literatuuronderzoek werd verwacht, maar de schalen zijn volgens Luna Rubio toch bruikbaar voor onderzoek naar de activiteiten van begeleiders op het gebied van onderwijsleerprocessen, keuzeprocessen of groepsprocessen.

In deel drie worden zowel de opzet van het kwalitatieve onderzoek als de resultaten van de kwantitatieve en kwalitatieve onderzoeken vermeld. In het vierde deel voorziet de schrijver het onderzoek van enige kanttekeningen en vermeldt zij de implicaties van het onderzoek. Tevens probeert zij een theoretische basis te leggen voor vervolgonderzoek naar leerlingbegeleiding.

Aan de mogelijkheid tot sociaal-wenselijke antwoorden van respondenten, besteedt Luna Rubio speciaal aandacht. In de schriftelijke vragenlijsten heeft zij enkele vragen toegevoegd, die te zamen een sociale-wenselijkheidsschaal vormen. De vragen over leerlingbegeleiding die significant correleren met de sociale-wenselijkheidsschaal zijn bij verdere analyses niet gebruikt, om te vermijden dat de respondenten een te rooskleurig beeld geven van hun activiteiten. Door deze aanpak werden verscheidene, inhoudelijk relevante items en sommige schalen in het geheel verwijderd. Luna Rubio had naar mijn mening er beter voor kunnen kiezen om bij de schaalconstructie geen rekening te houden met sociale wenselijkheid en er tijdens analyses voor te controleren.

De sociale-wenselijkheidsschaal is bovendien onzorgvuldig samengesteld. Er is namelijk sprake van overlap tussen de inhoud van deze schaal en de schalen voor leerlingbege-

leiding, wat het sterkst tot uitdrukking komt bij de volgende items van de sociale-wenselijkheidsschaal: 'de problemen die leerlingen vertonen worden op onze school altijd serieus genomen' en 'de persoonlijke ontwikkeling van de leerlingen is op onze school net zo belangrijk als hun leerprestaties'. Het gevolg is dat, zoals Luna Rubio (p. 93) zelf ook aangeeft, niet duidelijk wordt in hoeverre sociale wenselijkheid in het onderzoek een rol speelt.

Net als Bakker (1986) maakt Luna Rubio bij de bespreking van haar resultaten geen onderscheid tussen de verschillende soorten leerlingbegeleiders. Een dergelijk onderscheid zou in toekomstig onderzoek waarschijnlijk verhelderend zijn. Logischerwijs kan dan worden verwacht dat decanen zich vooral bezighouden met keuzeprocessen, vakdocenten en 'remedial teachers' met onderwijsleerprocessen, vertrouwenspersonen met persoonlijke ontwikkeling en mentoren met alle processen. Vervolgens zou dan ook moeten worden gekeken of er verschillen tussen scholen zijn. Wanneer deze verschillen er zijn, zou kunnen worden nagegaan in hoeverre die verschillen samenhangen met het beleid van een school ten aanzien van leerlingbegeleiding. Het zou dan kunnen blijken dat op bepaalde scholen vakdocenten zich ook met keuzeprocessen bezighouden, of dat op sommige scholen mentoren alleen maar gericht zijn op onderwijsleerprocessen. Bij een dergelijke opzet zou het verstandig zijn om per school bij meer dan een functionaris een vragenlijst af te nemen.

In de activiteiten van de respondenten geeft Luna Rubio weinig inzicht, aangezien frequentietabellen, gemiddelden en standaarddeviaties ontbreken. In dit opzicht heeft zij haar vraagstelling onvoldoende beantwoord. Zij voert wel verscheidene t-toetsen uit. De uitkomsten daarvan en alle opmerkingen die zij maakt op basis van dergelijke analyses, moeten met de nodige twijfel worden beschouwd. Dergelijke toetsen horen namelijk niet bij een explorerend onderzoek met een

selecte steekproef. Of de genoemde opleiding, zoals Luna Rubio zelf hoopt, kan worden verbeterd door deze studie, is zeer de vraag. Wat zeker is, is dat op grond van deze studie niet kan worden geconcludeerd in hoeverre scholen te kort schieten in het begeleiden van leerlingen.

Er zijn indicaties dat op een deel van de scholen voor voortgezet onderwijs leerlingbegeleiding nog steeds niet structureel is geregeld (vgl. Bos, Vermeulen & Cremers-van Wees, 1993). Het ontbreekt op die scholen aan voldoende facilitering voor leerlingbegeleiders, aan beleid op dit gebied of aan overleg tussen de verschillende functionarissen. In het algemeen, en zeker in de onderbouw, geven docenten tegenwoordig vaker begeleiding en staan er vaker mentoruren op het rooster. De kans dat een leerling in het voortgezet onderwijs wekelijks vrij heeft tijdens dat uur, is in vijftien jaar dan ook een stuk kleiner geworden.

Dit proefschrift levert dus nieuwe instrumenten op die hopelijk vaker gebruikt zullen worden. Het onderzoek vind ik af en toe knullig uitgevoerd en dat is zonde, want op deze manier heeft het onderzoek weinig zeggingskracht en levert het nauwelijks nieuwe inzichten in de activiteiten van leerlingbegeleiders op.

D.R. Veenstra

Literatuur

- Bakker, J.C. (1986) *Docenten en leerlingen. Onderzoek naar tijdsbesteding, begeleidingsgedrag en motivatie van avo-docenten in het kader van leerlingbegeleiding*. Amsterdam: VU Uitgeverij (dissertatie Vrije Universiteit).
- Bos, K.T., Vermeulen, C.J.A.J., & Cremers-van Wees, L.M.C.M. (1993) *Leerling- en keuzebegeleiding bij de basisvorming*. Enschede: OCTO.
- Hüpscher-Post, A.D. (1983) *Leerlingbegeleiding in het voortgezet onderwijs*. Rotterdam: Vakgroep Onderwijs sociologie en Onderwijsbeleid, Erasmus Universiteit Rotterdam.

Mededelingen

Promoties

Universiteit van Amsterdam

Promovendus: Drs. M. Vermeulen

Titel proefschrift: De school als arbeidsorganisatie

Promotor: Prof.dr. A.M.L. van Wieringen

Co-promotor: Dr. S. Karsten

Datum: 11 juni 1997

Universiteit van Amsterdam

Promovendus: Drs. T.M. Bechger

Titel proefschrift: Methodological aspects of educational comparison: the case of reading literacy

Promotor: Prof.dr. C.M. de Glopper

Co-promotor: Prof.dr. J.J.C.M. Hox

Datum: 4 juli 1997

Vrije Universiteit

Promovendus: Drs. E.F.J.M. Brand

Titel proefschrift: Differentiation of attention problems in children

Promotor: Prof.dr. D.J. Bakker en

Prof.dr. J.F. Orlebeke

Datum: 16 juni 1997

Vrije Universiteit

Promovendus: Drs. J.R. Koopmans

Titel proefschrift: The genetics of health-related behaviors. A study of adolescent twins and their parents

Promotor: Prof.dr. J.F. Orlebeke

Datum: 17 juni 1997

Universiteit Twente

Promovendus: Drs. Th.J. Bastiaens

Titel proefschrift: Werken en leren met

Electronic Performance Support Systems

Promotor: Prof.dr. W.J. Nijhof

Datum: 12 november 1997

Rijksuniversiteit Leiden

Promovendus: Drs. H.J. Vermeer

Titel proefschrift: Sixth-grade students' mathematical problemsolving behavior.

Motivational variables and gender differences

Promotor: Prof.dr. M. Boekaerts

Datum: 16 oktober 1997

Recent gestart onderzoek

Universiteit van Amsterdam

Titel project: Effect-evaluatie Kaleidoscoop en Piramide

Projectleiding: Dr. P. Leseman en

Drs. A. Veen

Titel project: Integratie van OET/OALT in het basisonderwijs

Projectleiding: Drs. M. Robijn

Universiteit Twente

Titel project: Didactisch model simulator-trainingen

Projectleiding: Prof.dr. J.M. Pieters

Titel project: Multi-niveau IRT-modellen met onbetrouwbaar gemeten covariaten

Projectleiding: Prof.dr. W.J. van der Linden

en Dr. C.A.W. Glas

Oratie

Dr. G.F. Heyting, benoemd tot hoogleraar in de faculteit der Pedagogische en Onderwijskundige Wetenschappen aan de Universiteit van Amsterdam, met de leeropdracht 'Opvoedkunde, in het bijzonder de grondslagen en geschiedenis van de pedagogische wetenschappen' heeft op 27 juni 1997 haar ambt aanvaard met het uitspreken van een rede, getiteld: *Het vanzelfsprekende en het discutabele; een schets van opvoedkundig grondslagenonderzoek.*

Afscheidscolleges

Prof.dr. F.J. Mönks heeft op 6 juni 1997 zijn werkzaamheden als hoogleraar in de Ontwikkelingspsychologie aan de Katholieke Universiteit Nijmegen beëindigd met een afscheidscollege, getiteld: *Differentiatie en integratie.*

Prof.dr. J.H.G.I. Giesbers heeft op 27 juni 1997 zijn werkzaamheden als hoogleraar in de Onderwijskunde aan de Katholieke Universiteit Nijmegen beëindigd met een afscheidscollege, getiteld: *Waarde(n)vaste schoolorganisatie en schoolleiding.*

Benoemingen

Dr. M. Mulder is per 1 april 1997 benoemd als Universitair Hoofddocent bij de Vakgroep Onderwijsorganisatie en -Management van de Faculteit der Toegepaste Onderwijskunde aan de Universiteit Twente.

Drs. G. van Engelenburg is per 15 april 1997 benoemd als Universitair Docent bij de Vakgroep Onderwijskundige Meetmethoden en Data-analyse van de Faculteit der Toegepaste Onderwijskunde aan de Universiteit Twente.

SWIDOC

Met ingang van 1 september 1997 zal het SWIDOC te zamen met vier andere KNAW-instituten opgaan in het Nederlands Instituut voor Wetenschappelijke Informatie: NIWI, dat gevestigd is op het adres: Joan Muyskenweg 25, 1096 CJ Amsterdam.

Het postadres is: NIWI, Postbus 41950, 1009 DD Amsterdam; telefoon: 020-4628600; fax: 020-6685079; web site: <http://www.niwi.knaw.nl>

Inhoud andere tijdschriften

Comenius

17e jaargang, nr. 2, 1997

Het welbevinden van jongeren uit verschillende gezinstypen, door E. Spruijt en M. de Goede

Moet besnijdenis mogen? door P. van der Ploeg

Opvattingen en handelen van cursisten en docenten in de basiseducatie, door M. Lunenberg en M. Volman

Van tobende puber naar lachende leerlingen, door A. Dieleman

Forum

'Voor enkele toevallige geïnteresseerden'.

Onderzoeksperspectieven voor een historische opvoedingssociologie, door R. Vanderstraeten

Tijdschrift voor Orthopedagogiek

36e jaargang, nr. 5, 1997

Paradigma-verschuiving in de visie op zorg voor mensen met een verstandelijke handicap

(inaugurele rede), door A.Th.G. van Gennep
De lokroep van het dominante paradigma.
Commentaar bij de oratie van Prof.dr. A.Th.G. van Gennep

Toch nog leren lezen in het basisonderwijs?
door W. Verhagen, C.A.J. Aarnoutse en J.H.L. Oud

Verblijfsduur als verklaring voor de groei van het speciaal onderwijs, door Y.J. Pijl

Tijdschrift voor Orthopedagogiek

36e jaargang, nr. 6, 1997

Risicogezinnen en predictie en preventie van kindermishandeling, door H.E.M. Baartman
Stimulering van het denken. Classificatie van programma's en enkele Europese trends, door J.H.M. Hamers en M.Th. Overtoom

Ei of ij, au of ou? Het effect van lees- en schrijfvoering met elkaar vergeleken, door W.H.J. van Bon en J. van Staalduinen
WSNS, beleid en statistiek, door A.J.J.M. Ruijsenaars en D.G.M. van der Aalsvoort

Ontvangen boeken

Day, Chr., Veen, D. van & Walraven, G. (Eds.), *Children and youth at risk and urban education*. Garant, Leuven/Apeldoorn, 1997, f 59,50.

Driesen, L., *Straffen?* Garant, Leuven/Apeldoorn, 1996, f 39,90.

Goorhuis-Brouwer, S.M., *Het wonder van de taalverwerving*. Uitgeverij De Tijdstroom, Utrecht, 1997, f 29,75.

Graas, D., *Zorgenkinderen op school. Geschiedenis van het speciaal onderwijs in Nederland, 1900-1950*. Garant, Leuven/Apeldoorn, 1996, f 59,50.

Groot, R. de, Heethuis, H., & Koekkoek, M. (red.), *Hulp op maat. Leervervaardentraining in de basisvorming*. Garant, Leuven/Apeldoorn, 1996, f 28,-.

Hilst, K. van der, *Basisboek psycho-sociale begeleiding. Probleemoplossen in de persoonlijke levensloop*. Uitgeverij SWP, Utrecht, 1997, f 34,50.

Hintjes, J., *Analytische ethiek. Een kritische introductie*. Uitgeverij Giotto, Vaals, 1997, f 25,-.

Kleijer, H., & Tillekens, G. (red.), *Burgerschap zonder moraal. Een sociologische kijk op de verantwoordelijke burger*. Garant, Leuven/Apeldoorn, 1997, f 39,50.

W.C.M. Resing en N. Bleichrodt*

Samenvatting

In dit artikel staat de gereviseerde Schoolgedrag Beoordelingslijst (SCHOBL-R) centraal. De vragen in het onderzoek dat met de SCHOBL-R werd uitgevoerd betreffen de factorstructuur van het instrument en de stabiliteit hiervan over leeftijd, de mogelijkheid om het gedrag van het individuele kind in zoveel mogelijk inhoudelijk valide gedragsclusters te beschrijven, de sekse-specificiteit en het leeftijdsbereik van de vragenlijst en de samenhang tussen sociaal-emotioneel gedrag zoals gemeten met de SCHOBL-R en intelligentiescores en schoolprestaties.

Uit onderzoek blijkt dat de gereviseerde Schoolgedrag Beoordelingslijst een duidelijke vier-factorstructuur heeft die stabiel blijft voor kinderen van 4 tot en met 11 jaar. De gedragsdomeinen die door de vragenlijst bestreken worden zijn: Extraversie, Werkhouding, Aangenaam Gedrag en Emotionele Stabiliteit. Deze factoren komen overeen met vier van de vijf factoren van de Big Five. Wanneer de beide parallelversies van de vragenlijst worden afgenomen, kan een kind beschreven worden in termen van een dertiental gedragsclusters. De betrouwbaarheid van het instrument is goed. Meisjes en jongens worden door leerkrachten anders beoordeeld. De hoogste verbanden tussen intelligentie en schoolprestaties enerzijds en sociaal-emotioneel functioneren anderzijds worden gevonden voor Werkhouding. De werkhouding van kinderen blijkt vrij sterk samen te hangen met enerzijds het logisch redeneren en anderzijds de behaalde schoolprestaties.

1 Inleiding

Aan het eind van de jaren zeventig ontwikkelde Zaal (1978, 1980, 1981) de Schoolgedrag Beoordelingslijst voor kinderen, ook wel de SCHOBL genaamd. Het schoolgedrag, dat door de leerkrachten met behulp van deze vragenlijst in kaart wordt gebracht, kan worden beschreven als 'de wijze waarop kinderen doorgaans omgaan met de leerkracht en met elkaar' (vgl. Bleichrodt, Resing & Zaal, 1993). Daarbij gaat het om voor het individuele kind karakteristieke gedragingen die een zekere consistentie dienen te vertonen over tijd en over situaties. Bij de omschrijving van wat het 'sociaal-emotionele gedragsdomein' wordt genoemd, is aangesloten bij persoonlijkheidstheorieën waarin individuele trekken ('traits') een centrale plaats innemen. Deze trekken zelf verwijzen op hun beurt weer naar concrete en specifieke gedragingen (vgl. Guilford, 1959; Zaal, 1978).

De Schoolgedrag Beoordelingslijst is samengesteld op basis van typeringen en uitspraken van leerkrachten over sociaal-emotioneel gedrag. Bij het definiëren van de items voor de vragenlijst is niet uitgegaan van probleemgedrag, maar van 'normaal gedrag in de klas', dit in tegenstelling tot een instrument als bijvoorbeeld de Child Behavior Checklist (CBCL: Verhulst, Koot, Akkerhuis & Veerman, 1990), de Conners Teacher Rating Scale (Blöte & Curfs, 1986) of de Amsterdamse Kindergedragslijst (De Jong & Das-Smaal, 1991). Er is geprobeerd het hele scala aan 'normaal' schoolgedrag, voor zover dat ligt op het sociaal-emotionele vlak, in kaart te brengen. Binnen de toepassing van de vragenlijst is echter zeker ook plaats voor de beschrijving van onaangepast en probleemgedrag. Hierbij geldt echter de restrictie dat 'afwijkend gedrag' is opgevat in termen van de frequentie waarin bepaalde gedragingen binnen de normale schoolpopulatie worden aangetroffen. De normen zijn dus niet gebaseerd op specifieke klini-

* Met dank aan J.N. Zaal voor zijn bijdrage aan de constructie van en onderzoek met de oorspronkelijk Schoolgedrag Beoordelingslijst (SCHOBL) en aan P. Dekker voor zijn bijdrage aan de data-analyses.

sche groepen en de beschrijving van het gedrag komt niet voort uit het gedrag van specifieke groepen gedragsgestoorde kinderen. Een dergelijke definitie van van 'normaal' afwijkend probleemgedrag is niet ongebruikelijk en heeft ook voordelen, aangezien een instrument dat gebaseerd is op 'normaal gedrag in de populatie' voor een grotere gebruikersgroep geschikt is. De lijst zal echter niet noodzakelijk alle mogelijke vormen van afwijkend gedrag vertonen (vgl. Guilford, 1959; Anthony, 1970).

Een eerste reden om bij de definitie van het sociaal-emotionele gedragsdomein niet uit te gaan van probleemgedrag is van psychometrische aard. Er ontstaan tal van moeilijkheden bij het analyseren en interpreteren van gegevens indien de variabelen slechts een geringe frequentie van voorkomen hebben. Bovendien heeft het laten beoordelen op weinig voorkomende gedragingen een ongunstige invloed op de bereidheid van leerkrachten het gedrag van alle leerlingen te beoordelen. Daarnaast bestaat het gevaar dat het laten beoordelen op problematisch gedrag een soort 'halo-effect' teweeg brengt: de leerling wordt slechts op één algemene goed-slecht dimensie beoordeeld.

Een tweede reden om niet van probleemgedrag, maar van 'normaal' gedrag in de populatie uit te gaan, was de constatering dat de structuren van gedragsbeoordelingen voor uiteenlopende groepen onder diverse omstandigheden veelal een overeenkomstig patroon laten zien (vgl. Cattell & Coan, 1957; Digman, 1963; Norman, 1963; Zaal, 1978). Ook uit meer recente studies op het gebied van de persoonlijkheidsbeschrijving (vgl. o.a. Digman, 1989, 1990; Goldberg, 1993; Halverson, Kohnstamm & Martin, 1994; Hofstee & De Raad, 1991) blijkt dat de structuur van persoonlijkheidsbeoordelingen, voor zover empirisch onderzocht door middel van factor- of clusteranalyse, sterke overlap vertoont met de factoren die worden gerapporteerd in eerder aangehaalde studies. Steeds vaker is er sprake van een vijftal persoonlijkheidsdimensies, ook wel de Big Five of het Vijf-factoren model genoemd (zie ook Norman, 1963). Deze vijf-factorstructuur is terug te vinden in persoonsbeschrijving van zowel volwassenen (Elshout & Akkerman, 1975; Hofstee & De Raad, 1991; McCrae & Costa, 1985) als kinderen (Digman, 1963,

1994; Digman & Inouye, 1986; Halverson, Kohnstamm & Martin, 1994; John et al., 1994; Kohnstamm, 1992a, 1992b; Kohnstamm, Halverson, Havill & Mervielde, 1996; Van Lieshout & Haselager, 1992; Mervielde, 1994). Om een dergelijke structuur te kunnen vinden, dient volgens Kohnstamm (1992a) tenminste aan twee voorwaarden te zijn voldaan: 1) het scala aan gedragsbeoordelingen dient zo groot mogelijk te zijn en 2) willen de factoren zich kunnen manifesteren dan dient de onderzoeksgroep voldoende omvang te hebben.

De vijf factoren die wel worden onderscheiden, zowel in de vrije beschrijving door ouders van het gedrag van hun kinderen (vgl. o.a. Kohnstamm, Halverson, Havill & Mervielde, 1996; Kohnstamm, Mervielde, Besevegis & Halverson, 1995) als in de beschrijving door leerkrachten van het gedrag van kinderen op school (o.a. Digman, 1963, 1989; Mervielde, 1994) zijn: (I) Extravert - Introvert; (II) Aangenaam - Onaangenaam; (III) Zorgvuldig - Onzorgvuldig; (IV) Emotioneel stabiel - Emotioneel instabiel; (V) Ideeënrijk - Ideeënarm. Hofstee (1992) en Hofstee en De Raad (1991) geven een vergelijkbare omschrijving van de vijf gronddimensies: (I) Extraversie / Surgency / Spontaniteit / Activiteit; (II) Agreeableness / Goedheid / Vriendelijkheid / Altruïsme / Respect; (III) Conscientiousness / Zorgvuldigheid; (IV) Emotionele stabiliteit / Onverstoorbaarheid versus Emotionaliteit / Neuroticisme; (V) Intellect / Openess to experience / Culture / Creativiteit / Autonomie.

Bij het samenstellen van de oorspronkelijke itemverzameling van de SCHOBL is gebruik gemaakt van de modellen van Stott (1962, 1967), Cattell (1957) en Becker en Krug (1964). Deze auteurs gebruikten een groot aantal paarsgewijs tegengestelde adjectieven om persoonlijkheidseigenschappen van kinderen te beoordelen. Bij het maken van de items hebben persoonlijkheidsfactoren die binnen deze modellen zijn onderscheiden geen rol gespeeld. De ongeveer 200 paarsgewijze geordende adjectieven, die uit de literatuur zijn verzameld, zijn aan ervaren leerkrachten uit de eerste groepen van het basisonderwijs voorgelegd met het verzoek deze om te zetten in overte, dat wil zeggen concreet waarneembare, gedragingen van leerlingen in de klas. Dit werd de leerkrachten gevraagd in de vorm van een gestruc-

2 Methode

2.1 De onderzoeksgroepen

Normen, psychometrische kwaliteiten, factorstructuur en valideringsgegevens voor de herziene versie van de Schoolgedrag Beoordelingslijst zijn vastgesteld op basis van een tweetal steekproeven. Voor de kinderen uit de eerste steekproef (I) is een van de parallelvormen (vorm B) van de SCHOBL-R ingevuld ($N=1296$), terwijl voor de kinderen uit de tweede steekproef (II) zowel de A- als de B-vorm is ingevuld ($N=367$).

Steekproef I bestond uit de kinderen die deel uitmaakten van de normeringssteekproef van de Revisie Amsterdamse Kinder Intelligentie-test (Bleichrodt, Drenth, Zaal & Resing, 1987). Van de in totaal 1415 kinderen uit deze RAKIT steekproef werd voor 1296 kinderen door leerkrachten uit het basisonderwijs de SCHOBL-R vorm B ingevuld. De kinderen waren afkomstig van 104 verschillende basisscholen en varieerden in leeftijd van 4 jaar en 2 maanden tot 11 jaar en 2 maanden. Er is sprake van een aselechte proportioneel gestratificeerde steekproef met trekking in twee trappen, systematisch respectievelijk aselechte. De scholen waren de primaire steekproefeenheden, de leerlingen de secundaire. Teneinde de representativiteit van de steekproef te optimaliseren en storende steekproeffluctuaties zoveel mogelijk te beperken, werden als stratificatiecriteria gehanteerd: regio, urbanisatiegraad, schoolgrootte, leeftijd en sekse. Per school zijn door de testleiders uit de schoollijsten per leeftijdsgroep telkens twee jongens en twee meisjes (4 en 5-jarigen) of een jongen en een meisje (6-11 jarigen) aselechte betrokken.

In steekproef II waren 367 leerlingen opgenomen, verdeeld over 46 basisscholen. Deze selectie van scholen vormde een aselechte steekproef uit de scholen die eerder participeerden in het RAKIT onderzoek. Per school zijn aselechte acht leerlingen geselecteerd, verdeeld over vier leeftijdsgroepen: 4-, 6-, 8- en 10-jarigen. Per leeftijdsgroep is telkens één jongen en één meisje geselecteerd. De leerkrachten van deze leerlingen vulden, met een tussenpoos van 14 dagen, per leerling beide parallelvormen van de SCHOBL-R in.

tureerd interview. Geprobeerd werd voor elk adjectief meer dan één concrete gedragsuitspraak te bedenken. De uiteindelijke itemverzameling bestond uit ongeveer 460 concrete gedragsuitspraken die vervolgens werden omgezet in 228 bipolaire items, bestaande uit telkens twee, aan elkaar tegengestelde, concrete gedragsuitspraken. Op grond van clusteranalyses en principale componentenanalyse werden uiteindelijk 104 bipolaire gedragsuitspraken geselecteerd, waarmee twee parallelvormen, vorm A en vorm B, voor het meten van schoolgedrag werden samengesteld (vgl. Zaal, 1978, 1981).

In deze bijdrage staat de beschrijving van de gereviseerde versie van de SCHOBL in de vorm van de SCHOBL-R centraal. Betrof het aanvankelijke leeftijdsbereik kinderen uit het begin van de basisschool (4 tot 6 jaar), de herziene versie is voorgelegd aan leerkrachten van leerlingen in de leeftijd van 4 tot en met 11 jaar. Om deze reden zijn enkele 'kinderlijke' items gewijzigd. Tevens is, om inhoudelijk-theoretische redenen, een beperkt aantal items gewijzigd. De vragen, die in het kader van het revisie-onderzoek naar voren kwamen, zullen in het navolgende worden beantwoord:

- Wat is de onderliggende factorstructuur van het instrument en is deze stabiel over leeftijden? Welke gedragsclusters kunnen binnen de factorstructuur worden onderscheiden?
- Is er sprake van sekse-specifiek schoolgedrag, zoals beoordeeld met het instrument?
- Kan de vragenlijst worden gebruikt voor kinderen van uiteenlopende leeftijden in het basisonderwijs?
- Welke aspecten van het sociaal-emotionele gedrag hangen samen met intelligentiescores en schoolprestaties van de leerlingen en welke niet?

Tevens is onderzocht of de Schoolgedrag Beoordelingslijst bruikbaar is voor allochtone leerlingen en voor leerlingen uit het Speciaal Onderwijs. Deze laatste twee aandachtspunten zijn elders nader beschreven (Bleichrodt, Resing & Zaal, 1993; Resing & Bleichrodt en Bleichrodt & Resing, in voorbereiding).

2.2 De meetinstrumenten en de herziening van de vragenlijst

2.2.1 De *SCHOBL-R*

Het sociaal-emotioneel functioneren van leerlingen in de klas is vastgesteld aan de hand van de *SCHOBL-R*. Deze schoolgedragslijst kent twee parallelvormen (A en B) die elk bestaan uit 52 vragen over concreet waarneembaar schoolgedrag van kinderen. Elke vraag bestaat uit een tweetal uitspraken die een tegenstelling vormen.

Bijvoorbeeld:

Praat honderduit 3 2 1 - 1 2 3 *Zegt geen woord*

Op de schaal tussen de beide uitspraken kan de leerkracht aangeven (1) welk van beide uitspraken de leerling het meest typeert en (2), binnen één van beide uitspraken, in welke mate het gedrag van de leerling wordt getypeerd: helemaal (3), redelijk (2), of net iets beter dan het tegenovergestelde gedrag (1). Het is dus de bedoeling dat de leerkracht eerst kiest welk van beide bipolaire polen de beste beschrijving van het gedrag van het kind oplevert en vervolgens een fijnere nuancering aanbrengt binnen de gekozen kant van de schaal.

2.2.2 De *RAKIT*

Om te onderzoeken welke gedragsaspecten al dan niet samenhangen met het intellectuele functioneren van het kind, werd bij alle kinderen uit steekproef I de *RAKIT* (Bleichrodt et al., 1987) afgenomen. De *RAKIT* is een voor Nederland ontwikkelde en genormeerde kinder-intelligentietest waarmee zowel een beeld kan worden verkregen van het algemeen niveau van cognitieve ontwikkeling van een kind (het *RAKIT* totaal IQ) als van sterke en zwakke kanten binnen de cognitieve ontwikkeling, in de vorm van de scores op vier onderscheiden factoren. Voor kinderen van 4 tot 5 jaar zijn dit de factoren Verbaal Leren en Vlotheid, Ruimtelijk Perceptueel Redeneren, Sequentieel Geheugen en Kwantiteit en voor kinderen van 5 jaar en ouder de factoren Perceptueel Leren, Verbaal Leren, Ruimtelijke Oriëntatie en Tempo en Verbale Vlotheid. Op voorhand wordt verwacht dat met name de werkhouding van het kind (werkt het consciëntieus, gemotiveerd en langdurig aan een taak of niet) een gematigde samenhang zal laten zien met het

intellectuele niveau van het kind (vgl. ook Zaal, 1981).

2.2.3 *Schoolprestaties*

Ter nadere validering van de *SCHOBL-R* is de samenhang tussen het schoolgedrag en de schoolprestaties van de leerlingen nader onderzocht. Aan de betreffende leerkrachten is gevraagd de prestaties op diverse schoolvakken te beoordelen. Tevens is bij de kinderen uit steekproef I een aantal CITO-toetsen op het gebied van lezen en rekenen afgenomen, afhankelijk van de leeftijd van de kinderen. Voor de 4-jarige kinderen was een dergelijke toets niet beschikbaar, kinderen van 5 jaar kregen de CITO-Begrippentoets voorgelegd, kinderen van 6 tot 8 jaar een leestoets van het CITO en kinderen van 8 tot en met 11 jaar zowel een lees- als een rekentoets. In Tabel 8 is een overzicht gegeven van de diverse toetsen die werden gebruikt.

3 Resultaten

3.1 *Factoriële structuur*

De oorspronkelijke *SCHOBL* kende vier hoofd-gedragsclusters: Vrijmoedigheid, Werkhouding, Sociale Omgang en Emotionaliteit en een 12-tal gedragschalen (zie Zaal, 1981). Bij de constructie van de oorspronkelijke lijst (Zaal, 1978) was een expliciet uitgangspunt dat sociaal-emotioneel gedrag gemeten diende te worden voor zover dat niet binnen het intellectuele domein lag. Beschrijvingen over hoe slim of hoe intelligent een kind zich in de klas gedraagt werden van aanvang af niet in de itemlijsten opgenomen. Op basis hiervan wordt derhalve verwacht dat niet vijf maar vier van de factoren uit het Vijf-factoren model zullen worden aangetroffen.

Om de factoriële structuur van het herziene instrument te onderzoeken, werd principale componentenanalyse met varimax-rotatie uitgevoerd op twee onafhankelijke datasets (steekproef I en steekproef II). Van een confirmatieve factoranalysetechniek werd afgezien op grond van theoretische overwegingen. Bookstein (1986, p. 228) concludeert dat "LISREL will frequently, but erratically, fail to produce meaningful results at all". McCrae, Zonderman, Costa, Bond en Paunonen (1996)

Tabel 1
 Hoofdladingen van de items uit de A-vorm gebaseerd op steekproef II (N=367) en van de items uit de B-vorm gebaseerd op steekproef I (N=1296), respectievelijk steekproef II (N=367 tussen haakjes)

Extraversie		Werkhouding		Aangenaam Gedrag		Emotionele Stabiliteit	
item	lading	item	lading	item	lading	item	lading
A5	.61	A4	.61	A2	.53	A1	.63
A7	.52	A8	.73	A9	.61	A6	.64
A15	.73	A13	.73	A12	.54	A11	.66
A19	.71	A16	.63	A18	.54	A17	.57
A25	.74	A22	.55	A21	.56	A23	.74
A27	.57	A28	.72	A26	.59	A29	.46
A30	.73	A32	.82	A35	.70	A37	.57
A33	.66	A36	.84	A41	.60	A47	.61
A34	.57	A39	.67	A44	.57		
A40	.65	A42	.81	A48	.65		
A43	.65	A46	.60				
A50	.68	A51	.72				
B5	.68 (.65)	B4	.66 (.68)	B2	.59 (.58)	B1	.39 (.45)
B7	.70 (.70)	B8	.66 (.70)	B9	.69 (.52)	B6	.58 (.61)
B15	.69 (.70)	B13	.68 (.69)	B12	.49 (.31)	B11	.38 (.52)
B19	.72 (.77)	B16	.55 (.64)	B18	.68 (.35)	B17	.57 (.73)
B25	.54 (.50)	B22	.83 (.83)	B21	.58 (.55)	B23	.56 (.67)
B27	.64 (.60)	B28	.78 (.78)	B26	.62 (.71)	B29	.51 (.52)
B30	.58 (.60)	B32	.53 (.57)	B35	.62 (.68)	B37	.45 (.41)
B33	.69 (.65)	B36	.69 (.77)	B41	.56 (.62)	B47	.42 (.46)
B34	.52 (.70)	B39	.67 (.66)	B44	.55 (.63)		
B40	.68 (.66)	B42	.59 (.63)	B48	.67 (.69)		
B43	.52 (.52)	B46	.52 (.48)				
B50	.65 (.66)	B51	.54 (.53)				

constateren dat in veel onderzoek met betrekking tot het Vijf-factoren model confirmatieve factoranalyses, zoals LISREL, zelfs indien van het minst restrictieve model wordt uitgegaan, geen ondersteuning voor een dergelijk model bieden. Items in persoonlijkheidsvragenlijsten blijken vrijwel nooit geheel factorzuiver te zijn. Er blijken betekenisvolle ladingen aanwezig te zijn op andere factoren. Belangrijke interpersoonlijke trekken vallen frequent tussen de orthogonale assen in (vgl. bijv. De Raad, Hendriks & Hofstee, 1992). McCrae et al. (1996) stellen ook dat het theoretisch mogelijk is dat items (bij)ladingen op 2 of meer factoren hebben. Wil men dan toch gebruik maken van een confirmatief model dan dienen in elk geval positieve secundaire ladingen en oblique rotatie ingebouwd te worden, maar ook dan valt de fit van het model over het algemeen tegen, tenzij men meerdere, inhoudelijk slecht te interpreteren en zeer specifieke factoren aan het model toevoegt. Een grote steekproef leidt bovendien volgens McCrae et al. (1996) onvermijdelijk tot verwerping van het model. Zij pleiten (p. 558) dan ook voor exploratieve fac-

toranalyse en replicatie op een tweede, onafhankelijke steekproef, eventueel aangevuld met procrustesrotatie op theoretische gronden. Ook pleiten zij voor een op theorie steunende aanname met betrekking tot het aantal te roteren factoren.

Op grond van bovenstaande overwegingen is gekozen voor een exploratieve principale componentenanalyse met varimax-rotatie op twee onafhankelijke datasets. Met een voldoende groot aantal variabelen ($N=52$) zal componentenanalyse vrijwel vergelijkbare resultaten opleveren als factoranalyse; bovendien zijn de bewerkingen helder en repliceerbaar (vgl. McCrae et al., 1996).

Op basis van het verloop van de eigenwaarden en op grond van theoretische overwegingen (Zaal (1978) onderscheidde vier gedragsclusters; beschrijvingen uit het intellectuele domein zijn op voorhand niet opgenomen) is gekozen voor rotatie naar vier factoren. Op beide steekproeven is een exploratieve componentenanalyse uitgevoerd, waarna gekeken is naar de mate van overeenstemming tussen beide factoroplossingen. Een grote overeen-

stemming impliceert repliceerbaarheid en ondersteunt de validiteit van de factorstructuur. Rotatie naar vier factoren leverde in beide analyses een duidelijke factorstructuur op, met hoge ladingen van de items op één van de vier factoren en lage ladingen op de overige (drie) factoren. De hoofdloadingen op de vier factoren zijn voor beide factoroplossingen te vinden in Tabel 1. Voor meer gedetailleerde informatie betreffende de bijladingen wordt verwezen naar Bleichrodt, Resing en Zaal (1993). Als maat voor de factoriële invariantie zijn Tuckers phi-coëfficiënten berekend. Deze coëfficiënten staan vermeld in Tabel 2.

Tabel 2
Phi-coëfficiënten ter bepaling van de mate van factoriële invariantie van de factormatrices van Vorm B van de SCHOBL-R voor twee onafhankelijke steekproeven (steekproef I (N=1296) en steekproef II (N=367))

Factoren	steekproef I/steekproef II
Extraversie	.99
Werkhouding	.98
Aangenaam Gedrag	.96
Emotionele Stabiliteit	.96

Als ondergrens voor het als vergelijkbaar mogen beschouwen van twee factoren wordt, in navolging van Ten Berge (1977), een phi-coëfficiënt van .85 gehanteerd. De overeenstemming tussen beide, onafhankelijk en zonder procrustesrotatie verkregen, factorstructuren is hoog, zelfs als de ondergrens tot .90 of hoger zou worden opgetrokken. Een procrustesrotatie op de tweede steekproef zou tot vrijwel dezelfde of nog hogere phi-coëfficiënten leiden. Uit het verloop van de eigenwaarden kan worden afgeleid dat voor beide steekproeven ongeveer 50 procent van de variantie wordt verklaard door deze vier factoren.

Gebaseerd op een inhoudelijke analyse van de items en rekening houdend met de terminologie van het Vijf-factoren model zijn de factoren als volgt benoemd:

I: *Extraversie*. De factorschaal Extraversie beschrijft de polariteit 'introvert - extravert'. Het 'extraverte' kind laat merken dat het er is, praat veel, is actief, toont initiatief, is vrolijk en openhartig, heeft zelfvertrouwen, is eigenwijs, domineert andere kinderen en is bijdehand. Aan de ene

kant het vrijmoedige, actieve, energieke kind en aan de andere kant het geredde, verlegen, gesloten en teruggetrokken kind.

II: *Werkhouding*. Representatief voor de factorschaal Werkhouding zijn gedragingen die betrekking hebben op het leren en werken op school: het kind 'met een positieve werkhouding' kan zich lang concentreren, prestatieniveau en -tempo zijn constant, het vertoont grote inzet bij het werk en werkt netjes. Daar tegenover staat het afleidbare, ongeïnteresseerde en zich slordig gedragende kind, dat nu eens goed en dan weer slecht presteert en dat vaak moet worden gewaarschuwd.

III: *Aangenaam Gedrag*. De gemeenschappelijke component van de beoordelingen die hoog laden op deze factorschaal wordt gevormd door de interactie met andere kinderen. De polariteit kent een 'goed - slecht' karakter. Plagerig en egoïstisch staan tegenover goedmoedig en vrijgevig, prikkelbaar en snel boos tegenover kalm en toegeeflijk, kritisch en egocentrisch tegenover instemmend en onzelfzuchtig.

IV: *Emotionele Stabiliteit*. Deze factorschaal beschrijft de emotionele aspecten van het gedrag. Hoge ladingen hebben uitspraken als 'snel in tranen', 'bezorgd om alles', 'maakt overal een probleem van' en 'snel van streek'. Tegenover elkaar staan het sentimentele en gevoelige kind dat snel bescherming zoekt bij de leerkracht en het stoere, ongevoelige en soms onverschillige kind dat zijn eigen boontjes wel zal doppen.

Nagegaan is of de factorstructuur stabiliteit vertoont over leeftijd. De betekenis-constantie van de SCHOBL-R is onderzocht door de uitkomsten van de principale componentenanalyses, die uitgevoerd zijn op de diverse leeftijds-subgroepen, onderling te vergelijken. Opnieuw zijn exploratieve principale componentenanalyses uitgevoerd voor drie leeftijdsgroepen: 4.2 - 6.2 jaar, 6.2 - 8.2 jaar en 8.2 - 11.2 jaar (steekproef I). Daarnaast zijn de scores voor jongens en meisjes apart geanalyseerd. Vervolgens is nagegaan in welke mate de gevonden factorstructuren gelijk blijven voor de drie leeftijdsgroepen en voorts voor de jongens- en meisjesgroep. Tabel 3 geeft, per factor, voor de

Tabel 3

Phi-coëfficiënten ter bepaling van de mate van factoriële invariantie van de factormatrices van Vorm B van de SCHOBL-R voor drie leeftijdsgroepen (4.2 - 6.2 (N=350), 6.2 - 8.2 (N=379) en 8.2 - 11.2 (N=567) en voor jongens (N=635) en meisjes (N=661)

Factoren	4.2-6.2 / 6.2-8.2	4.2-6.2 / 8.2-11.2	6.2-8.2 / 8.2-11.2	meisjes / jongens
Extraversie	.97	.96	.98	.99
Werkhouding	.98	.97	.99	.98
Aangenaam Gedrag	.96	.93	.98	.98
Emotionele Stabiliteit	.95	.92	.96	.97

diverse groepen afzonderlijk, een overzicht van de berekende Tuckers phi-coëfficiënten.

De uitkomsten wijzen op een grote mate van overeenstemming in factorstructuur voor de verschillende leeftijdsgroepen (.92 - .99) en vrijwel identieke factorstructuren voor jongens en meisjes (.97 - .99).

Ook op de 52 items uit parallelvorm A (afgenomen bij steekproef II) is een principale componentenanalyse met rotatie naar vier factoren uitgevoerd. In Tabel 2 staan de hoofdloadingen (per item) op de vier factoren vermeld. De A-vorm kent eenzelfde, inhoudelijk overeenkomende structuur als de B-vorm. De parallelvorm-betrouwbaarheid varieert van .88 voor Extraversie, .90 voor Werkhouding, .85 voor Aangenaam Gedrag en .71 voor Emotionale Stabiliteit. Gecorrigeerd voor attenuatie bedragen deze correlaties respectievelijk .97, .99, .99 en .91.

3.2 Interne consistentie en intercorrelaties tussen de factoren

Tabel 4 geeft een overzicht van de interne consistentie van de factorschalen, voor steekproef

Tabel 4

Coëfficiënten van interne consistentie (homogeniteitscoëfficiënten) voor de factorschalen van Vorm A en B en voor de totale SCHOBL-R

Factorschalen	Steekproef I		Steekproef II	
	Vorm B	Vorm A	Vorm B	Vorm A + B
	totaal N=1296	totaal N=358	totaal N=358	totaal N=358
Extraversie	.90	.91	.90	.95
Werkhouding	.90	.93	.90	.96
Aangenaam Gedrag	.89	.87	.86	.93
Emotionele Stabiliteit	.69	.82	.74	.88

I (vorm B) en voor steekproef II (vorm A, B en A plus B). In de Tabel zijn de homogeniteitscoëfficiënten weergegeven.

De gevonden coëfficiënten vertonen voor beide steekproeven een grote mate van overeenkomst. Voor jongens en meisjes (in de Tabel niet apart uitgesplitst) zijn de coëfficiënten vrijwel gelijk. Hetzelfde geldt voor de verschillende leeftijdsgroepen. Wanneer we de homogeniteiten voor de gehele SCHOBL-R in ogenschouw nemen, dan blijken de coëfficiënten voor de factoren Extraversie, Werkhouding en Aangenaam gedrag zeer hoog te zijn, respectievelijk .95, .96 en .93, terwijl de coëfficiënt voor de factor Emotionele Stabiliteit iets lager is (.88). Laatstgenoemde factor bevat echter het geringste aantal items. Per versie liggen de coëfficiënten iets lager (variërend van .70 tot .93). Met name de factor Emotionele Stabiliteit (met 8 items per versie) kent een iets lagere homogeniteitscoëfficiënt.

In Tabel 5 zijn, voor steekproef I (B-versie) de intercorrelaties tussen de factorschalen, voor drie verschillende leeftijdsgroepen vermeld (4.2 - 6.2 jaar (N=350); 6.2 - 8.2 jaar (N=379) en 8.2 - 11.2 jaar (N=567).

Emotionele Stabiliteit blijkt de meest onafhankelijke schaal te zijn. De hoogste correlaties worden gevonden tussen Aangenaam Gedrag enerzijds en Extraversie en Werkhouding anderzijds. De negatieve correlaties tussen Extraversie en Aangenaam Gedrag zijn vooral te verklaren uit het feit dat leerkrachten drukke en bazige kinderen vaak tevens als onbehouden en egoïstisch beoordelen. Het positieve verband tussen Werkhouding en Aangenaam Gedrag is vooral het gevolg van een overlap in beoordeelde gedragingen zoals ongehoorzaam en impulsief enerzijds en onbehouden en egoïstisch anderzijds. Uit Tabel 5 blijkt dat er

Tabel 5

Correlaties tussen de factorschalen van de B-vorm van de SCHOBL-R: 4.2 – 6.2 jaar (N=350), 6.2 – 8.2 jaar (N=379), 8.2 – 11.2 jaar (N=567)

Factorschalen	Extraversie			Werkhouding			Aangenaam Gedrag		
	4.2-6.2	6.2-8.2	8.2-11.2	4.2-6.2	6.2-8.2	8.2-11.2	4.2-6.2	6.2-8.2	8.2-11.2
Werkhouding	-.02	-.02	-.12						
Aangenaam Gedrag	-.51	-.32	-.41	.33	.48	.53			
Emotionele Stabiliteit	-.06	-.14	-.19	.19	.14	.07	.09	.11	.11

slechts geringe verschuivingen in het patroon van intercorrelaties optreden met verandering van leeftijd. Het verband tussen Extraversie en Aangenaam Gedrag is bij oudere kinderen iets minder sterk, terwijl het verband tussen Aangenaam Gedrag en Werkhouding juist iets toeneemt met leeftijd. Over het algemeen kan gesteld worden dat de intercorrelaties niet al te hoog zijn en dat de betrouwbaarheid van de vier schalen zo hoog is dat voldoende unieke variantie per factorschaal resteert.

3.3 Sekse-verschillen

Om na te gaan of jongens en meisjes andere scores behalen op de SCHOBL-R dan wel anders door leerkrachten worden beoordeeld op hun schoolgedrag, zijn – voor steekproef I – voor drie leeftijdsgroepen de gemiddelde ruwe scores op de vier factoren, uitgesplitst naar sekse, berekend. Op deze gegevens is vervolgens een multivariate variantie-analyse uitgevoerd met sekse en leeftijd als factoren. Er werd een significant sekse-effect gevonden voor de factorschalen Werkhouding, Aangenaam Gedrag en Emotionele Stabiliteit. Meisjes zijn positiever beoordeeld wat betreft hun werkhouding en sociale omgang en blijken over het algemeen volgens de leerkrachten emotioneler te reageren. Er werd geen significant leeftijdseffect gevonden bij een indeling in drie leeftijdsgroepen: 4.2 – 6.2 jaar, 6.2 – 8.2 jaar en 8.2 – 11.2 jaar. Er is geen sprake van een significante interactie tussen sekse en leeftijd. Op de gegevens van steekproef II is een vergelijkbare multivariate analyse uitgevoerd. De resultaten van deze analyse komen vrijwel overeen: er is geen sprake van een leeftijds-effect en wel van een sekse-effect. Op basis van deze resultaten is besloten aparte normen voor jongens en meisjes te berekenen. Het was niet nodig aparte leeftijdsnormen te ontwikkelen. Indien de gemiddelde scores op de vier

factoren vergeleken worden met de ruwe gemiddelde scores die jonge kinderen op de SCHOBL in het onderzoek van Zaal (1978) behaalden, kan worden geconcludeerd dat er in de loop van de tijd geen noemenswaardige verschuivingen zijn opgetreden.

3.4 Gedragsschalen

Om in praktijksituaties het gedrag van het kind zo goed en zo volledig mogelijk te kunnen beschrijven, is onderzocht of er binnen de vier hoofdfactoren van de SCHOBL-R inhoudelijk zinvolle en goed van elkaar te onderscheiden gedragsclusters aanwezig zijn. Hofstee en De Raad (1991) geven aan dat er binnen het Vijf-factoren model specifiekere gradaties van clusters van persoonlijkheidseigenschappen aanwezig zijn. Voor steekproef II, waarbij beide versies van de SCHOBL-R zijn afgenomen, is derhalve nagegaan of er een verdere onderverdeling in gedragsschalen mogelijk is binnen de vier hoofdfactoren. Om dit te kunnen onderzoeken is op de 104 items (parallelvorm A plus B) een hiërarchische clusteranalyse uitgevoerd. Op basis van deze analyse konden dertien gedragsclusters worden onderscheiden. Deze staan, geordend per factor, vermeld in Figuur 1. De gedragsschalen zijn voor de gebruiker uitsluitend te interpreteren indien zowel de A- als de B-vorm van de SCHOBL-R voor een kind is ingevuld.

Ook op de gedragsschalen worden verschillen in beoordeling tussen jongens en meisjes geconstateerd. De gemiddelde scores op de gedragsschalen die te maken hebben met Werkhouding (gehoorzaam, geconcentreerd, ambitieus en weloverwogen) en Aangepast Gedrag (voorkomend, onzelfzuchtig, meevolend) liggen hoger voor meisjes dan voor jongens, terwijl de gemiddelde scores voor de gedragsschalen die te maken hebben met

FACTORSCHALEN		GEDRAGSSCHALEN			
EV Extraversie	Og	Ongeremd	-	Terughoudend	
	Vp	Gesloten	-	Vrijpostig	
	Bz	Volgzaam	-	Bazig	
	Zo	Zelfonderschattend	-	Zelfoverschattend	
WH Werkhouding	Gh	Ongehoorzaam	-	Gehoorzaam	
	Gc	Ongeconcentreerd	-	Geconcentreerd	
	Ab	Gemakzuchtig	-	Ambitius	
	Wo	Impulsief	-	Weloverwogen	
AG Aangenaam Gedrag	Vk	Onbehouwen	-	Voorkomend	
	Oz	Egoïstisch	-	Onzelfzuchtig	
	Mv	Kil	-	Meevoelend	
ES Emotionele Stabiliteit	Ob	Onevenwichtig	-	Onbewogen	
	Ov	Zwaartillend	-	Onverschillig	

Figuur 1. Factor- en gedragsschalen uit de SCHOBL-R

Extraversie (ongeremd, vrijpostig, bazig, zelfoverschattend) en Emotionele Stabiliteit (onbewogen, onverschillig) juist voor jongens hoger liggen.

Een 2e-orde principale componentenanalyse op de scores op de gedragsschalen (geanalyseerd over beide parallelvormen) laat een duidelijk te interpreteren vier-factorpatroon zien. Dit factorpatroon is, voor jongens en voor meisjes, vrijwel identiek. Voor de totale groep staat dit ladingenpatroon vermeld in Tabel 6. Voor zowel jongens als voor meisjes bedraagt de verklaarde variantie ongeveer 85 procent. De intercorrelaties tussen de gedragsschalen die tot een factor gerekend kunnen worden zijn over het algemeen vrij hoog, maar voor prak-

tisch gebruik kan het gedrag van het individuele kind op inhoudelijk zinvolle wijze in fijnere gedragsclusters worden beschreven.

De vier gedragsschalen, die samen vrijwel overeenkomen met de factor Extraversie, hebben hoge ladingen op deze factor, variërend van .75 tot .88. Gesloten – Vrijpostig is de meest factor-zuivere schaal, de overige hangen, weliswaar in geringe mate, ook samen met één van de andere factoren. Ook de vier gedragschalen, die samen vrijwel geheel de factor Werkhouding vormen, hebben hoge ladingen op deze factor, variërend van .65 tot .92. De meest zuivere representanten van deze factor zijn de gedragsschalen Ongeconcentreerd – Geconcentreerd en Gemakzuchtig – Ambitius. De gedragsschaal Impulsief – Weloverwogen ligt op het grensgebied tussen Werkhouding en Extraversie met vrij hoge ladingen op beide factoren. De drie gedragsschalen binnen Aangenaam Gedrag hebben hoge ladingen op deze factor, maar Onbehouwen – Voorkomend heeft een nevenlading op de factor Extraversie. Dit wordt in belangrijke mate veroorzaakt door de overeenkomst in te beoordelen gedrag uit de schaal Terughoudend – Ongeremd. De beide gedragsschalen binnen Emotionele Stabiliteit laden hoog op deze factor. Met name Zwaartillend – Onverschillig heeft nog (lage) nevenladingen op de factoren Werkhouding en Aangenaam Gedrag.

Tabel 6
2e-orde principale componentenanalyse (varimaxrotatie naar vier factoren) op de scores op de 13 gedragsschalen van de SCHOBL-R

Gedragsschalen	Totale groep			
	EV	WH	AG	ES
Ongeremd	.85	.41	.13	-.07
Vrijpostig	.85	.19	-.23	-.14
Bazig	.81	-.09	.35	-.12
Zelfoverschattend	.75	-.17	.13	.37
Gehoorzaam	.33	.74	.39	.13
Geconcentreerd	.07	.92	.16	.06
Ambitius	-.10	.84	.21	-.26
Weloverwogen	.55	.65	.37	.03
Voorkomend	.59	.37	.61	-.01
Onzelfzuchtig	.27	.22	.87	
Meevoelend	-.17	.38	.78	.18
Onbewogen	-.24	-.28	-.11	.85
Onverschillig	.20	.35	.32	.79

Tabel 7

Relatie beoordelingen van schoolgedrag (Vorm B van de SCHOBL-R) met intelligentie (RAKIT) voor vier leeftijdsgroepen: 4.2 - 5.2 (N=172), 5.2 - 6.2 (N=178), 6.2 - 8.2 (N=379) en 8.2 - 11.2 (N=567)

RAKIT-intelligentie	SCHOBL-R factoren			
	Extraversie	Werkhouding	Aangenaam gedrag	Emotionele Stabiliteit
<i>4.2 - 5.2 jaar</i>				
RAKIT totaal IQ	.20**	.33**	-.06	.01
RAKIT verkort IQ	.23**	.32**	-.06	-.03
Verbaal leren en vlotheid	.17*	.19*	-.04	-.03
Ruimtelijk perceptueel redeneren	.14	.37**	-.00	.06
Sequentieel geheugen	.17*	.14	-.16	-.06
Kwantiteit	.07	.14	-.03	.06
<i>5.2 - 6.2 jaar</i>				
RAKIT totaal IQ	.15*	.48**	.05	.14
RAKIT verkort IQ	.12	.43**	.05	.10
Perceptueel redeneren	.04	.46**	.07	.13
Verbaal leren	.16*	.28**	-.02	.05
Ruimtelijke oriëntatie en tempo	.20**	.38**	-.00	.13
Verbale vlotheid	.22**	.25**	.01	.08
<i>6.2 - 8.2 jaar</i>				
RAKIT totaal IQ	.19**	.38**	-.02	.11*
RAKIT verkort IQ	.19**	.36**	-.02	.09
Perceptueel redeneren	.11*	.38**	.04	.09
Verbaal leren	.08	.25**	.01	.14**
Ruimtelijke oriëntatie en tempo	.25**	.25**	-.12*	.03
Verbale vlotheid	.31**	.09	-.15**	-.01
<i>8.2 - 11.2 jaar</i>				
RAKIT totaal IQ	.17**	.28**	.06	.07
RAKIT verkort IQ	.16**	.26**	.08	.08
Perceptueel redeneren	.06	.27**	.10*	.08
Verbaal leren	.13**	.24**	.06	.10*
Ruimtelijke oriëntatie en tempo	.21**	.11**	-.02	.00
Verbale vlotheid	.26**	.03	.12**	-.05

** significant op 1% niveau; * significant op 5% niveau.

3.5 De relatie met intelligentie en schoolprestaties

Ter validering van het ontwikkelde instrument is de relatie met intelligentie enerzijds en met prestaties op school anderzijds onderzocht. Verwacht wordt dat werkhouding een gematigde samenhang vertoont en dat de overige factoren weinig samenhang zullen laten zien met de beide criteria. Immers, bij de selectie van de oorspronkelijke items zijn gedragingen uit het intellectuele domein bewust niet opgenomen. Voor steekproef I werden de relaties tussen deze criteria en de vier factoren nader onderzocht. In Tabel 7 zijn de gevonden correlaties weergegeven. In de Tabel zijn vier leeftijdsgroepen onderscheiden aangezien de RAKIT-factorstructuur voor de jongste kinderen enigszins afwijkt van die voor de oudere groepen.

De relaties tussen sociaal-emotioneel gedrag en intelligentie vertonen een redelijk consistent patroon over de vier gepresenteerde leeftijdscategorieën. De als meer extravert beoordeelde kinderen zijn gemiddeld iets intelligenter dan hun klasgenoten die als minder extravert worden beoordeeld. Met name geldt dit voor de prestaties op de Verbale Vloheidsfactor. Ook Ruimtelijke Oriëntatie en Tempo en in mindere mate Verbaal Leren vertonen een positieve correlatie met Extraversie. De SCHOBL-R factor die, zoals verwacht, de hoogste correlaties vertoont met intelligentie is de factor Werkhouding. Dit is het geval voor zowel het totale RAKIT IQ, als voor de afzonderlijke intelligentiefactoren met uitzondering van Verbale Vlotheid voor de beide oudste groepen kinderen. Voor de beide andere factoren is nauwe-

Tabel 8

Relatie van beoordelingen van schoolgedrag (SCHOBL-R) met CITO toets resultaten voor zes leeftijds-groepen

Leeftijd	Aantal	CITO toetsen	SCHOBL-R factoren			
			Extraversie	Werkhouding	Aangenaam Gedrag	Emotionele Stabiliteit
5.2- 6.2	156	Begrippentoets 1	.00	.39**	.14	.10
		Begrippentoets 2	.07	.25**	.09	.07
6.2-7.2	168	Technisch lezen	-.07	.31**	.12	.10
7.2-8.2	164	Lees en Begrijp	.07	.20*	.00	.07
8.2-9.2	149	Begrijpend lezen 3	.09	.38**	.04	.13
		Rekenen 3	.02	.32**	.07	.20**
9.2-10.2	150	Begrijpend lezen 4	-.01	.18*	.09	.04
		Rekenen 4	.01	.32**	.09	.07
10.2-11.2	146	Begrijpend lezen 5	.11	.20*	-.03	-.04
		Rekenen 5	.22**	.20*	-.02	-.07

** significant op 1% niveau; * significant op 5% niveau

lijks een verband met intelligentie gevonden.

De leerkrachten hebben de schoolprestaties van de kinderen op een aantal aspecten beoordeeld op 7-puntsschalen. Voor kinderen van 4.2 tot 6.2 jaar werd gevraagd de kinderen onder andere te beoordelen op hun woordenschat, taalbeheersing, voorbereidend lezen en rekenen, logisch redeneren, et cetera. Voor de oudere kinderen werd onder andere gevraagd deze te beoordelen op hun reken-, taal-, lees-, en totaalprestaties. De belangrijkste correlaties zijn wederom die tussen Werkhouding en schoolprestatie-beoordelingen. Voor de jongste kinderen variëren de correlaties tussen de gevraagde schoolprestaties en Werkhouding tussen .38 en .50, met uitzondering van Grove Motoriek (.23). Aangenaam Gedrag en Emotionele Stabiliteit vertonen voor deze leeftijdsgroep geen verband met schoolprestatie-beoordelingen, de correlaties met Extraversie zijn zeer gematigd (rond .15), met uitzondering van Woordenschat (.27) en Taalbeheersing (.26). Voor de beide oudste leeftijdsgroepen is het beeld vergelijkbaar. Werkhouding correleert tussen .45 en .57 met de beoordeelde schoolprestaties, Taal, Begrijpend Lezen, Rekenen en Totaal schoolprestaties, terwijl de correlaties met de factor Extraversie aanzienlijk lager zijn (rond .15). Alleen voor de oudste groep is een licht positief verband gevonden van rond .15 tussen beoordelingen van schoolprestaties en Emotionele Stabiliteit.

Om de samenhang tussen het sociaal-emotionele schoolgedrag en schoolprestaties niet

alleen op het niveau van het leerkrachtoordeel, maar ook op het objectieve niveau van schooltoetsen te kunnen onderzoeken, werden bij de kinderen één of meer, voor de groep waarin ze zaten representatieve, CITO-toetsen afgenomen. De gevonden verbanden staan vermeld in Tabel 8.

Alleen de factor Werkhouding vertoont nu een betekenisvol verband met de prestaties op de CITO-toetsen. Opvallend is dat de correlaties uit Tabel 8 niet alleen lager zijn in vergelijking tot de correlaties tussen de sociaal-emotionele gedragsfactoren en leerkrachtbeoordelingen, maar ook in vergelijking tot de correlaties met de intelligentiegegevens. Opmerkelijk is verder dat de eerder vermelde verbanden tussen Lezen en Extraversie en, in mindere mate tussen Rekenen en Aangenaam Gedrag, hier niet teruggevonden worden.

4 Conclusies en discussie

De SCHOBL-R heeft een duidelijke onderliggende factorstructuur die stabiel blijft over het leeftijdsbereik van 4 tot 11 jaar. Voor de factoren Extraversie, Werkhouding en Aangenaam Gedrag zijn bijna alle factorladingen hoger dan .50 en de meeste zelfs hoger dan .60. Alleen voor de factor Emotionele Stabiliteit zijn de factorladingen over het algemeen iets lager. De vier onderscheiden hoofdfactoren zijn inhoudelijk goed te interpreteren en komen overeen

met de eerste vier factoren uit het Vijf-factoren model. De vijfde factor (Ideeënrijk, Intellect, het Openstaan voor ervaringen, Creativiteit en Autonomie) wordt niet teruggevonden, hetgeen in overeenstemming is met de verwachting. Bij de constructie van de vragenlijst was een expliciet uitgangspunt dat sociaal-emotioneel gedrag, voor zover dat niet binnen het intellectuele domein ligt, gemeten diende te worden. Beschrijvingen over hoe intelligent, hoe creatief of hoe slim een kind in de klas is zijn dan ook van aanvang af niet in de itemlijsten voorgekomen.

De hoofdfactoren zijn onderling wel gecorrelleerd, maar dit wordt in de literatuur over persoonlijkheidsbeschrijvingen vrijwel altijd teruggevonden en is dus in overeenstemming met de verwachtingen. Voor de gedragsschalen geldt dit nog in sterkere mate. Deze schalen zijn tot stand gekomen op basis van clusteranalyse en met name die clusters die grenzen aan het gebied van een andere factor hangen onderling vrij sterk samen. Toch is een fijner onderscheid in te beschrijven gedrag inhoudelijk zinvol en voor de praktijk goed bruikbaar.

De SCHOBL-R heeft een zeer goede betrouwbaarheid, zeker als enige tijd na elkaar zowel de A- als de B-vorm worden afgenomen. Ook de parallelvorm betrouwbaarheid is hoog. De scores op de gedragsschalen zijn voor individuele kinderen uitsluitend te berekenen indien beide parallelvormen in combinatie worden afgenomen. Voor het berekenen van de hoofdfactoren is dit niet per se nodig.

De relaties tussen sociaal-emotioneel gedrag en intelligentie vertonen een redelijk consistent patroon voor de vier verschillende leeftijdsgroepen. De meer extraverte kinderen worden als gemiddeld intelligenter beoordeeld dan hun introverte klasgenoten. Dit geldt vooral voor de Verbale Vlotheidsfactor. Dit is in overeenstemming met de verwachting; de tests die hoog laden op de factor Verbale Vlotheid vereisen dat het kind in korte tijd veel zegt (veel woorden noemt bij Ideeënproductie; een uitvoerig verhaal vertelt bij de Vertelplaat). Uit nadere analyses blijkt dat het verband vooral veroorzaakt wordt door de hoge correlatie tussen Extraversie en Ideeënproductie. Ook Ruimtelijke Oriëntatie en Tempo en in mindere mate Verbaal Leren vertonen een positieve samenhang met Extraversie.

De hoogste verbanden tussen intelligentie en sociaal-emotioneel functioneren worden, zoals verwacht, gevonden voor de factor Werkhouding. Dit geldt zowel voor het algemene niveau van cognitief functioneren, het totale RAKIT IQ, als de afzonderlijke intelligentiefactoren met uitzondering van Verbale Vlotheid. De factor Perceptueel Redeneren, waarin ook een deel logisch redeneren zit, vertoont de hoogste samenhang. Geconcludeerd kan worden dat het vrijmoedige kind over het algemeen betere prestaties zal leveren op verbale vlotheidstaken en, in mindere mate, op ruimtelijke oriëntatie en dat kinderen met een positieve werkhouding gemiddeld beter zullen presteren op taken die te maken hebben met perceptueel redeneren, verbaal leren en ruimtelijke oriëntatie.

Schoolprestaties en schoolprestatiebeoordelingen blijken vooral samen te hangen met de factor Werkhouding. Een positieve werkhouding gaat gepaard met beter presteren op school. Opvallend is dat de correlaties tussen leerkrachtbeoordelingen van schoolprestaties en Werkhouding hoger zijn dan de correlaties tussen werkelijke toetsprestaties en Werkhouding, zelfs in vergelijking met de correlaties tussen Werkhouding en intelligentiefactoren. De sterke samenhang tussen beoordelingen van schoolprestaties en beoordelingen van Werkhouding kan wellicht voor een deel verklaard worden door het feit dat beide beoordelingen van een leerling in de meeste gevallen door dezelfde leerkracht zijn gegeven. Maar dit geldt vanzelfsprekend niet voor de relatie tussen Werkhouding, zoals beoordeeld door de leerkracht, en RAKIT-scores.

Over leeftijden heen blijft de factorstructuur, in de vorm van de vier hoofdfactoren vrijwel identiek. Ook voor jongens en meisjes wordt een vrijwel overeenkomstig patroon van factoren gevonden. De verschillen in gemiddelde SCHOBL-R scores zijn bovendien vrijwel gelijk voor de diverse leeftijdsgroepen. Sekse-verschillen in beoordeeld schoolgedrag blijken stabiel over leeftijd te zijn. Dergelijke gegevens lijken erop te wijzen dat leerkrachten bij het beoordelen van het sociaal-emotionele gedrag van de leerling mogelijk vooral het gemiddelde kind in het basisonderwijs als referentiekader nemen.

De verschillen in gemiddelde SCHOBL-R

scores tussen jongens en meisjes zullen voor een belangrijk deel berusten op verschillen binnen het klasseverband. Jongens en meisjes worden door de leerkrachten duidelijk anders beoordeeld. Het is op grond van de hier gevonden resultaten niet mogelijk aan te geven of dit helemaal toe te schrijven valt aan ander gedrag dat jongens en meisjes laten zien, of dat leerkrachten op sekse stereotyperende wijze meisjesgedrag anders beoordelen dan vergelijkbaar jongensgedrag. Uit de literatuur blijkt overigens vaak dat meisjes als ijveriger, consciëntieuzer, prettiger in de omgang en emotioneler worden beoordeeld dan jongens.

Deze groepsverschillen in gemiddelde gedragsbeoordelingen tussen jongens en meisjes roepen derhalve onvermijdelijk de vraag op of er sprake is van verschillen in daadwerkelijk gedrag of van gemeenschappelijke beeldvorming (sekse stereotypering) bij de beoordelen de leerkrachten. Dit is een onvermijdelijk dilemma aangezien het onderhavige onderzoek nu eenmaal uitgaat van beoordeeld gedrag en niet van objectieve registratie van gedrag. Dit blijft voldoende reden om de verschillen tussen de groepen met de nodige terughoudendheid te interpreteren. Die mogelijkheid wordt de gebruiker geboden omdat er aparte normtabellen voor jongens en meisjes zijn ontwikkeld.

Er is echter, ondanks het feit dat er van beoordelingen sprake is, voldoende bewijs voor de betrouwbaarheid, objectiviteit en validiteit van de leerkrachtbeoordelingen. Zaal (1978) heeft in het verleden een goede overeenstemming aangetoond tussen beoordelingen van verschillende leerkrachten, het herhaaldelijk beoordelen van een en hetzelfde kind door dezelfde leerkracht, en dergelijke. Verder zijn er verschillen gevonden tussen beoordelingen van basisschoolleerlingen en van leerlingen uit het LOM-onderwijs, waarvan verwacht mag worden dat zij afwijkend gedrag zullen vertonen aangezien er vaak sprake is van ernstige sociaal-emotionele gedragsproblematiek al dan niet in combinatie met lage schoolprestaties (vgl. Bleichrodt, Resing & Zaal, 1993; Resing en Bleichrodt (in voorbereiding)). De leerkrachten blijken dus niet zonder meer altijd uitsluitend de eigen klas als referentiekader te gebruiken bij hun beoordelingen.

De gevonden sekseverschillen komen overeen met een algemeen verwachtingspatroon.

Tegelijkertijd zien we echter ook dat beoordelingen van Werkhouding positief samenhangen met objectief gemeten schoolprestaties en intelligentiescores. Verder onderzoek naar de relatie tussen beoordelingen en objectief gemeten gedragscriteria blijft wenselijk. Mogelijkheden daartoe kunnen vanzelfsprekend gevonden worden door het objectief scoren van gedrag in de klas, al of niet met behulp van video-opnamen. Tevens is het van belang de validiteit van de SCHOBL-R nader te onderzoeken door de samenhang vast te stellen met meer klinische gedragsvragenlijsten, zoals de CBCL, en met oudervragenlijsten binnen het sociaal-emotionele gedragsdomein die in ontwikkeling zijn (Elphick, Slotboom & Kohnstamm, 1995, 1997).

Literatuur

- Anthony, E.J. (1970). The behavior disorders of children. In P.H. Mussen (Ed.), *Carmichael's manual of child psychology. Vol. II* (pp. 667-764). New York: Wiley.
- Becker, W.C., & Krug, R.S. (1964). A circumplex model for social behavior in children. *Child Development, 35*, 371-396.
- Berge, J.M.F. ten (1977). *Optimizing factorial invariance*. Groningen: VRB Drukkerijen.
- Bleichrodt, N., Drenth, P.J.D., Zaal, J.N., & Resing, W.C.M. (1987). *Revisie Amsterdamse Kinder Intelligentie Test*. Handleiding. Lisse: Swets & Zeitlinger.
- Bleichrodt, N., & Resing, W.C.M. (in voorbereiding). Boordeling van schoolgedrag bij allochtone leerlingen in het basisonderwijs.
- Bleichrodt, N., Resing, W.C.M., & Zaal, J.N. (1993). *Beoordeling van schoolgedrag: SCHOBL-R*. Lisse: Swets & Zeitlinger.
- Blöte, A.W., & Curfs, L.M.G. (1986). Het gebruik van de Conners Teacher Rating Scale in Nederland: enige psychometrische gegevens. *Nederlands Tijdschrift voor de Psychologie, 41*, 226-236.
- Bookstein, F.L. (1986). The elements of latent variable models: a cautionary lecture. In M.E. Lamb, A.L. Brown & B. Rogoff (Eds.), *Advances in developmental psychology (Vol. 4)* (pp. 203-230). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cattell, R.B. (1957). *Personality and motivation. Structure and measurement*. New York: Harcourt, Brace & World.

- Cattell, R.B., & Coan, R.A. (1957). Child personality structure as revealed in teachers' ratings. *Journal of Clinical Psychology, 13*, 315-327.
- Digman, J.M. (1963). Principal dimensions of child personality as inferred from teacher's judgments. *Child Development, 34*, 43-60.
- Digman, J.M. (1989). Five robust trait dimensions: Development, Stability, and Utility. *Journal of Personality, 27*, 195-214.
- Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.
- Digman, J.M. (1994). Child personality and temperament: Does the Five-factor model embrace both domains? In C.F. Halverson, G.A. Kohnstamm & R.P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 323-338). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Digman, J.M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50*, 116-123.
- Elphick, E., Slotboom, A., & Kohnstamm, G.A. (1995). *Parental expertise on individual differences between children: Construction of the questionnaire "Beoordelings-Lijst Individuele verschillen tussen Kinderen (BLIK)".* Interne publicatie Vakgroep Ontwikkelings- en Onderwijspsychologie. Leiden: RU Leiden.
- Elphick, E., Slotboom, A., Kohnstamm, G.A. (1997). Persoonlijkeitsbeoordelingen door ouders van jonge adolescenten: een onderzoek naar de generaliseerbaarheid van het vijf-factor model. *Nederlands Tijdschrift voor de Psychologie, 52*, 151-162.
- Elshout, J.J., & Akkerman, A.E. (1975). *Vijf persoonlijkheidsfactoren test, 5 PFT.* Nijmegen: Berkhout.
- Goldberg, L.R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*, 26-34.
- Guilford, J.P. (1959). *Personality.* New York: McGraw-Hill.
- Halverson, C.F., Kohnstamm, G.A., & Martin, R.P. (Eds.) (1994). *The developing structure of temperament and personality from infancy to adulthood.* Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hofstee, W.K.B. (1992). *Persoonlijkheidseigenschappen: de caleidoscoop van individuele verschillen.* Rede Vrije Universiteit Brussel. Faculteit voor Psychologie en Opvoedkundige Wetenschappen.
- Hofstee, W.K.B., & Raad, B. de (1991). Persoonlijkheidsstructuur: de ABC taxonomie van Nederlandse eigenschapstermen. *Nederlands Tijdschrift voor de Psychologie, 46*, 262-274.
- John, O.P., Caspi, A., Robins, R.W., Moffitt, T.E., & Stouthamer-Loeber, M. (1994). The "little five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development, 65*, 160-178.
- Jong, P.F. de, & Das-Smaal, E.A. (1991). De Amsterdamse Kindergedragslijst: een korte gedragsbeoordelingslijst voor kinderen. *Nederlands Tijdschrift voor de Psychologie, 46*, 76-83.
- Kohnstamm, G.A. (1992a). Ontwikkelingen in de persoonlijkheidspsychologie: de opkomst van de Big Five. *Pedagogische Studiën, 69*, 3-11.
- Kohnstamm, G.A. (1992b). Factoren in gedragsbeoordelingen van leerlingen. *Pedagogische Studiën, 69*, 12-22.
- Kohnstamm, G.A., Halverson, C.F., Havill, V.L., & Mervielde, I. (1996). Parent's free descriptions of child characteristics: a cross-cultural search for the developmental antecedents of the Big Five. In S. Harkness & C.M. Super (Eds.), *Parents' cultural belief systems: Their origins, expressions, and consequences* (pp. 27-55). New York: Guilford Press.
- Kohnstamm, G.A., Mervielde, I., Besevegis, E., & Halverson, C.F. (1995). Tracing the Big Five in parents' free descriptions of their children. *European Journal of Personality, 9*, 283-304.
- Lieshout, C.F.M. van, & Haselager, G.T.J. (1992). Persoonlijkheidsfactoren in Q-sort persoonsbeschrijvingen van kinderen: relatie tot het vijf-factor model. *Pedagogische Studiën, 69*, 23-39.
- McCrae, R.R., & Costa, P.T. (1985). Updating Norman's adequate taxonomy: Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology, 49*, 710-721.
- McCrae, R.R., Zonderman, A.B., Costa, P.T., Bond, M.H., & Paunonen, S.V. (1996). Evaluating replicability of factors in the revised NEO personality inventory: Confirmatory factor analysis versus procrustes rotation. *Journal of Personality and Social Psychology, 70*, 552-566.
- Mervielde, I. (1994). A five-factor model classification of teachers' constructs on individual differences among children ages 4 to 12. In C.F. Halverson, G.A. Kohnstamm & R.P. Martin (Eds.), *The developing structure of temperament and*

- personality from infancy to adulthood* (pp. 387-397). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Norman, W.T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, 66, 574-583.
- Raad, B. de, Hendriks, A.A.J., & Hofstee, W.K.B. (1992). Towards a refined structure of personality traits. *European Journal of Personality*, 6, 301-319.
- Resing, W.C.M., & Bleichrodt, N. (in voorbereiding). Beoordeling van schoolgedrag bij bijzondere groepen leerlingen: het gebruik van de SCHOBL-R in het speciaal onderwijs.
- Stott, L.H. (1962). Personality at age four. *Child Development*, 33, 287-311.
- Stott, L.H. (1967). *Child development. An individual longitudinal approach*. New York: Holt, Rinehart & Winston.
- Verhulst, F.C., Koot, J.M., Akkerhuis, G.M., & Veerman, J.W. (1990). *Praktische handleiding voor de CBCL*. Assen/Maastricht: Van Gorcum.
- Zaal, J.N. (1978). *Sociaal-emotioneel gedrag in de klas*. Academisch proefschrift. Groningen: Wolters-Noordhoff.
- Zaal, J.N. (1980). Relaties tussen sociaal-emotioneel functioneren en schoolprestaties van leerlingen in kleuter- en basisonderwijs. *Pedagogische Studiën*, 57, 257-267.
- Zaal, J.N. (1981). *Schoolgedrag Beoordelingslijst, SCHOBL*. Groningen: Wolters Noordhoff.

Manuscript aanvaard 22-5-1997

Auteurs

W.C.M. Resing is universitair docent op het terrein van de cognitieve ontwikkeling en de psychodiagnostiek en werkzaam bij de sectie Ontwikkelings- en Onderwijspsychologie van de Rijksuniversiteit te Leiden

N. Bleichrodt is bijzonder hoogleraar op het terrein van de alloctonen problematiek in arbeidssituaties en werkzaam bij de vakgroep Arbeids- en Organisationspsychologie van de Vrije Universiteit te Amsterdam

Correspondentie-adres: W.C.H. Resing, Faculteit Sociale Wetenschappen, Vakgroep Psychologie,

sectie Ontwikkelings- en Onderwijspsychologie, Wassenaarseweg 52, 2333 AK, Leiden
E-mail: Resing@rulfsw.leidenuniv.nl

Abstract

School behaviour measured again

W.C.M. Resing & N. Bleichrodt. Pedagogische Studiën, 1997, 74, 313-327.

The central questions in this article focus on the factorial structure of a revised instrument measuring school behaviour (SCHOBL-R) in children from primary schools, the stability of this instrument over age, the possibility to describe the school behaviour of children in concrete behavioral clusters, the sex-specificity of the instrument, and the relationship of the SCHOBL-R with intelligence scores and school performances. The studies reported here indicate that the SCHOBL-R has a clear four factor structure which remains stable for children between four and eleven years old. The four components, Extraversion, Attitude towards School Work, Agreeableness, and Emotional Stability correspond to four of the five factors of the Big Five. Administering both parallel versions of the SCHOBL-R provides teachers, psychologists and others with the opportunity to diagnose and describe in more detail (using 13 behavioral clusters) the school behaviour of children from primary schools.

Leren lesgeven met het directe instructiemodel op de tweedegraads lerarenopleiding

S. Veenman, J. Bakermans, Y. Franzen, M. van Hoof*

Samenvatting

Om de implementatie-effecten te bepalen van een programma over effectieve instructie voor de tweedegraads lerarenopleiding is een quasi-experimenteel onderzoek uitgevoerd. Vóór en na de training zijn bij de studenten observaties verricht. Zij blijken na het volgen van het trainingsprogramma de geselecteerde instructievaardigheden over het algemeen beter uit te voeren dan vóór het programma. Het programma blijkt geen invloed te hebben op het taakgerichte gedrag van de leerlingen. Volgens het oordeel van de mentoren passen de getrainde studenten de gewenste instructievaardigheden niet beter toe dan de niet-getrainde studenten. Het programma wordt door de studenten en docenten positief beoordeeld.

1 Achtergrond

Door de huidige cognitieve visie op leren – waarbij leren opgevat wordt als een actief, constructief, cumulatief en doelgericht proces (Shuell, 1988) – wordt bij het leren op school meer dan voorheen de nadruk gelegd op de verantwoordelijkheid van de leerling voor zijn eigen leren. Deze opvatting vinden we bijvoorbeeld terug in de nota van de Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs (1994) als gesproken wordt over de school als studiehuis. De leraar zal volgens deze gedachte meer

begeleider dan instructeur zijn. Het actief en constructief verwerven van kennis en vaardigheden door de leerlingen dient begeleid te worden door aangepaste vormen van instructie, uitleg, hulp en ondersteuning van de leraar. Wanneer de leerling nog over weinig voorkennis beschikt of wanneer het gaat om complexe problemen zal de leraar in de beginfase van het onderwijs de leeractiviteiten van de leerlingen op een meer directe wijze aansturen, om vervolgens via de fase van de gedeelde sturing en afnemende ondersteuning de verantwoordelijkheid voor het leren steeds meer in handen te leggen van de leerlingen.

Een beproefde instructie-aanpak voor de aansturing van de leeractiviteiten in de beginfase van het onderwijs is het directe instructiemodel. Dit model is door de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen en het Christelijk Pedagogisch Studiecentrum (CPS) als uitgangspunt gekozen voor een trainingsprogramma over effectieve instructie voor de tweedegraads lerarenopleiding (voorheen de NLO's). Hierbij is gebruik gemaakt van de ervaringen van een eerder voor de PABO ontworpen programma (zie Veenman, Leenders, Meyer & Sanders, 1993). In beide programma's staat het directe instructiemodel centraal. Dit model bestaat uit de volgende hoofdfasen: 1) het geven van een terugblik; 2) uitleg van nieuwe leerstof; 3) begeleide inoefening; 4) zelfstandige verwerking; 5) het geven van terugkoppeling en het corrigeren van fouten en 6) periodieke herhaling van de leerstof. Deze fasen zijn uitgewerkt in twee deelmodellen: een model voor uitvoerend handelen en een model voor strategisch handelen (cf. Veenman, 1992). Het model voor uitvoerend handelen 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. Bijvoorbeeld, het

* Dit onderzoek werd mede mogelijk gemaakt door het Christelijk Pedagogisch Studiecentrum (CPS) te Amersfoort. De auteurs bedanken de volgende personen voor hun bijdrage aan deze studie: G. de Boer (CPS), Y. Leenders (CPS), A. v.d. Horst (CPS), L. Bouts (RTD KUN) en de deelnemende studenten, docenten en mentoren van de tweedegraads lerarenopleiding.

aanleren van rekenkundige algoritmen of spel-lingeregels. De leerstof die hierbij gebruikt wordt, laat zich meestal goed structureren en opdelen in kleine stappen. In het model voor strategisch handelen neemt het aanleren van leerstrategieën een belangrijke plaats in. Ook wordt aandacht besteed aan metacognitieve vaardigheden en zelfregulatie-activiteiten van leerlingen. Dit model is vooral geschikt voor kennisgebieden en vaardigheden die zich minder makkelijk laten structureren, waar de kennis of vaardigheden niet stapsgewijs aangeleerd kunnen worden. Voorbeelden hiervan zijn tekstverklaren en het oplossen van wiskundige problemen. Voor een verdere beschrijving van beide modellen en hun uitgangspunten wordt verwezen naar de PABO-studie (Veenman e.a., 1993).

2 Onderzoeksvragen

Het onderzoek kent de volgende vragenstellingen: 1) passen de studenten de in het trainingsprogramma uitgewerkte instructievaardigheden toe in de lessen die zij op de stageschool geven? 2) heeft implementatie van de gekozen instructievaardigheden invloed op het taakgericht gedrag van de leerlingen van de stageschool? en 3) hoe hebben de studenten en docenten het werken met het trainingsprogramma ervaren?

3 Opzet van het onderzoek

Het onderzoek bestond uit drie deelonderzoeken: a) een observatie-onderzoek waarbij gebruik werd gemaakt van observatoren en dat gericht was op de toepassing van de vaardigheden uit het trainingsprogramma door de studenten op de stageschool en op de registratie van het taakgericht gedrag van de leerlingen (vraagstelling 1 en 2), b) een observatie-onderzoek waarbij gebruik werd gemaakt van de beoordeling van de mentoren van de studenten en dat eveneens gericht was op de toepassing van de vaardigheden uit het trainingsprogramma op de stageschool (vraagstelling 1) en c) een deelstudie met interviews en vragenlijsten dat gericht was op de waardering van het trainingsprogramma door de studenten en de do-

centen (vraagstelling 3). Het observatie-onderzoek werd opgezet volgens een quasi-experimenteel 'pretest-posttest control group design'.

4 Methode en instrumenten

4.1 De onderzoeksgroep

De onderzoeksgroep bestond uit tweede- en derdejaars studenten van drie tweedegraads lerarenopleidingen. De opleiding op locatie A lag in het zuiden van het land, op locatie B in het oosten en op locatie C in het zuid-oosten. Van locatie A werd één klas geselecteerd voor deelname aan het onderzoek, van locatie B en C ieder twee klassen. Uit deze vijf klassen volgden 64 studenten het programma over directe instructie (experimentele groep) en 71 studenten uit parallelklassen volgden de normale lessen didactiek of onderwijspsychologie (controlegroep). Het trainingsprogramma werd op elke locatie gegeven door een docent Onderwijskunde. Deze gaf ook les aan de parallelklassen. Uit de experimentele klassen werden voor het observatie-onderzoek 27 studenten geselecteerd (locatie A: 14 studenten, locatie B: 6 studenten; locatie C: 7 studenten). Uit de controlegroep werden 24 studenten geselecteerd (locatie A: 10 studenten; locatie B: 4 studenten; locatie C: 10 studenten).

Vóór en na afloop van het trainingsprogramma werd aan de mentoren gevraagd de verrichtingen van de studenten te beoordelen. Van de 51 verspreide vragenlijsten tijdens de voormeting werden er 33 door de mentoren ingevuld (experimentele groep $n = 18$, controlegroep $n = 15$). Tijdens de nameting werden 39 vragenlijsten teruggestuurd (experimentele groep $n = 19$, controlegroep $n = 20$). De gemiddelde response bedroeg 71%.

De vragenlijsten die gebruikt werden voor het verzamelen van informatie over de waardering van het trainingsprogramma, werden toegestuurd aan 64 studenten die het programma hadden gevolgd. In totaal werden 44 vragenlijsten teruggestuurd (response = 69%). Bij de drie betrokken docenten werden interviews afgenomen om hun opvattingen over het trainingsprogramma te achterhalen.

4.2 De Directe Instructieschaal

Zowel de getrainde observatoren als de mentoren van de stageschool maakten gebruik van de Directe Instructie Schaal (DIS) om het instructiegedrag van de studenten te beoordelen. Deze beoordelingen werden gegeven aan de hand van een vijfpuntsschaal. De DIS vormde een bewerking van de schaal die eerder gebruikt was in het PABO-onderzoek (Veenman e.a., 1993). Voor beide instructiemodellen werd één schaal geconstrueerd vanwege de mogelijkheid van vloeiende overgangen tussen de modellen voor uitvoerend en strategisch handelen. In plaats van een tweedeling kunnen beide modellen gepositioneerd worden als liggend op de uiteinden van één continuüm (cf. Rosenshine & Meister, 1992).

Voorafgaande aan de observaties volgden de drie observatoren een training van circa 40 uur. Hierbij werd gebruik gemaakt van video-opnames en werden op een niet bij het onderzoek betrokken scholengemeenschap 17 proef-observaties verricht. Vervolgens werd aan de hand van 12 lessen (3 lessen Frans, 6 lessen wiskunde en 3 lessen geschiedenis) de tussen-observator-betrouwbaarheid geschat met behulp van variantie-analyse (Winer, 1971). Voor de 26 items van de DIS varieerde deze betrouwbaarheid van 0.60 tot 0.99 (mediaan = 0.82). Twee items met schattingen lager dan 0.60 werden uit de schaal verwijderd.

Op grond van een lesindeling in drie fasen werden de resterende 24 items van de DIS op inhoudelijke gronden verdeeld in drie subschalen: presentatie van de leerstof (waarbij de leraar de leeractiviteiten structureert; inclusief dagelijkse terugblik), begeleidde inoefening (gedeelde structurering van de leeractiviteiten) en verwerking van de leerstof (waarbij de leerling de leeractiviteiten structureert en waarbij de leraar de leerlingen stimuleert tot taakgerichtheid en de structuur van het directe instructiemodel toepast). Maten voor interne consistentie (Cronbachs alpha) werden berekend voor de gehele schaal en voor de drie subschalen. Deze varieerden voor de schaal en de subschalen van de observatoren van 0.63 tot 0.86; voor die van de mentoren van 0.58 tot 0.87.

4.3 Taakgerichtgedrag van de leerlingen

Na 10 minuten een 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 aantal taakgerichte leerlingen werd vervolgens gedeeld door het aantal leerlingen dat in de klas aanwezig was. Deze score werd vervolgens omgezet in een percentage om een indicatie te krijgen van het taakgerichte gedrag van de leerlingen. Deze procedure werd toegepast na minuut 10, 20, 30 en 40. Aldus werden per les vier schattingen verkregen van het taakgerichte gedrag van de leerlingen. Elke les duurde ongeveer 50 minuten. De betrouwbaarheid voor het coderen van het taakgerichte gedrag bedroeg 0.85.

4.4 De lesopdracht

Om de invloed van het programma op het onderwijsgedrag onder enigszins gecontroleerde condities te kunnen vaststellen, werd voor de te geven les op de stageschool een aantal richtlijnen opgesteld. Deze richtlijnen waren zo geformuleerd dat de student kon kiezen voor een van beide instructiemodellen. In de les diende nieuwe leerstof gepresenteerd te worden waarna de leerlingen de kans dienden te krijgen deze zelfstandig te verwerken.

4.5 De uitvoering van het programma

Het programma *Effectieve instructie in het voortgezet onderwijs* (Veenman, Bakermans, Franzen & Van Hoof, 1994) bestaat uit zes hoofdstukken. Hoofdstuk 1 handelt over leren leren. Uitgelegd worden begrippen als metacognitie, zelfregulatie en leerstrategieën. In hoofdstuk 2 wordt aandacht geschonken aan het gebruik van leerstrategieën. Enkele kenmerken van een goede strategiegebruiker worden besproken. Verder worden voor de verschillende vakken enkele uitgewerkte voorbeelden van strategieën gegeven. In hoofdstuk 3 wordt een lesvoorbereidingsformulier gepresenteerd dat geschikt is voor toepassing van het instructiemodel voor uitvoerend en strategisch handelen. In de hoofdstukken 4 en 5 komen de twee instructiemodellen aan bod. Deze hoofdstukken vormen de kern van het programma. De twee instructiemodellen worden uitgelegd en geïllustreerd voor de verschillende vakken in het voortgezet onderwijs. In hoofdstuk 6

worden enkele klassenmanagementvaardigheden behandeld die van belang zijn voor effectieve instructie.

Het trainingsprogramma werd gegeven door drie ervaren docenten onderwijskunde met kennis van het directe instructiemodel. Vóór de start van het programma volgden zij een een-daagse workshop waarin aandacht werd besteed aan de inhoud van het programma en aan mogelijke manieren om het programma uit te voeren en in te passen in het reguliere onderwijs van de opleiding.

Het aantal lessen dat de docenten aan onderdelen van het trainingsprogramma besteedden bedroeg per locatie gemiddeld 8,5 uur (vijf à zes lessen van 100-120 minuten). Over het algemeen was het aantal lessen dat de studenten op de stageschool konden geven beperkt. De mentoren werden door de lerarenopleidingen nauwelijks geïnformeerd over de doelen en de inhoud van het trainingsprogramma. De korte handleiding voor de mentoren werd nauwelijks gebruikt of geraadpleegd. Slechts een klein aantal mentoren gaf de studenten feedback over de uitvoering van het directe instructiemodel.

4.6 Dataverzameling en -verwerking

Voor de aanvang van het programma werd elke student uit de observatiestudie tijdens één les geobserveerd (oktober-december 1994). Na afloop van het programma werd deze student opnieuw tijdens één les geobserveerd (januari-februari 1995). De observatoren hadden geen voorkennis van de groep waartoe de studenten behoorden. Het programma werd uitgevoerd tussen de voor- en nameting (november-december 1994). In dezelfde periode werden de vragenlijsten afgenomen. De interviews met de docenten vonden plaats in maart 1995.

De observatiegegevens van de DIS van de observatoren werden omgezet in somscores en gemiddelden. Omdat het aantal mentoren per meetmoment varieerde en niet alle mentoren zich capabel achten de DIS volledig in te vullen, werd bij de analyse van de gegevens van de mentoren alleen gebruik gemaakt van gemiddelden. De vier scores voor het taakgerichte gedrag van de leerlingen werden per geobserveerde les gemiddeld tot één schatting. Bij de toetsingen werd een significantieniveau van 5% aangehouden. De eenheid van analyse is de

student (of de klas waaraan hij/zij heeft lesgegeven). Voor verdere informatie over de uitvoering van het onderzoek wordt verwezen naar Bakermans, Franzen en Van Hoof (1995).

5 Resultaten

Eerst werd getoetst of er vóór de aanvang van de training significante verschillen bestonden tussen de controle- en de experimentele groep voor wat betreft de DIS-scores. Initiële verschillen werden niet gevonden. In dit opzicht konden de twee groepen als gelijk aan elkaar worden beschouwd. Voor de taakgerichte leertijd werden wel significante initiële verschillen aangetroffen. De leerlingen die onderwezen werden door de studenten uit de experimentele groep waren tijdens de voormeting significant meer taakgericht (84%) dan de leerlingen die onderwezen werden door de studenten uit de controlegroep (77%). Met dit gegeven dient bij de interpretatie van de gegevens rekening te worden gehouden.

Voor een vergelijking van de scores van de voormeting met de nameting werd een *t*-toets uitgevoerd. Op grond van de initiële verschillen betreffende de taakgerichtheid van de leerlingen werd voor de vergelijking van de experimentele groep met de controlegroep een covariantie-analyse (ANCOVA) uitgevoerd met de scores voor taakgerichte leertijd en voor de DIS als covariaten.

Tabel 1 bevat een samenvatting van DIS-scores van de observatoren, alsmede de resultaten van de toetsingen. Hieruit blijkt dat de studenten uit de experimentele groep tijdens de nameting significant hoger scoren op de totaalscore voor de DIS en op de drie subschalen 'presentatie', 'begeleide inoefening' en 'verwerking' dan op de voormeting ($p < 0.01$). Voor de studenten uit de controlegroep zijn geen significante verschillen gevonden tussen de voor- en nameting. Studenten uit de experimentele groep scoren tijdens de nameting op 13 van de 24 items (54%) van de DIS significant hoger dan tijdens de voormeting ($p < 0.05$). Studenten uit de controlegroep scoren op de nameting op slechts 4 van de 24 items significant hoger en op 2 items zelfs significant lager.

Wanneer gekeken wordt naar de gecorri-

Tabel 1

Overzicht van de gemiddelden en standaarddeviaties op de Directe Instructie Schaal (DIS) voor de experimentele groep en de controlegroep volgens het oordeel van de observatoren, alsmede de resultaten van de *t*-toets voor verschillen tussen de voor- en nameting en van de covariantie-analyses

DIS/subschalen/items	Experimentele groep					Controlegroep					F
	Voormeting		Nameting			Voormeting		Nameting			
	M	SD	M	SD	Adj.M	M	SD	M	SD	Adj.M	
DIS totaal (24 items, $\alpha=.86$)	56.8	9.4	70.4**	8.8	69.8	54.8	11.1	58.5	9.2	59.1	17.6**
<i>Subschaal Presentatie</i> (9 items, $\alpha=.75$)	20.5	4.0	26.3**	4.9	26.0	19.2	5.5	21.3	3.9	21.5	13.5**
Dagelijkse terugblik	2.6	1.6	4.6**	0.8	4.7	1.5	1.0	3.3**	1.5	3.3	11.4**
Activeren van de voorkennis	2.3	1.5	2.4	1.4	2.4	2.5	1.6	2.0	1.6	2.0	<1.0
Aangeven van lesdoelen	1.9	0.6	2.0	1.0	2.0	1.7	0.6	1.6	0.7	1.6	1.6
Bieden van structuur (bijv. via een stappenplan)	2.9	0.9	3.6**	0.9	3.6	2.8	1.3	3.1	0.8	3.1	3.8
Modelleren van strategieën	1.1	0.6	1.7*	1.0	1.6	1.2	0.4	1.5	1.0	1.5	<1.0
Belang onderstrepen van strategieën	1.4	0.9	2.2**	1.4	2.2	1.5	0.9	1.5	0.5	1.5	5.7*
Geven van concrete voorbeelden	3.9	0.8	4.0	0.7	4.0	3.5	1.1	3.9*	0.9	4.0	<1.0
Geven van samenvattingen	1.2	0.6	1.9*	1.4	1.9	1.7	1.2	1.2*	0.5	1.2	4.4*
Activeren van denkprocessen	3.3	1.1	4.0*	1.1	3.9	2.9	1.2	3.0	1.2	3.1	5.3*
<i>Subschaal Begeleide Inoefening</i> (5 items, $\alpha=.76$)	15.6	3.8	19.2**	3.4	18.9	15.2	3.7	16.2	4.5	16.5	4.5*
Geven van concrete op- drachten	4.0	1.1	4.8**	0.5	4.8	4.2	0.8	4.3	1.0	4.2	6.5*
Leerlingen verwoorden hun strategiegebruik	2.5	1.4	3.2*	1.6	3.1	2.3	1.4	2.4	1.6	2.5	1.8
Controle op begrijpen	2.3	1.3	2.3	1.0	2.3	1.9	0.9	2.0	1.1	2.1	<1.0
Geven van procesterugkoppeling	3.3	1.0	4.3**	0.7	4.3	3.2	1.0	3.7*	1.2	3.8	3.4
Corrigeren van fouten	3.5	1.0	4.5**	0.6	4.5	3.6	1.1	3.8	1.2	3.8	5.8*
<i>Subschaal Verwerking</i> (10 items, $\alpha=.63$)	20.7	3.2	24.9**	3.9	24.7	20.4	3.6	21.0	3.0	21.1	11.6**
Rolwisselend lesgeven	1.2	0.6	1.3	0.7	1.2	1.3	0.6	1.4	0.7	1.4	<1.0
Stimuleren van de betrokken- heid van de leerlingen	3.7	0.6	4.3**	0.9	4.3	3.3	0.6	3.9**	0.7	3.9	2.0
Controle op uitvoering (leer- lingen beginnen meteen)	3.6	0.9	3.8	1.3	3.8	3.5	0.8	3.8	1.0	3.7	<1.0
Leerlingen helpen elkaar	2.1	0.8	1.9	0.6	1.9	2.1	1.1	1.6*	0.6	1.6	3.3
Verwijzing naar andere toe- passingssituaties	1.4	0.9	1.7	1.3	1.7	1.4	0.8	1.5	1.2	1.5	<1.0
Actief verwerken van de leerstof	3.2	0.7	4.1**	0.9	4.1	3.0	1.0	3.1	0.9	3.1	13.7**
Leerlingen leren hun werk te plannen (zelfregulatie)	1.0	0.2	1.7**	1.2	1.7	1.3	0.7	1.1	0.3	1.1	4.8*
Leerlingen leren hun werk te bewaken (zelfregulatie)	1.0	0.0	1.2	0.6	1.2	1.0	0.2	1.1	0.3	1.1	<1.0
Leerlingen leren hun werk te evalueren (zelfregulatie)	1.0	0.0	1.2	0.6	1.2	1.1	0.3	1.0	0.2	1.1	<1.0
Gebruik van de DI-structuur	2.4	1.1	3.7	0.8	3.6	2.4	1.0	2.5	0.7	2.6	25.4**

Opm.: Experimentele groep $n = 27$; controlegroep $n = 24$. *M* = gemiddelde; *SD* = standaarddeviatie; *Adj. M* = gemiddelde gecorrigeerd voor verschillen op de voormeting. Somscores en gemiddelden zijn gebaseerd op een 5-puntsschaal met als schaaluiters: 1 = geen toepassing van de vaardigheid, 5 = duidelijke toepassing van de vaardigheid. * $p < .05$; ** $p < .01$

geerde scores op de nameting dan blijkt dat de studenten uit de experimentele groep significant hoger scores op de DIS-totaal ($F(1,50) = 17.6, p < .01$) en op de drie subschalen 'presen-

tatie' ($F(1,50) = 13.5, p < .01$), 'begeleide inoefening' ($F(1,50) = 4.5, p < .05$) en 'verwerking' ($F(1,50) = 11.6, p < .01$) dan de studenten uit de controlegroep. Voor de subschaal

Tabel 2

Overzicht van de gemiddelden en standaarddeviaties op de Directe Instructie Schaal (DIS) voor de experimentele groep en de controlegroep volgens het oordeel van de mentoren, alsmede de resultaten van de t-toets voor verschillen tussen de voor- en nameting en van de covariantie-analyses

DIS/subschalen/items	Experimentele groep					Controlegroep					F
	Voormeting		Nameting			Voormeting		Nameting			
	M	SD	M	SD	Adj.M	M	SD	M	SD	Adj.M	
DIS totaal (24 items, $\alpha=.87$)	2.7	0.7	3.1	0.6	3.1	2.8	0.7	3.2	0.5	3.3	1.5
<i>Subschaal Presentatie</i> (9 items, $\alpha=.71$)	2.9	0.8	3.3	0.8	3.3	2.8	0.8	3.3	0.8	3.3	<1.0
Dagelijkse terugblik	2.7	1.5	3.5	1.1	3.5	2.5	1.4	3.7*	1.1	3.6	<1.0
Activeren van de voorkennis	2.7	1.3	3.6*	1.1	3.5	2.9	1.6	3.5	1.1	3.6	<1.0
Aangeven van lesdoelen	2.8	1.2	2.7	1.4	2.6	2.8	1.5	2.7	1.1	2.7	<1.0
Bieden van structuur (bijv. via een stappenplan)	3.1	1.2	3.7	1.1	3.6	3.3	1.1	3.0	1.3	3.0	1.6
Modelleren van strategieën	3.4	1.2	3.5	1.1	3.5	3.1	1.1	3.4	1.2	3.4	<1.0
Belang onderstrepen van strategieën	2.1	1.4	2.8	1.3	2.8	2.4	1.0	2.9	1.2	2.9	<1.0
Geven van concrete voorbeelden	3.5	1.5	3.5	1.5	3.6	3.7	1.4	3.9	1.3	3.9	<1.0
Geven van samenvattingen	2.4	1.2	2.7	1.3	2.6	2.0	1.4	2.7	1.3	2.8	<1.0
Activeren van denkprocessen	3.3	1.3	3.5*	1.1	3.5	2.9	0.9	3.5*	1.2	3.5	<1.0
<i>Subschaal Begeleide Inoefening</i> (5 items, $\alpha=.58$)	2.8	0.8	3.4*	0.7	3.3	3.1	0.6	3.6*	0.5	3.7	4.2*
Geven van concrete opdrachten	3.7	1.3	4.1	0.9	4.1	3.9	0.9	4.3	0.7	4.4	1.2
Leerlingen verwoorden hun strategiegebruik	1.9	1.2	2.9*	1.4	2.8	2.1	1.0	3.1*	1.2	3.2	<1.0
Controle op begrijpen	2.8	1.1	3.1	1.3	3.1	2.7	1.1	3.4*	0.9	3.4	<1.0
Geven van procesterugkoppeling	1.7	1.2	3.1**	1.3	3.1	2.4	1.1	2.8	1.0	2.9	<1.0
Corrigeren van fouten	3.3	0.9	3.9*	1.1	3.8	4.3	0.7	4.3	0.8	4.5	3.1
<i>Subschaal Verwerking</i> (10 items, $\alpha=.85$)	2.4	0.8	2.8	0.7	2.8	2.7	0.8	3.0	0.6	3.0	1.2
Rolwisselend lesgeven	1.6	1.0	1.6	1.1	1.5	1.4	1.1	2.0	1.1	2.1	1.5
Stimuleren van de betrokkenheid van de leerlingen	3.1	1.0	3.7*	0.9	3.6	3.3	1.0	3.8	0.9	3.9	<1.0
Controle op uitvoering (leerlingen beginnen meteen)	3.0	1.2	3.9*	1.2	4.0	4.1	1.0	4.3	0.7	4.1	<1.0
Leerlingen helpen elkaar	2.8	1.5	2.9	1.3	2.9	3.4	1.3	3.3	1.2	3.3	1.0
Verwijzen naar andere toepassingssituaties	2.2	1.3	2.0	1.2	2.0	2.0	1.3	2.8	1.4	2.8	3.0
Actief verwerken van de leerstof	2.0	1.4	2.3	1.2	2.4	2.3	1.3	2.5	1.0	2.5	<1.0
Leerlingen leren hun werk te plannen (zelfregulatie)	2.3	1.3	2.8	1.4	2.7	2.5	1.1	3.2**	1.1	3.2	1.2
Leerlingen leren hun werk te bewaken (zelfregulatie)	1.7	1.1	2.5*	1.2	2.5	2.4	1.1	2.8	1.1	2.7	<1.0
Leerlingen leren hun werk te evalueren (zelfregulatie)	1.8	1.2	2.7*	1.3	2.7	2.1	1.1	3.1**	1.1	3.1	<1.0
Gebruik van de DI-structuur	2.5	0.8	2.7	1.0	2.6	2.6	1.1	2.8	1.1	2.9	<1.0

Op.m.: Voormeting: experimentele groep $n = 18$; controlegroep $n = 15$. Nameting: experimentele groep $n = 19$; controlegroep $n = 20$. M = gemiddelde; SD = standaarddeviatie; $Adj. M$ = gemiddelde gecorrigeerd voor verschillen op de voormeting. De gemiddelde scores zijn gebaseerd op een 5-puntsschaal met als schaal-uiteersten: 1 = geen toepassing van de vaardigheid, 5 = duidelijke toepassing van de vaardigheid. * $p < .05$; ** $p < .01$

'presentatie' zijn bij 4 van de 9 items significante verschillen tussen de experimentele en de controlegroep aangetroffen, voor die van 'begeleide inoefening' bij 2 van de 5 items en voor

de subschaal 'verwerking' bij 3 van de 10 items.

Met behulp van variantie-analyses is de invloed van verschillende variabelen op de

gecorrigeerde scores voor de nameting nagegaan. Het betreft de controlevariabelen 'plaats' (locatie A, B, C), 'geslacht' (man, vrouw), 'vak' (talen: Nederlands, Engels, Duits; exacte vakken: wiskunde, natuurkunde, techniek; overige vakken: aardrijkskunde, geschiedenis, economie), en 'schooltype' (laag: IVBO, VBO, VBO-MAVO; middel: MAVO, MAVO-HAVO; hoog: HAVO, VWO). Systematische verschillen tussen de locaties en het geslacht van de studenten zijn niet gevonden. Een significant interactie-effect is aangetroffen tussen de 'treatment' (de onderzoeksgroepen) en de variabele 'vak' voor de subschaal 'verwerking' ($F(2,50) = 3.5; p < 0.05$) en tussen de 'treatment' en 'schooltype' voor de subschaal 'presentatie' ($F(2,49) = 4.3; p < 0.05$). De scores van de studenten uit het talencluster van de experimentele groep vertoonden op de subschaal verwerking een stijging, die van de studenten van de controlegroep een daling. Verder bleek dat de scores van de studenten uit de experimentele groep die les gaven aan HAVO/VWO- of VWO-klassen op de subschaal 'presentatie' toenamen, terwijl de scores van de studenten uit de controlegroep die aan dezelfde klassen les gaven afnamen.

Tabel 2 geeft een overzicht van de scores van de studenten uit de experimentele en de controlegroep op de DIS volgens het oordeel van de mentoren en van de resultaten van de uitgevoerde toetsen. Uit de resultaten van de *t*-toets blijkt dat de studenten uit de experimentele groep volgens de mentoren alleen op de subschaal 'begeleide inoefening' tijdens de

nameting significant hoger scoren dan tijdens de voormeting ($p < 0.05$). Voor de controlegroep is dit eveneens het geval ($p < 0.05$). Voor de DIS in zijn totaliteit en voor de subschalen 'presentatie' en 'verwerking' worden door de mentoren geen significante verschillen tussen de voor- en nameting voor de experimentele en de controlegroep waargenomen. Studenten uit de experimentele groep scoren volgens de mentoren tijdens de nameting op 9 van de 24 items (38%) van de DIS significant hoger dan tijdens de voormeting; studenten uit de controlegroep op 6 van de 24 items (25%).

Uit de vergelijking van de scores tussen de experimentele en de controlegroep blijkt dat de mentoren alleen voor de subschaal 'begeleide inoefening' een significant verschil opmerken: de studenten uit de controlegroep scoren op deze subschaal significant hoger dan de studenten uit de experimentele groep. Dit is tegengesteld aan de verwachting. Hoewel niet significant, scoorden de studenten uit de controlegroep op deze subschaal van meet af aan hoger dan de studenten uit de experimentele groep. Tussen de boven beschreven controlevariabelen en de onderzoeksgroepen zijn geen significante interactie-effecten aangetroffen.

Uit Tabel 3 blijkt dat het trainingsprogramma geen significant effect heeft op het taakgerichte gedrag van de leerlingen die les hebben gehad van studenten uit de experimentele groep. Tijdens de voormeting bedraagt het percentage leerlingen dat taakgericht bezig is 84% en tijdens de nameting 87%. Slechts op taakmoment 4 is sprake van een significant effect.

Tabel 3

Overzicht van de gemiddelde percentages taakgerichte leerlingen voor de studenten uit de experimentele en de controlegroep volgens het oordeel van de observatoren, alsmede de resultaten van de *t*-toets voor verschillen tussen de voor- en nameting en van de covariantie-analyses

Taakgerichtheid	Experimentele groep					Controlegroep					F
	Voormeting		Nameting			Voormeting		Nameting			
	M	SD	M	SD	Adj.M	M	SD	M	SD	Adj.M	
Gemiddeld percentage taakgerichte leerlingen, (lestotaal)	84.2	10.1	86.7	9.2	86.8	76.9	14.2	84.4*	8.6	84.3	<1.0
Moment 1 (na 10 minuten)	86.6	20.2	87.6	14.6	87.9	81.9	20.1	89.5	9.8	89.2	<1.0
Moment 2 (na 20 minuten)	86.0	11.5	88.2	12.0	88.1	77.7	22.4	86.5*	10.2	86.6	<1.0
Moment 3 (na 30 minuten)	86.8	9.1	86.9	11.6	85.8	75.5	13.8	83.9*	11.4	85.0	<1.0
Moment 4 (na 40 minuten)	77.5	15.5	84.2*	11.8	83.8	72.1	17.2	77.7	22.6	78.1	1.3

Opm.: M = gemiddelde; SD = standaarddeviatie; Adj. M = gemiddelde gecorrigeerd voor verschillen op de voormeting. * $p < .05$

Bij leerlingen die les kregen van studenten uit de controlegroep is er een significant verschil tussen het gemiddeld percentage taakgerichte leerlingen: tijdens de voormeting 77% en tijdens de nameting 84% ($p < 0.05$). Deze toename wordt vooral gerealiseerd op de observatiemomenten 2 en 3 (na 20 en 30 minuten in de les). Tussen de nameting-scores van de experimentele en de controlegroep bestaan geen significante verschillen (de taakgerichte score van de voormeting is hierbij gebruikt als covariaat). Ook hier zijn geen significante interactie-effecten tussen de controle-variabelen en de onderzoeksgroepen gevonden.

Uit de resultaten van de vragenlijsten en de interviews kan geconcludeerd worden dat zowel de studenten als de docenten tevreden zijn over het trainingsprogramma. Studenten vinden dat het trainingsprogramma goed aansluit bij hun voorkennis en bij hun praktijk-situatie. Het model voor strategisch handelen wordt door de studenten als bruikbaar ervaren, maar zij geven de voorkeur aan het model voor uitvoerend handelen. Op alle locaties spreken de studenten de verwachting uit dat ze (een aantal van) de geleerde vaardigheden willen blijven toepassen in de klas. De docenten geven aan dat zij het trainingsprogramma als waardevol ervaren en willen derhalve (delen van) het programma in het onderwijsplan voor het nieuwe schooljaar gebruiken.

6 Discussie

Uit de gegevens van de observatoren blijkt dat de studenten die het trainingsprogramma hebben gevolgd de gewenste instructievaardigheden over het algemeen beter toepasten dan de studenten die het trainingsprogramma niet hebben gevolgd. Deze resultaten komen grotendeels overeen met de PABO-studie (Veenman e.a., 1993). Ook PABO-studenten bleken de vaardigheden uit het directe instructiemodel op de stageschool toe te passen.

In tegenstelling tot de gegevens van de observatoren vinden de mentoren dat de getrainde studenten de gewenste instructievaardigheden niet beter toepassen dan de niet-getrainde studenten. Het verschil in uitkomsten tussen de observatoren en de mentoren laat zich mogelijk als volgt verklaren. Hoewel de

mentoren een korte samenvatting hebben gekregen van de inhoud van de instructievaardigheden die in het trainingsprogramma werden benadrukt, hebben slechts weinig mentoren kans gezien deze samenvatting te bestuderen en deze te gebruiken als leidraad bij de nabespreking van de lessen met de studenten. Bij de invulling van de observatielijst zijn de mentoren afgegaan op de eenregelige omschrijvingen van de instructievaardigheden in de DIS en hebben deze naar eigen inzichten geïnterpreteerd. Verder is gebleken dat tijdens de observaties van de lessen door de observatoren niet alle mentoren tijdens de door de student gegeven les in de klas aanwezig zijn geweest. Zij vulden achteraf de observatielijst in op grond van globale impressies. Om deze redenen wordt in deze studie meer waarde gehecht aan de scores van de observatoren dan aan de scores van de mentoren. Ook bleek dat tijdens de stageperiode de mentoren nauwelijks inhoudelijke feedback hebben gegeven aan de studenten over het gebruik van de instructievaardigheden voor uitvoerend of strategisch handelen. Mogelijk zouden door een meer inhoudelijke begeleiding de mentoren wel verschillen hebben kunnen constateren tussen getrainde en niet-getrainde studenten. De PABO-studie kende voor de mentoren geen controlegroep, zodat een rechtstreekse vergelijking met die studie niet mogelijk is.

Met betrekking tot de taakgerichte leertijd van de leerlingen zijn geen significante verschillen tussen de leerlingen van getrainde en niet-getrainde studenten gevonden. De leerlingen van de studenten uit de experimentele groep scoorden op de voormeting significant hoger dan de leerlingen van de studenten uit de controlegroep (84% om 77%). Het percentage leerlingen dat taakgericht bezig was nam voor de experimentele groep toe met 2,5%, voor de controlegroep met 7,5%. Tussen de locaties in de experimentele groep bleken bij de voormeting belangrijke verschillen te bestaan. Op locatie A bedroeg het percentage taakgerichte leerlingen 80%, op locatie B 91% en op locatie C 88%. Alleen op locatie A was een belangrijke vooruitgang te constateren, namelijk van 80% naar 87%. Op locatie B liep dit percentage terug van 91% naar 83%, op locatie C bleef dit nagenoeg gelijk. Door de hoge beginwaarden op de locaties B en C is sprake van een

plafond-effect, waardoor op de nameting een regressie optreedt naar het gemiddelde. Verder is op deze locaties sprake van een selectie-effect. Op het moment van de voormeting waren de studenten op deze twee locaties nog niet met hun stages begonnen. De docenten kozen voor deze studenten de stagescholen, waardoor de mogelijkheid niet uitgesloten kan worden dat de studenten uit de experimentele groep in de klassen kwamen met de meer taakgerichte leerlingen. De toename in taakgerichte leertijd in de controlegroep kan verklaard worden door een grotere vertrouwdheid van de studenten met de leerlingen tijdens de stageperiode. In de voormeting zagen de studenten de klas voor het eerst. Vanwege de initiële verschillen tussen de experimentele en de controlegroep bleek de invloed van de training op het taakgerichte gedrag van de leerlingen in dit onderzoek moeilijk vast te stellen omdat statistische correctie voor initiële verschillen de invloed van deze verschillen op de nameting niet geheel corrigeert (Reichardt, 1979). In de PABO-studie werden bij de aanvang van de studie geen significante verschillen tussen de controle- en de experimentele groep gevonden. In deze studie bleek het volgen van het trainingsprogramma wel een significant positief effect te hebben op het taakgericht gedrag van de leerlingen. Belangrijke verschillen met de PABO-studie zijn dat in deze studie de studenten les gaven aan de hand van vooraf gestandaardiseerde reken- en taallessen, dat de vooren nameting in de reguliere stageperioden vielen en dat de klassen waarin zij les gaven onderling meer vergelijkbaar waren. In toekomstig onderzoek dient nadere aandacht gegeven te worden aan de relatie tussen het trainingsprogramma en de taakgerichte leertijd van de leerlingen.

Op grond van de initiële verschillen in taakgerichte leertijd tussen de experimentele en de controlegroep in de NLO-studie kan niet uitgesloten worden dat de leerlingen in de klassen van de getrainde studenten leergieriger waren dan de leerlingen in de klassen van de niet-getrainde studenten. Het zou kunnen zijn dat de getrainde studenten de gewenste vaardigheden gemakkelijker konden uitvoeren dan de niet-getrainde studenten.

Ondanks de ervaring die de docenten reeds hadden met het directe instructiemodel is de

voorschooling van de docenten voor de uitvoering van het programma te kort geweest. Uit de gesprekken met de studenten is gebleken dat de docenten het voordoen of modelleren van de instructievaardigheden voor hun studenten weinig toepasten. De voorschooling heeft de betrokken docenten onvoldoende uitgerust om de gewenste instructievaardigheden modelmatig voor hun studenten te demonstreren. Ook bleek uit de gesprekken met de docenten en de studenten dat een goede uitvoering van het programma met bijbehorende opdrachten voor oefening op de opleiding of de stageschool meer tijd behoeft. De betrokken docenten zagen geen kans hun studenten te begeleiden op de stageschool. Indien in de toekomst meer tijd wordt uitgetrokken voor het programma ontstaat meer ruimte om aandacht te geven aan de bruikbaarheid van het instructiemodel voor strategisch handelen. Nu hadden de studenten een voorkeur voor het instructiemodel voor uitvoerend handelen. Dit komt omdat in het huidige onderwijs de rol van de leraar als verstreker van informatie meer dominant is dan de rol van de leraar als begeleider van leerprocessen (cf. Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs, 1994).

Uit dit onderzoek blijkt dat de getrainde studenten in HAVO- en VWO-klassen op de subschaal 'presentatie' vooruit gingen en de ongetrainde studenten achteruit. In de andere klassen (IVBO, VBO, MAVO/HAVO) boekten zowel getrainde als niet-getrainde studenten vooruitgang, waarbij de getrainde studenten het meest vooruit gingen. Een duidelijke verklaring voor dit interactie-effect is niet te geven. Het zou kunnen zijn dat HAVO/VWO-leerlingen meer gericht zijn op het opnemen en verwerven van informatie en dat de getrainde studenten van deze interesse beter gebruik hebben gemaakt dan de niet-getrainde studenten. Verder blijkt dat de getrainde studenten uit het talencluster op de subschaal 'verwerking' een vooruitgang boekten en de ongetrainde studenten uit het talencluster een achteruitgang. De getrainde studenten uit dit cluster slaagden er beter in dan de ongetrainde studenten om de gewenste instructievaardigheden te realiseren doordat zij de leerlingen actiever bij de leertaken betrokken.

Concluderend kan worden gesteld dat ondanks tekortkomingen in de uitvoering van

de training, het trainingsprogramma voor de betrokken studenten en docenten van de tweedegraads lerarenopleiding een succesvolle en waardevolle onderneming is geweest.

Literatuur

- Bakermans, J., Franzen, Y., & Hoof, M. van (1995). *Effectieve instructie in het voortgezet onderwijs: Een evaluatie-onderzoek naar het trainingsprogramma*. (Doctoraalscriptie) Nijmegen: Vakgroep Onderwijskunde, Katholieke Universiteit Nijmegen.
- Reichardt, C.S. (1979). The statistical analysis of data from nonequivalent group designs. In T.C. Cook & D.T. Campbell (Eds.), *Quasi-experimentation: Design and analysis issues for field settings* (pp. 147-205). Chicago: Rand McNally.
- Rosenshine, B., & Meister, C. (1992). The use of scaffolds for teaching higher-level cognitive strategies. *Educational Leadership*, 49(7), 26-33.
- Shuell, T.J. (1988). The role of the student in learning from instruction. *Contemporary Educational Psychology*, 13, 276-295.
- Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs. (1994). *De tweede fase vernieuwt: Scharnier tussen basisvorming en hoger onderwijs, deel 2*. Den Haag: Porsius.
- Veenman, S. (1992). Effectieve instructie volgens het directe instructiemodel. *Pedagogische Studiën*, 69(4), 242-269.
- Veenman, S., Leenders, Y., Meyer, P., & Sanders, M. (1993). Leren lesgeven met het directe instructiemodel. *Pedagogische Studiën*, 70(1), 2-16.
- Veenman, S. (Red.), Bakermans, J., Franzen, Y., & van Hoof, M. van (1994). *Effectieve instructie in het voortgezet onderwijs: Leren onderwijzen met behulp van het directe instructiemodel*. Hoevelaken: Christelijk Pedagogisch Studiecentrum
- Winer, B.J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.

Manuscript aanvaard 29-1-1997

Auteurs

S. Veenman is als universitair hoofddocent werkzaam bij de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen.

J. Bakermans is als ontwikkelingsconsulente (CD-i en Internet) werkzaam bij de Vakopleiding Procesindustrie (VaPro) te Leidschendam.

Y. Franzen is als docente pedagogiek/onderwijskunde werkzaam bij de Fontys Hogescholen (Pabo) te Sittard en Roermond.

M. van Hoof is als cursusleidster werkzaam bij het Instituut voor Toegepaste Voorlichtingskunde (ITV) te Wageningen.

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

Abstract

Learning to teach according to the direct instruction model: Effects of a pre-service secondary-education training program

S. Veenman, J. Bakermans, Y. Franzen & M. van Hoof. *Pedagogische Studiën*, 1997, 74, 328-337.

A quasi-experimental, treatment-control group investigation was designed to test the effects of a pre-service training course for secondary-education teachers. Pre- and post-training comparison of classroom observations by trained observers revealed significantly more effective instruction by the student teachers after training. No treatment effect was found for pupil engagement rates. The ratings from the supervising teachers did not show significantly better use of the recommended instructional skills by the trained student teachers than by the untrained student teachers. The course was rated by the student teachers and teacher educators as valuable.

Wat doen kleuters en leerkrachten in groep 2 van het basisonderwijs?

D. Manders en C. Aarnoutse

Samenvatting

In dit onderzoek wordt nagegaan welke activiteiten en talige gedragingen leerlingen en leerkrachten van groep 2 van het basisonderwijs uitvoeren en hoeveel tijd deze activiteiten en gedragingen in beslag nemen. Op acht 'doorsnee' scholen werd gedurende twee ochtenden en twee middagen een kind in groep 2 geobserveerd. Daarbij werd gebruik gemaakt van een observatie-instrument dat bestond uit vier variabelen: tijd, setting, activiteit en gedrag van het kind. Uit het onderzoek blijkt dat er in groep 2 gedurende een dag veel psychomotorische activiteiten plaatsvinden en relatief weinig cognitieve. De belangrijkste conclusie is dat er in de praktijk van de kleutergroep weinig sprake is van interactie tussen de leerkracht en de leerlingen en tussen de leerlingen onderling. Uit de gegevens kan worden afgeleid dat het taal-aanbod in groep 2 uitgebreid en verbeterd kan worden.

1 Inleiding

In dit onderzoek gaan we na hoe in groep 2 van een aantal basisscholen onderwijs, in het bijzonder taalonderwijs, wordt gegeven. Onderzocht wordt welke activiteiten leerlingen en leerkrachten uitvoeren en hoeveel tijd deze activiteiten in beslag nemen. Na een korte schets van een aantal publicaties over de vernieuwing van het onderwijs in groep 1 en 2 bespreken we de opzet en uitkomsten van enkele observatiestudies die eerder in groep 1 en 2 zijn uitgevoerd.

De positie van het onderwijs in groep 1 en 2 staat de laatste jaren volop in de belangstelling, getuige de indrukwekkende reeks van publicaties, commissies, rapporten en beleidsmaatregelen. Wat betreft publicaties in de vorm van artikelen en boeken verwijzen we korthedshalve naar Bus en Verhoeven (1992), Cohen de

Lara (1996), Janssen-Vos (1992, 1996), Janssen-Vos en Laevers (1996), Janssen-Vos, Pompert en Vink (1995), Jongerius en Beermink (1993), Van Kuyk (1991), Van Oers en Janssen-Vos (1992), Verhoeven (1994), en Van Zeeland (1992). In verschillende rapporten (ARBO, 1990; Inspectie van het onderwijs, 1992 en 1994; Commissie Evaluatie Basisonderwijs, 1994) wordt geconstateerd dat het voormalige kleuter- en lager onderwijs weliswaar in 1985 zijn samengevoegd, maar dat doelgerichtheid in het onderwijs aan jonge kinderen nog te weinig aanwezig is en dat het ononderbroken ontwikkelingsproces van leerlingen zeker in de eerste groepen van de basisschool nog niet gerealiseerd is. In haar rapport "Probleemidentificatie en aanzet voor een actieplan taal" constateert een commissie onder leiding van Aarnoutse (1995) dat in de groepen 1 en 2 systematisch aandacht moet worden besteed aan de taalontwikkeling van de kinderen, met name van die kinderen die afkomstig zijn uit sociaal zwakke milieus en uit tweetalige groepen. In de nota "Een impuls voor het basisonderwijs" beklemtoont de staatssecretaris Netelenbos (1995) van OC&W het belang van goed (taal)onderwijs in de onderbouw (groep 1 tot en met 4). Zij geeft het Expertisecentrum Nederlands de opdracht om het onderwijs in de Nederlandse taal in die groepen te verbeteren.

De eerste stap op weg naar verbetering van het (taal)onderwijs in groep 1 en 2 is nauwkeurig onderzoek naar de wijze waarop dat onderwijs wordt gegeven. Zonder een duidelijk beeld van de werkelijkheid of van het dagelijks reilen en zeilen in die groepen is het gevaar groot dat een mogelijke vernieuwing niet of onvoldoende aansluit bij of rekening houdt met de beginsituatie van de leerkrachten en leerlingen. Deze situatie bepaalt immers het aantal en de inhoud van de stappen of fasen die nodig zijn om een vernieuwing met succes te realiseren.

Volgens de eerder genoemde "Probleem-identificatie en aanzet voor een actieplan taal" is er nog niet veel observatie-onderzoek uitgevoerd naar het (taal)onderwijs in de betreffende groepen. In het onderhavige onderzoek zullen we proberen een beeld te geven van het onderwijs, in het bijzonder van het taalonderwijs, in groep 2. Hierbij maken we uiteraard zoveel mogelijk gebruik van de kennis die in eerdere observatie-onderzoeken is verzameld. In het volgende beschrijven we twee vrij recente observatiestudies die voor ons onderzoek van belang zijn, namelijk de studie van Van Gennip en Ruijs (1992) en van Reezigt, Guldemond en Ros (1995). Enkele andere onderzoeken naar de dagelijkse tijdsbesteding in de groepen 1 en 2 zullen we eveneens vermelden. Voor een overzicht van de resultaten uit 13 onderzoeken die sinds 1985, het jaar dat de Wet op het Basisonderwijs in werking trad, zijn uitgevoerd verwijzen we naar Reezigt (1996).

Van Gennip en Ruijs (1992) onderzochten de pedagogisch-didactische kwaliteit van het onderwijsaanbod in de groepen 1- 2 en 3 van 20 basisscholen. Na interviews met de directies en leerkrachten observeerden ze gedurende één ochtend het onderwijs in jaargroep 1-2 en 3. In totaal werd in 20 kleutergroepen geobserveerd. Het observatie-instrument was opgebouwd volgens de settings die in die groepen meestal voorkomen. Onder settings verstaan zij "een geheel van met elkaar samenhangende activiteiten van leerkrachten en kinderen" (p. 25). Voorbeelden van settings zijn volgens hen de opening van de dag, de spelwerkijd, buiten spelen en een kringactiviteit.

Op basis van de observaties en interviews beschrijven Van Gennip en Ruijs hoe het ochtendprogramma in groep 1 en 2 er in hoofdlijnen uit ziet. Volgens hen duurt een ochtend in een kleutergroep gewoonlijk drie uur en vormen de opening en afsluiting een vast onderdeel van het ochtendprogramma. Verder is er altijd een spelwerkijd die ongeveer een uur duurt. In dat uur werken en/of spelen de kinderen, individueel en/of in kleine groepen met ontwikkelingsmaterialen, met spelactiviteiten en met thematische of projectactiviteiten. Gemiddeld zitten de kleuters ongeveer anderhalf uur per ochtend in de kringopstelling. In de kring vinden veel en verschillende activiteiten plaats zoals het kringgesprek, vertellen, voor-

lezen, boekpromotie, auditieve training, poppenkast, muziek maken, dansen en voorbereidend rekenen. Als laatste vast onderdeel noemen Van Gennip en Ruijs verder het samen eten en drinken. Naast deze vaste onderdelen is er ruimte voor activiteiten die per dag wisselen zoals taalontwikkeling, voorlezen, vertellen en muzikale vorming.

De gemiddelde tijdsduur van de settings die dagelijks voorkomen bedraagt volgens de observaties van Van Gennip en Ruijs voor de dagopening, eten en drinken, buiten spelen en de afsluiting respectievelijk 17, 17, 49 en 7 minuten. De gemiddelde tijdsduur van de gezamenlijke activiteiten met een vormende waarde bedraagt voor voorlezen, het kringgesprek, expressie en bewegingsonderwijs respectievelijk 20, 21, 19 en 49 minuten. De gemiddelde tijdsduur van settings die 's ochtends gericht zijn op spelen, leren en werken bedraagt voor de spelwerkijd, lezen en taal, rekenen en schrijven respectievelijk 54, 18, 19 en 0 minuten. Uit deze gegevens blijkt dat 's ochtends relatief veel tijd wordt uitgetrokken voor buiten en/of binnen spelen. Verder valt op dat 's ochtends vrij veel tijd wordt besteed aan taalactiviteiten zoals voorlezen, kringgesprek, expressie en taal en lezen. Tot slot valt de duur van de spelwerkijd op. Volgens Van Gennip en Ruijs maken de kleuters dan in vrijwel alle geobserveerde groepen gebruik van speelmateriaal, van constructie- en compositiematerialen of van vormgevende materialen. Opgemerkt moet worden dat de gemiddelde tijdsduur zoals boven vermeld *niet* betrekking heeft op de activiteiten van *alle* leerkrachten en leerlingen, maar op de activiteiten van die leerkrachten en leerlingen bij wie die activiteiten werden waargenomen.

Het is jammer dat in het onderzoek van Van Gennip en Ruijs alleen 's ochtends in groep 1-2 is geobserveerd. Een ander zwak punt is de geringe uitwerking of specificatie van de settings in activiteiten. Een en ander heeft tot gevolg dat hun onderzoeksgegevens met de nodige voorzichtigheid moeten worden gehanteerd en dat het moeilijk is om deze gegevens met die van ander observatie-onderzoek te vergelijken.

Reezigt, Guldemond en Ros (1995) observeerden vijf leerkrachten van groep 1 en 2 ieder gedurende zes ochtenden; de leerkrachten

Tabel 1

Gemiddelde tijd aan activiteiten in groep 1-2 besteed en standaarddeviatie (in procenten), 30 observaties, vijf leerkrachten (volgens Reezigt, Guldemond & Ros, 1995)

Activiteit	gemiddelde tijd	SD
1. buiten spelen	21.0	12.9
2. individueel werken/leerkracht betrokken	14.8	8.6
3. overgangen tussen lessen	10.0	5.4
4. inlooptijd	9.3	6.5
5. eten en drinken	7.2	2.4
6. groepsgesprek/onderwijsleergesprek	6.5	7.2
7. vertellen in de kring	5.1	4.5
8. opdrachten geven	4.3	3.2
9. niet-inhoudelijke activiteiten	4.2	4.6
10. individueel werken/leerkracht niet betrokken	3.8	4.9
11. gymles/geleid door de leerkracht	3.5	7.3
12. gymles/niet geleid door de leerkracht	2.6	6.9
13. administratieve routines	2.2	2.2
14. wachttijd	1.9	2.9
15. leerkracht uit de klas	1.6	1.9
16. voorlezen	1.0	2.5
17. instructie aan de klas	1.0	2.7

werkten op onderwijsvoorrangsscholen. Reezigt et al. maakten voor hun observaties gebruik van het Special Strategies Observation System (SSOS) van Schaffer en Nesselrodt (1992). De resultaten van hun onderzoek berusten op de observaties van in totaal 30 ochtenden. Eén ochtend duurde meestal 210 minuten. Tabel 1 geeft een overzicht van de tijd die gemiddeld besteed wordt aan de verschillende activiteiten. De tijd is hierbij omgerekend in procenten.

Tabel 1 laat zien dat aan buiten spelen gemiddeld de meeste tijd wordt besteed, gevolgd door individueel werken. Overgangen tussen de lessen en de inloop 's morgens vroeg nemen ook veel tijd in beslag. Eten en drinken, het groepsgesprek en onderwijsleergesprek en vertellen in de kring volgen daarna. Alle overige activiteiten nemen minder dan 5% van de tijd per ochtend in beslag. De hoge standaarddeviaties wijzen op grote verschillen tussen de leerkrachten. Reezigt et al. vinden het vreemd dat voorlezen en vertellen in de kring niet vaker voorkomen, en voor zover ze voorkomen maar kort duren. Vooral omdat hun onderzoek betrekking had op onderwijsvoorrangsscholen die vaak gericht zijn op specifieke taalontwikkelingsactiviteiten.

Reezigt et al. hebben hun observatiegegevens vergeleken met de uitkomsten van andere onderzoeken naar de dagelijkse tijdsbestedingen in de groepen 1 en 2. Het betreft een vergelijking met de observatie-onderzoeken van

Van der Meer, Appelhof, Ten Doesschate en Kramer (1986), van Harskamp, Pijl en Snippe (1992) en van Van Gennip en Ruijs (1992). Uit het onderzoek van Van der Meer et al. blijkt dat gemiddeld 60% van de tijd aan onderwijs wordt besteed en dat organisatie, ontspanning, verzorging en reactiviteiten respectievelijk 13%, 11%, 15% en 1% in beslag nemen. Het onderzoek van Harskamp, Pijl en Snippe (1992) laat zien dat in groep 2 29% van de schooltijd besteed wordt aan buiten spelen en aan spelen in de gymnastiekzaal, terwijl in 10% van de tijd de kinderen in de klas spelen (poppenhoek en dergelijke). Omdat in deze onderzoeken vaak observatiecategorieën worden gebruikt die moeilijk met elkaar te vergelijken zijn en soms slechts globaal of niet omschreven zijn, laten we deze vergelijking hier achterwege (vgl. Commissie Evaluatie Basisonderwijs, 1994).

Het sterke punt van het onderzoek van Reezigt et al. in vergelijking met dat van Van Gennip en Ruijs is dat de observatiecategorieën betrekking hebben op activiteiten en niet op vrij globale settings. Bovendien zijn de uitkomsten gebaseerd op meerdere observaties (ochtenden) per leerkracht, wat de stabiliteit van de resultaten ten goede komt.

Uit het bovenstaande blijkt dat precieze gegevens over de activiteiten die leerkrachten en leerlingen dagelijks in groep 1 en 2 uitvoeren en over de tijd die deze activiteiten kosten voor een belangrijk deel ontbreken. De studie

van Reezigt, Guldemond en Ros (1995) biedt nog het meeste houvast, al heeft die alleen betrekking op de ochtenden. Wat betreft het onderwijs in de Nederlandse taal ontbreken nog veel gegevens. In dit onderzoek staan twee vragen centraal namelijk:

- a. Welke activiteiten voeren leerlingen en leerkrachten in groep 2 uit en hoeveel tijd nemen deze activiteiten in beslag?
- b. Welke talige gedragingen voeren leerlingen en leerkrachten uit in groep 2 en hoeveel tijd vragen deze gedragingen?

2 Methode van onderzoek

2.1 Opzet van het onderzoek

De bedoeling was om op tien 'gewone' basisscholen een onderzoek in groep 2 uit te voeren, waarbij telkens één leerling zou worden geobserveerd. Elk kind zou gedurende vier dagdelen worden geobserveerd aan de hand van een vooraf ontwikkeld observatie-instrument. Gedurende de geobserveerde dagdelen zouden ook opnamen worden gemaakt van alles wat het betreffende kind zou zeggen. Daartoe zou elk kind een microfoon tijdens de vier dagdelen dragen. Naast de observaties en opnamen zou ook een interview met de leerkrachten van groep 2 worden gehouden.

2.2 Selectie van de scholen en leerlingen

Om praktische redenen is gekozen voor scholen uit Helmond en omgeving. In totaal werden at random 20 'gewone' basisscholen per brief verzocht om aan het onderzoek mee te doen. Onder 'gewone' basisscholen worden alle basisscholen verstaan die niet tot de traditionele vernieuwingsscholen (Jenaplan-, Montessori-, Freinetscholen e.d.) behoren en niet onder het onderwijsvoorrangsbeleid van de overheid vallen. Na een telefonische ronde waren acht scholen bereid om aan het onderzoek mee te werken. De twaalf scholen die niet aan het onderzoek konden meedoen, gaven daarvoor diverse redenen op die meestal neerkwamen op de belasting die deelname met zich mee zou brengen.

Wat de kenmerken van de scholen en de kinderen betreft kan het volgende worden opgemerkt: vier van de acht scholen bevonden zich in een plattelandsgemeente en vier scholen

in een verstedelijkte gemeente of stad. Het aantal leerlingen en leerkrachten varieerde van 159 tot meer dan 350 leerlingen en van 9 tot 25 leerkrachten. De scholen waren qua leerlinggewicht in redelijke mate met elkaar vergelijkbaar. De gemiddelde gewichten varieerden van 1.09 tot 1.20. Wat betreft de denominatie: zeven scholen waren rooms-katholiek en één school was openbaar.

Met betrekking tot de selectie van de kinderen werd aan de betreffende leerkracht van groep 2 gevraagd om een 'doorsnee' leerling te kiezen. Het te observeren kind mocht geen probleemkind zijn en niet meertalig. Alle acht geselecteerde kinderen hadden een leerlinggewicht van 1.00 en waren volgens de beroepclassificatie van het ITS (Van Westerlaak, Kropman & Collaris, 1975) afkomstig uit een midden- of hoger milieu. Alle leerlingen waren 5 of 6 jaar; er waren zes meisjes en twee jongens. Op één leerling na hadden alle kinderen op een peuterspeelzaal gezeten. Drie kinderen waren de oudste in het gezin, twee kinderen het middelste kind en drie kinderen waren het jongste kind. Op alle acht scholen zaten de kinderen in heterogene kleutergroepen, wat betekent dat de leerlingen van groep 1 en 2 door elkaar zaten. De geselecteerde kinderen zaten in groepen die varieerden van 20 tot 32 leerlingen. De verhouding tussen het aantal kinderen van groep 1 en 2 was in de meeste klassen ongeveer gelijk.

2.3 Meetinstrumenten

Zoals eerder vermeld zijn er twee soorten van meetinstrumenten ontwikkeld namelijk een observatie-instrument en een interview.

2.3.1 Observatie-instrument

Om zicht te krijgen op wat er in een kleutergroep dagelijks gebeurt, is een week lang op een school in Nijmegen geobserveerd. Hierbij was sprake van oriënterende, niet-gestructureerde observaties. Tijdens die observaties werd alles genoteerd wat zich in de klas afspeelde met als doel te achterhalen welke settings en activiteiten in groep 2 voorkomen. Met behulp van deze observaties en de eerder beschreven observatie-onderzoeken werd het observatie-instrument ontwikkeld (zie Tabel 2).

Het observatie-instrument bestaat uit vier variabelen: tijd, setting, activiteit en gedrag van

het kind. De variabele *tijd* houdt in dat het geselecteerde kind elke dertig seconden wordt geobserveerd en dat de codes voor die observatie in die tijd worden ingevuld. *Setting* wordt in navolging van Van Gennip en Ruijs (1992) omschreven als een afgerond geheel van met elkaar samenhangende activiteiten van leerkrachten en van kinderen. We onderscheiden de volgende settings: de kring, de speelwerk-tijd, bewegen, organisatie, dagelijkse dingen en video. Binnen een setting kunnen verschillende activiteiten voorkomen. Zo kan er 's morgens in de kring verteld worden door de kinderen, voorgelezen door de leerkracht of een versje worden opgezegd. Binnen die setting is dan sprake van drie activiteiten. Binnen een activiteit is een kind ergens mee bezig, het vertoont een bepaald *gedrag*. Meestal is dat gedrag gericht op een bepaald object. Dat object kan iemand of iets zijn. Zo kan bij de activiteit 'voorlezen' een kind bijvoorbeeld het volgende gedrag vertonen: het luistert naar de leerkracht die voorleest, het fluistert met zijn buurman, het geeft antwoord op een vraag van de leerkracht of het stelt een vraag aan de leerkracht. Het gedrag van kinderen is onder te verdelen in talig gedrag en niet talig gedrag. Een kind vertoont talig gedrag als het spreekt, luistert, leest of schrijft. Voorbeelden van niet talig gedrag zijn: met blokken spelen, knutselen en rennen over de speelplaats.

Zoals uit het bovenstaande en uit Tabel 2 blijkt, zijn de settings, activiteiten en gedragingen in verschillende soorten of categorieën verdeeld. In het volgende geven we een beschrijving van de verschillende soorten settings, activiteiten en gedragingen. Als eerste beschrijven we de settings die we observeren:

Met '*kring*' bedoelen we dat gedeelte van de dag dat de kinderen in kringopstelling zitten. Soms vinden er in de kring activiteiten plaats die onder een andere setting vallen zoals het omkleden voor de gymnastiekles (hoort bij organisatie), de dagopening en fruit eten (horen bij dagelijkse dingen). Als een activiteit onder een andere setting geplaatst kan worden, wordt niet de setting 'kring' gecodeerd. Wel wordt die setting gecodeerd als er wordt voorgelezen of een prentenboekpresentatie plaatsvindt en de kinderen vóór de leerkracht op de grond zitten of in een halve kringopstelling.

'*Speelwerk-tijd*' wordt eveneens in navolging van Van Gennip en Ruijs gedefinieerd als een "periode waarin kinderen werken en/of spelen, individueel en/of in kleine groepjes, met ontwikkelingsmaterialen, met spelactiviteiten en met thematische of projectactiviteiten" (p. 57). Dat werken of spelen kan dan plaatsvinden aan tafel, op de grond, in de hoeken of in de gang of hal. In de speelwerk-tijd kunnen kinderen een opdracht krijgen (open of gesloten) of een activiteit zelf kiezen.

Onder '*bewegen*' vallen het buiten spelen, binnen spelen en de gymnastiekles. 'Buiten spelen' spreekt voor zichzelf. 'Binnen spelen' wil zeggen dat de kinderen vrij spelen in de speelzaal, al dan niet met materialen en/of toestellen. Zodra de leerkracht met de leerlingen geleide activiteiten doet in de speelzaal is sprake van een gymnastiekles.

Onder '*organisatie*' vallen de activiteiten die noodzakelijk zijn om een dag goed te laten verlopen. De leerkracht is bijvoorbeeld bezig met het organiseren van de speelwerk-tijd, het buiten spelen, de gymnastiekles et cetera.

De setting '*dagelijkse dingen*' heeft betrekking op die activiteiten die iedere dag terugkeren. 'Buiten spelen' wordt niet onder 'dagelijkse dingen' geplaatst, maar hoort bij 'bewegen'.

Wat de setting '*video*' inhoudt, spreekt voor zichzelf. De leerkracht en de leerlingen bekijken, meestal buiten de klas, een video-opname.

In Tabel 2 worden alle activiteiten die we observeren vermeld. In deze tabel zijn ze geplaatst onder de setting waarbinnen ze voorkomen. De activiteiten waarvan de inhoud niet geheel duidelijk is, lichten we hier toe.

Eerst geven we een omschrijving van de activiteiten in de *kring*. Van 'voorlezen' is sprake als een leerkracht een boek gebruikt waaruit hij het verhaal van een schrijver in diens woorden vertelt. Als er een prentenboek wordt voorgelezen, wordt er 'prentenboekpresentatie' gecodeerd. Bij 'vertellen van een verhaal door de leerkracht' wordt een verhaal voorgedragen zonder dat daarbij een boek wordt gebruikt (Van Zeeland, 1992).

Met 'verhaal navertellen' wordt bedoeld dat de kinderen een verhaal dat is voorgelezen in eigen woorden moeten navertellen. Vaak gebeurt dit door middel van een vraag- en antwoordspel waarbij de leerkracht de vragen

Tabel 2
Settings met bijbehorende activiteiten

A. kring	C. bewegen
1. voorlezen door de leerkracht	36. buiten spelen zonder materiaal
2. vertellen van een verhaal door de leerkracht	37. buiten spelen met materiaal
3. verjaardag vieren	38. speelkwartier
4. vertellen door de kinderen	39. binnen spelen zonder materiaal/toestellen
5. verhaal navertellen	40. binnen spelen met materiaal/toestellen
6. vertellen naar aanleiding van een verhaal	41. geleide gymnastiekles
7. prentenboek presentatie	42. kringspel
8. 'voorlezen' door een kind	
9. liedjes zingen	D. organisatie
10. schrijven in de kring	43. uitleg werkles
11. brief voorlezen	44. werkjes uitdelen
12. aanbieden van ontwikkelingsmateriaal	45. kiezen
13. versje aanbieden/opzeggen	46. bespreken van de werkles/het binnen of buiten spelen
14. kringgesprek	47. opruimen
15. voorbereidend rekenen	48. jassen halen door middel van opdrachten
16. wereldoriëntatie	49. jassen aandoen/uitdoen
17. catechese	50. naar buiten/naar binnen gaan
18. taalspel	51. omkleden
19. zangspel	52. naar de gymzaal/aula/zolder gaan en weer terug
20. stiltespel	53. uitdelen van spullen
21. kimspeel	54. wegzetten/halen van plantjes
22. raadspel	55. in de kring gaan zitten/aan tafel gaan zitten
B. speelwerktijd	56. wachten
23. bouwhoek	E. dagelijkse dingen
24. poppenhoek	57. dagopening
25. lees-schrijfhoek	58. dagsluiting
26. boekenhoek	59. fruit halen/uitdelen
27. tekenen/kleuren/knutselen/schilderen	60. liedje/versje fruit
28. werkblad maken	61. fruit eten
29. spelen met constructiemateriaal	F. video
30. spelen met ontwikkelingsmateriaal	62. video kijken
31. puzzelen	
32. lees-schrijfpoddracht	63. ongerichte activiteit
33. in een schrift/werkboek werken	
34. gezelschapsspeel	
35. computerspeel	

stelt. 'Vertellen naar aanleiding van een verhaal' houdt in dat de leerlingen op een verhaal mogen reageren.

De categorie 'schrijven in de kring' wordt gecodeerd zodra de leerkracht met de kinderen een schrijfactiviteit uitvoert zoals het samen schrijven van een verhaal, van woordjes of van een brief.

Bij het 'aanbieden van een ontwikkelingsmateriaal' geeft de leerkracht uitleg over het (nieuwe) materiaal en wordt er in de kring mee gespeeld.

In een kringgesprek wordt gepraat over een bepaald onderwerp in tegenstelling tot 'vertellen door de kinderen', waarbij de kinderen vertellen over hun eigen belevenissen en ervaringen. Vertellen over het weekend wordt gecodeerd als 'vertellen door de kinderen', terwijl een gesprek over bijvoorbeeld het thema

carnaval de code 'kringgesprek' krijgt.

Wereldoriëntatie en catechese zijn 'lessen' die ook bij kleuters aan bod komen. Vaak bestaan die lessen uit voorlezen, spelletjes, gesprekjes en dergelijke, kortom activiteiten die een eigen code hebben. Als er echter duidelijk sprake is van een lesje over 'de natuur', 'geschiedenis', 'aardrijkskunde' of 'godsdienst' worden codes voor wereldoriëntatie en catechese gegeven.

In een kleuterklas worden veel spelletjes gedaan. Van Zeeland (1992) verstaat onder taalspel allerlei korte klassikale activiteiten die iets met taal te maken hebben. Rijmen in de kring is bijvoorbeeld een taalspel. Van een zangspel is sprake als er een spel met bewegingen wordt gedaan waarbij gezongen wordt. Een stiltespel is een spel waarbij kinderen goed moeten luisteren. Een voorbeeld is 'Tik tik wie

ben ik?' waarbij een kind moet raden wie er op zijn schouder tikt aan de hand van de stem van een kind. Een kimspeel is een spel waarbij voorwerpen onder een handdoek worden gelegd. De kinderen moeten dan raden wat er weg is of wat erbij is gekomen. Bij een raadspel wordt een raadsel opgegeven door de juf of een ander kind.

De activiteiten die plaatsvinden in de *speelwerktijd* spreken allemaal voor zich. De activiteiten binnen de setting *bewegen* zijn al gedeeltelijk aan bod gekomen bij de beschrijving van die setting. Van speelkwartier is sprake als er buiten wordt gespeeld met alle groepen van de hele school.

Bij *organisatie* hoort onder andere de organisatie van de speelwerktijd. Voor dat alle kinderen aan het werk kunnen, vindt de uitleg van de werkles plaats. De leerkracht vertelt dan welke werkjes er gedaan moeten worden en legt soms ook klassikaal uit hoe een werkje aangepakt moet worden. Dat gebeurt vooral aan het begin van de week als alle kinderen het werkje nog moeten doen. Bij 'werkjes uitlegen' vertelt de leerkracht welke kinderen welk werkje moeten gaan doen. 'Kiezen' houdt in dat de kinderen om de beurt zelf een werkje mogen kiezen. Met het bespreken van de werkles of het binnen of buiten spelen wordt bedoeld dat de leerkracht en de kinderen samen bespreken hoe een activiteit verlopen is.

In de kleuterklas worden de jassen vaak gehaald door middel van opdrachtjes, bijvoorbeeld 'alle kinderen die een rode broek aan hebben, mogen hun jas gaan halen'. Daarom is jassen halen door middel van opdrachten een aparte categorie.

Met 'omkleden' wordt het omkleden bedoeld voor het spelen in de speelzaal of voor de gymnastiekles. In de kleuterklas staan vaak plantjes op de tafels. Die worden aan het begin van de ochtend of middag weggezet door een kind en later weer teruggezet. Daarom kan ook 'wegzetten/halen van plantjes' gecodeerd worden bij organisatie.

Activiteiten die binnen de setting *dagelijkse dingen* vallen zijn de dagopening en de dagsluiting. Bij de dagopening horen het oplezen van de namen, het zingen van een liedje, het opzeggen van een versje of een gebedje, het doornemen van de kalender en het weer en eventueel ook nog kort vertellen. Onder de

dagsluiting verstaan we het zingen van een liedje, opzeggen van een versje of het doen van een spelletje ter afsluiting van de dag.

Er is ook nog een categorie 'ongerichtte activiteit' die apart staat van de zes settings. Als de observator géén van de genoemde activiteiten kan coderen, geeft hij de code 'ongerichtte activiteit'. Er is bijvoorbeeld sprake van ongerichtte activiteit als een kind in de speelwerktijd rondloopt zonder met een activiteit bezig te zijn.

Zoals eerder vermeld is het gedrag van kinderen onder te verdelen in talig en niet talig gedrag. De hoofdcategorieën van *talig gedrag* die geobserveerd worden, zijn: luistert naar de juf, luistert naar een kind, luistert naar iets of iemand anders, spreekt en overig talig gedrag. Subcategorieën van de hoofdcategorie 'luistert naar de juf' zijn onder andere: luistert naar de juf die iets vertelt, luistert naar de juf die een vraag stelt aan een kind, luistert naar de juf die voorleest en luistert naar de juf die uitleg/instructie geeft. Bij de hoofdcategorie 'luistert naar een kind' horen onder andere de subcategorieën luistert naar een kind dat een vraag van de juf beantwoordt, luistert naar een kind dat tegen een ander kind praat en luistert naar een kind dat 'voorleest'. De hoofdcategorie 'spreekt' bevat onder andere de subcategorieën vraagt iets aan de juf, beantwoordt vraag van de juf, praat tegen kind(eren), vertelt in de kring en zegt versje op. Onder 'overig talig gedrag' verstaan we alle gedrag waardoor taal ontwikkeld wordt, behalve spreken en luisteren. Voorbeelden van overig talig gedrag zijn: zingen, schrijven, rijmen, verwoorden bij plaatjes en spelen met taal-ontwikkelingsmateriaal. Rijmen is ook een vorm van spreken, maar omdat het expliciet in dienst staat van taalontwikkeling wordt het onder de categorie 'overig talig gedrag' ondergebracht. Hetzelfde geldt voor het verwoorden bij plaatjes.

Het gedrag van de leerkracht is in dit observatie-instrument niet expliciet opgenomen. Dit gedrag zit verweven in het gedrag van het kind. We nemen als voorbeeld de bovengenoemde activiteit: voorlezen. Als het gedrag van het kind is 'luistert naar de leerkracht die voorleest' spreekt het voor zichzelf dat de leerkracht voorleest. Is het gedrag van het kind 'stelt een vraag aan de leerkracht' dan is het

gedrag van de leerkracht waarschijnlijk 'luistert naar de leerling die een vraag stelt'.

Een voorbeeld van een ingevuld deel van het observatie-instrument is weergegeven in Tabel 3. Als er in een vak geen code wordt ingevuld, wil dat zeggen dat er niets is veranderd aan setting, activiteit of gedrag.

Tabel 3
Voorbeeld van een ingevuld deel van het observatie-instrument

Tijd	Setting	Activiteit	Gedrag
0.00	E	57	37
.30			39
1.00			29

De eerste observatie (bij tijd is 0.00) die ingevuld is, wil zeggen: de setting die voorkomt is setting E 'dagelijkse dingen', de activiteit is 57 'dagopening' en het gedrag van de geobserveerde leerling is 37 'zegt een versje op'. Met behulp van de codes en de verklaringen in Tabel 2 en de onderverdeling van de hoofdcategorieën van talig gedrag in subcategorieën kan men zo alle observaties uitwerken.

De try-out van het instrument vond plaats op een onderzoeksschool. Gedurende drie dagen werd in de kleutergroep een leerling geobserveerd. Daaruit kwam naar voren dat het instrument vrij goed functioneerde. Het werkte met een intervalstelsel. De observator observeerde om de 30 seconden de leerling en noteerde binnen die tijd de bijbehorende codes. Het bleek dat de tijdseenheid van 30 seconden goed gekozen was. Als er minder tijd was om te coderen was dat niet haalbaar, terwijl een eenheid van 1 minuut te lang was. Zoals eerder vermeld zijn er ook opnamen gemaakt van het verbale gedrag van de kinderen. De resultaten van deze opnamen worden later gepubliceerd.

De eigenlijke observaties vonden in de periode van februari 1996 tot en met april 1996 plaats. Op elk van de acht scholen werd gedurende de maandagochtend, de dinsdagmiddag, de woensdagochtend en de donderdagmiddag één kind geobserveerd.

Interrater-betrouwbaarheid

Om vast te stellen in hoeverre de door de observator gemaakte registraties betrouwbaar waren, is de interrater-betrouwbaarheid bepaald. Voorafgaand aan de eigenlijke observa-

ties is gedurende één dagdeel door twee observatoren geobserveerd. De intersubjectieve overeenstemming tussen de observatoren werd berekend voor setting, activiteit en gedrag van het kind. De overall Kappa's van setting, activiteit en gedrag waren respectievelijk 0.97, 0.95 en 0.69. Deze overall Kappa's zijn voldoende hoog. De overeenstemming wat betreft setting en activiteit zijn volgens de vuistregels voor het beoordelen van de kappas 'almost perfect' en wat gedrag betreft 'substantial' (Cohen, 1960). De overall Kappa van het gedrag van het kind is lager dan de andere. De gedragingen van het kind veranderen van de drie te observeren variabelen het meest en zijn het moeilijkst te observeren.

Verloop van de observaties

Aan het begin van de maandagochtend werd de observator door de leerkracht aan de kinderen voorgesteld. Meestal was het kind dat geobserveerd ging worden van te voren al ingelicht. De observator deed het te observeren kind het zenderje om en begon met observeren. Hij zat op een zodanige plaats dat hij het observatie-kind steeds goed kon zien en wisselde ook regelmatig van plaats. De indruk van de observator was dat het observatie-kind en doorgaans ook de leerkracht al snel vergeten waren dat hun doen en laten werd gadeslagen. Deze indruk werd achteraf bevestigd door de leerkracht. Ieder geobserveerd dagdeel leverde een protocol op, variërend van 10 tot 20 pagina's voor de ochtenden en 7 tot 13 pagina's voor de middagen. In de protocollen werd beschreven welke activiteiten, settings en gedragingen om de 30 seconden geobserveerd waren. Daarnaast werden de opnamen die gemaakt waren met het zenderje beluisterd en werden de taaluitingen van de geobserveerde leerling in het protocol opgenomen. De opnamen vormden een geschikt middel om de betrouwbaarheid en validiteit van de observaties achteraf te controleren. Alle observaties i.c. registraties zijn op deze wijze gecontroleerd en gewijzigd als er geen sprake van overeenstemming was tussen de observaties en de opnamen. Opgemerkt moet worden dat wijzigingen achteraf nauwelijks nodig waren.

2.3.2 De interviews

Als aanvulling op de observaties zijn alle leerkrachten bij wie is geobserveerd ook geïnterviewd. Bij deze leerkrachten zijn twee interviews afgenomen. Het eerste interview was bedoeld om informatie te verkrijgen over de klas, de leerkracht en het kind dat geobserveerd werd. Ook probeerden we door middel van dit interview een beeld te krijgen van het ochtend- en middagprogramma in een kleutergroep. Verder gingen we in het interview in op het taal-leesonderwijs in groep 2 om aanvullende informatie te verkrijgen bij de observaties. Het doel van het tweede interview was om zaken die naar voren kwamen bij de observaties en die nog niet bevestigd waren aan de orde te stellen. Op de inhoud en de resultaten van deze interviews gaan we in dit artikel niet verder in.

2.4 Analyse van de gegevens

Van alle observaties zijn protocollen gemaakt. Analyse van de observaties vond plaats met het programma SPSS. Per leerling zijn de geobserveerde codes ingevoerd, waarna gemiddelde percentages berekend werden voor de verschillende settings en voor de hoofdcategorieën van talig gedrag en niet-talig gedrag. Omdat veel activiteiten slechts een klein percentage van de tijd voor hun rekening namen, was het niet zinvol voor iedere leerling apart de gemiddelde percentages voor alle activiteiten uit te rekenen. De gemiddelden zijn berekend over 32 dagdelen, over 16 ochtenden en over 16 middagen.

Daarna zijn de acht leerlingbestanden samengevoegd en de gemiddelde percentages en standaarddeviaties berekend voor de verschillende settings en activiteiten en hoofdcategorieën van talig gedrag, wederom over 32 dagdelen, 16 ochtenden en 16 middagen. De gemiddelden en standaarddeviaties voor de

gedragingen binnen de hoofdcategorieën van talig gedrag zijn alleen berekend over 32 dagdelen omdat het talig gedrag van de leerlingen gedurende de ochtend en de middag niet veel verschilde.

Vervolgens is per setting nagegaan hoe de verdeling van activiteiten en hoofdcategorieën van talig gedrag eruit zag. We wilden hierdoor te weten komen welke (taal)activiteiten en welk (talig) gedrag binnen de verschillende settings (vaak) voorkwamen. Ook voor de belangrijkste activiteiten werd gekeken naar de verdeling van de hoofdcategorieën van talig gedrag. Dit om te onderzoeken welk (talig) gedrag hoe vaak binnen bepaalde activiteiten voorkwam.

3 Resultaten

3.1 Tijd besteed aan settings en activiteiten

In deze paragraaf geven we een overzicht van de tijd die door de geobserveerde leerlingen besteed wordt aan de settings en activiteiten, gemiddeld over alle dagdelen. Daarbij gaan we er van uit dat een schooldag in groep 2 gemiddeld 330 minuten duurt en dat een ochtend meestal 210 minuten vergt. Omdat de ochtend en middag vaak een verschillend programma laten zien, vermelden we ook de gemiddelden voor beide dagdelen apart.

Tijd besteed aan settings

Tabel 4 geeft een overzicht van de tijd die door de geobserveerde leerlingen gemiddeld besteed wordt aan verschillende settings. De tijd is omgerekend in procenten. Omdat 1% van een schooldag (330 minuten) 3.3 minuten is, is 2% $2 \times 3.3 = 6.6$ minuten etc. We ronden af op hele getallen, dus: 1% is 3 minuten, 2% is 7 minuten. Door de afronding ontstaan hier en daar

Tabel 4

Tijd aan settings besteed, gemiddeld over 32 dagdelen in acht scholen

Setting	gemiddelde tijd	SD	minimum	maximum
speelwerktijd	28% (93 min.)	7	22%	45%
organisatie	23% (76 min.)	6	14%	33%
bewegen	21% (69 min.)	7	9%	32%
kring	19% (62 min.)	6	10%	30%
dagelijkse dingen	8% (26 min.)	2	4%	11%
video	1% (3 min.)	2	0%	5%
Totaal	100% (329 min.)			

Tabel 5

Tijd aan settings besteed, gemiddeld over 16 ochtenden en 16 middagen in acht scholen

Setting	gemiddelde tijd	SD	minimum	maximum
OCHTEND				
speelwerktijd	26% (55 min.)	5	19%	32%
organisatie	22% (46 min.)	6	16%	32%
bewegen	21% (44 min.)	5	14%	30%
kring	20% (42 min.)	3	15%	24%
dagelijkse dingen	10% (21 min.)	3	6%	17%
video	1% (2 min.)	3	0%	8%
Totaal	100% (210 min.)			
MIDDAG				
speelwerktijd	32% (38 min.)	16	14%	66%
organisatie	26% (31 min.)	9	11%	36%
bewegen	21% (25 min.)	14	0%	43%
kring	15% (18 min.)	11	0%	38%
dagelijkse dingen	4% (5 min.)	4	1%	11%
video	2% (2 min.)	4	0%	10%
Totaal	100% (119 min.)			

kleine foutjes. Dit geldt ook voor de andere Tabellen.

Tabel 4 laat zien dat aan de speelwerktijd gemiddeld de meeste tijd wordt besteed. Daarna volgt organisatie, waaraan in totaal 76 minuten wordt besteed. Bewegen en kring nemen ieder ongeveer het vijfde deel van de tijd voor hun rekening. De setting 'video' neemt weinig tijd in beslag; er werd slechts op enkele scholen video gekeken, terwijl er geobserveerd werd. Uit de standaarddeviaties en de percentages vermeld onder minimum en maximum blijkt dat er tussen de leerkrachten aanzienlijke verschillen bestaan wat betreft de tijd die aan de verschillende settings wordt besteed.

Tabel 5 geeft een overzicht van de tijd die in de ochtend en middag aan settings is besteed.

Vergelijking van de gemiddelden van de ochtenden en de middagen laat zien, dat aan speelwerktijd, organisatie en kring 's morgens meer tijd wordt besteed dan 's middags. Ook de 'dagelijkse dingen' nemen 's ochtends veel meer tijd in beslag dan 's middags (10% tegenover 4%). Dit komt waarschijnlijk doordat in de ochtend vrijwel altijd een uitgebreide dagopening plaatsvindt, terwijl aan de opening van de middag slechts kort of helemaal geen aandacht wordt besteed. Bovendien wordt er in de ochtend altijd gegeten en gedronken. Opvallend is dat de standaarddeviaties in de middag veel hoger zijn dan in de ochtend, wat wijst op grote verschillen tussen de leerkrachten wat betreft de indeling van de middag. Bij de set-

ting speelwerktijd is er bijvoorbeeld een minimum van 14% en een maximum van 66%.

Omdat in de observatie-onderzoeken die in de inleiding zijn beschreven alleen in de ochtend is geobserveerd, vergelijken we onze ochtend-gegevens met de resultaten van die onderzoeken. In ons onderzoek neemt de speelwerktijd de meeste tijd in beslag, namelijk 55 minuten. Van Gennip en Ruijs (1992) vinden ongeveer hetzelfde aantal minuten, al rekenen zij er ook het opruimen bij. Opruimen behoort in ons onderzoek tot organisatie en duurt 10 minuten. Organisatie neemt in ons onderzoek 46 minuten in beslag. Reezigt, Guldmond en Ros (1995) komen in hun onderzoek tot de conclusie dat organisatie een te groot deel van de ochtend beslaat. Uit hun resultaten komt echter naar voren dat organisatie 30 minuten kost. Volgens Reezigt et al. wordt aan buiten spelen de meeste tijd besteed, namelijk 48 minuten. Volgens Harskamp, Pijl en Snippe (1992) wordt 60 minuten aan bewegen besteed. In ons onderzoek vinden we dat 44 minuten aan bewegen wordt besteed. De kring duurt in ons onderzoek 42 minuten. Uit de resultaten van Reezigt et al. blijkt dat de kringactiviteiten 52 minuten duren. Eten en drinken wordt door hen echter ook tot de kringactiviteiten gerekend. De dagelijkse dingen beslaan in ons onderzoek 21 minuten. Dat is aanzienlijk minder dan de 41 minuten die Van Gennip en Ruijs vonden. Uit het bovenstaande blijkt opnieuw dat gegevens van observatie-onderzoeken waarin met

verschillende observatiecategorieën wordt gewerkt verschillen en moeilijk met elkaar zijn te vergelijken.

Tijd besteed aan activiteiten

We hebben ook de tijd die gemiddeld aan de verschillende activiteiten besteed wordt over de 32 dagdelen berekend. Veel activiteiten nemen slechts een klein percentage van de tijd voor hun rekening. In het volgende besteden we per setting alleen aandacht aan die activiteiten die relatief de nodige tijd vragen (minstens 1% van de tijd; dat wil zeggen 3 minuten of meer).

Zoals uit Tabel 6 blijkt neemt in de setting 'kring' de activiteit 'vertellen door de kinderen' gemiddeld 5% of 17 minuten van de tijd in beslag. Hierbij moet worden opgemerkt dat de observaties op de maandag- en woensdagoch-

tend plaatsvonden. Gedurende de maandagochtend werd op alle scholen door de kinderen verteld over het weekend. Als de observaties op bijvoorbeeld de dinsdagochtend hadden plaatsgevonden, zou vertellen door de kinderen waarschijnlijk een minder groot aandeel hebben gehad. Het 'vieren van een verjaardag' vraagt 3% van de tijd, terwijl het 'voorlezen door de leerkracht' gemiddeld 7 minuten in beslag neemt.

Volgens Reezigt, Guldemond en Ros (1995) komen activiteiten als vertellen door de kinderen en voorlezen in de kring slechts weinig voor en duren ze kort. De gemiddelde tijd die aan voorlezen wordt besteed, is volgens hen 1% of twee minuten per ochtend. Van Gennip en Ruijs (1992) observeerden evenals wij vaker vertellen en voorlezen in de kring. De gemiddelde tijdsduur van voorlezen bedroeg bij hen

Tabel 6
Tijd aan activiteiten besteed, verdeeld over de settings, gemiddeld over 32 dagdelen in acht scholen

Activiteiten	gemiddelde tijd	SD	minimum	maximum
KRING				
vertellen door de kinderen	5% (17 min.)	4	0%	11%
verjaardag vieren	3% (10 min.)	3	0%	8%
voorlezen door de leerkracht	2% (7 min.)	3	0%	7%
catechese	1% (3 min.)	3	0%	8%
SPEELWERKTIJD				
tekenen/kleuren/knutselen/schilderen	11% (36 min.)	9	0%	23%
spelen met constructiemateriaal	3% (10 min.)	6	0%	17%
spelen met ontwikkelingsmateriaal	3% (10 min.)	5	0%	15%
lees-schrijfpodracht	2% (7 min.)	4	0%	10%
lees-schrijfhoek	2% (7 min.)	3	0%	8%
in een schrift/werkboek werken	1% (3 min.)	2	0%	5%
werkblad maken	1% (3 min.)	4	0%	10%
bouwhoek	1% (3 min.)	3	0%	9%
BEWEGEN				
buiten spelen met materiaal	13% (43 min.)	9	0%	25%
binnen spelen met materiaal/toestellen	5% (17 min.)	5	0%	13%
speelkwartier	1% (3 min.)	3	0%	7%
geleide gymnastiekles	1% (3 min.)	2	0%	9%
buiten spelen zonder materiaal	1% (3 min.)	3	0%	7%
ORGANISATIE				
opruimen	5% (17 min.)	2	4%	9%
wachten	4% (13 min.)	3	0%	9%
omkleden	3% (10 min.)	2	0%	6%
halen, aan- en uitdoen van jassen	2% (7 min.)	2	0%	5%
werkjes uitdelen	1% (3 min.)	1	0%	3%
uitdelen van spullen	1% (3 min.)	1	0%	3%
naar buiten/naar binnen gaan	1% (3 min.)	1	1%	2%
DAGELIJKSE DINGEN				
fruit eten	3% (10 min.)	1	2%	5%
dagopening	3% (10 min.)	1	1%	5%
fruit halen/uitdelen	1% (3 min.)	1	0%	2%
VIDEO				
video kijken	1% (3 min.)	2	0%	4%

20 minuten per ochtend. Bij een onderzoek van Bus (1990) onder 22 leerkrachten las 86% eens per dag voor en 16% twee maal per dag.

Tijdens de *speelwerk*tijd wordt een groot gedeelte van de tijd getekend, gekleurd, geknutseld en geschilderd, namelijk gedurende 11% van de tijd of 36 minuten. 'Spelen met constructiemateriaal' komt op de tweede plaats en daarna volgen 'spelen met ontwikkelingsmateriaal', 'werken aan een lees-schrijfopdracht', 'spelen in de lees-schrijfhoek' et cetera.

Wat er tijdens de speelwerk

tijd gebeurt, is door de andere onderzoekers niet gericht geobserveerd. Van Gennip en Ruijs (1992) constateren wel dat er veel gebruik gemaakt wordt van speelleermaterialen (lotto's, puzzels, werkbladen), van constructie- en compositiematerialen (bijvoorbeeld bouwmateriaal) en van vormgevende materialen (teken- en schildermateriaal, knutselmateriaal). In ons onderzoek heeft het werken met vormgevend materiaal tijdens de speelwerktijd de overhand.

Bij de setting *bewegen* neemt het 'buiten spelen met materiaal' het grootste deel van de tijd in beslag, namelijk 13% of 43 minuten. Uit de observaties komt, evenals uit het onderzoek van Van Gennip en Ruijs (1992), naar voren dat binnen spelen alleen maar plaatsvindt als het slecht weer is. Als het weer het maar enigszins toelaat, gaan de leerkrachten met de kinderen naar buiten.

Binnen de setting *organisatie* neemt 'opruimen' de grootste plaats in, namelijk 5% of 17 minuten. De geobserveerde leerlingen zijn gemiddeld 13 minuten aan het wachten, meestal als ze al klaar zijn met opruimen. Er wordt 10 minuten besteed aan omkleden voor het binnen spelen of de gymnastiekles en 7 minuten aan het halen, aan- en uitdoen van de jassen. Omdat andere onderzoekers de setting organisatie niet zover hebben uitgesplitst, kunnen we hier geen vergelijking maken.

Binnen de setting *dagelijkse dingen* zijn 'fruit eten' en de 'dagopening' de activiteiten die de meeste tijd in beslag nemen, namelijk 10 minuten. Het halen en uitdelen van fruit neemt daarbij nog 3 minuten in beslag. Van Gennip en Ruijs (1992) constateren dat eten en drinken gemiddeld 17 minuten duurt. In ons onderzoek duurt de dagopening 10 minuten en bij Van Gennip en Ruijs in totaal 17 minuten.

De tijd dat de geobserveerde leerlingen naar de *video* keken, bedraagt gemiddeld 1% of 3 minuten.

3.2 Tijd besteed aan talig gedrag

In deze paragraaf geven we een overzicht van de tijd die door de geobserveerde leerlingen wordt besteed aan talig gedrag, gemiddeld over alle dagdelen. Bovendien vermelden we de tijd die binnen de settings en activiteiten aan talig gedrag wordt besteed. Omdat er weinig verschil bestaat tussen de tijd die de kinderen 's ochtends en 's middags aan talige gedragingen besteden, vermelden we hier niet de gemiddelden voor de dagdelen apart. Wat betreft de tijd die binnen de activiteiten aan talige gedragingen (luistert naar de juf die om stilte vraagt e.d.) wordt besteed, beperken we ons tot de meest voorkomende activiteiten.

Tijd besteed aan talig en niet talig gedrag

We vermelden de tijd die de geobserveerde leerlingen aan gedrag besteden, gemiddeld over de 32 dagdelen. Van al het geobserveerde gedrag is gemiddeld 53% niet talig en 47% talig. De geobserveerde leerlingen zijn 28% van de tijd of 92 minuten bezig met luisteren, waarvan 19% naar de juf, 8% naar een medeleerling en 1% naar iets of iemand anders (een ouder, de computer, een cassettebandje). Gesproken wordt er gemiddeld 12% van de tijd of 40 minuten van de 330 minuten. Het overig talig gedrag, waaronder schrijven, lezen en het bezig zijn met taalontwikkelings- of leeschrijfmateriaal neemt gemiddeld 22 minuten in beslag. In het onderzoek dat in de inleiding is beschreven, is het talig gedrag van de leerlingen niet geobserveerd. We kunnen onze gegevens dan ook niet vergelijken met die uit ander Nederlands onderzoek.

Tijd besteed aan talig en niet talig gedrag binnen de settings

De tijd die de kinderen van groep 2 binnen de onderscheiden settings aan talig en niet talig gedrag besteden, wordt in Tabel 7 vermeld.

In de *kring* is het percentage talig gedrag relatief hoog, namelijk 15% of 49 minuten en het percentage niet talig gedrag laag, namelijk 4%. Wat opvalt is dat in de kring veel geluisterd wordt. De tijd dat de leerlingen luisteren, bedraagt in totaal 43 minuten, waarvan 33

Tabel 7
Tijd aan talig en niet talig gedrag besteed, gemiddeld over 32 dagdelen in acht scholen

Gedrag	gemiddelde tijd
KRING	
luistert naar de juf	10% (33 min.)
luistert naar een kind	3% (10 min.)
luistert naar iets of iemand anders	0% (0 min.)
spreekt	1% (3 min.)
overig talig gedrag	1% (3 min.)
niet talig gedrag	4% (13 min.)
SPEELWERKTIJD	
luistert naar de juf	1% (3 min.)
luistert naar een kind	2% (7 min.)
luistert naar iets of iemand anders	0% (0 min.)
spreekt	5% (17 min.)
overig talig gedrag	4% (13 min.)
niet talig gedrag	16% (53 min.)
BEWEGEN	
luistert naar de juf	1% (3 min.)
luistert naar een kind	1% (3 min.)
luistert naar iets of iemand anders	0% (0 min.)
spreekt	3% (10 min.)
overig talig gedrag	1% (3 min.)
niet talig gedrag	15% (50 min.)
ORGANISATIE	
luistert naar de juf	5% (17 min.)
luistert naar een kind	1% (3 min.)
luistert naar iets of iemand anders	1% (3 min.)
spreekt	2% (7 min.)
overig talig gedrag	0% (0 min.)
niet talig gedrag	14% (46 min.)
DAGELIJKSE DINGEN	
luistert naar de juf	2% (7 min.)
luistert naar een kind	1% (3 min.)
luistert naar iets of iemand anders	0% (0 min.)
spreekt	1% (3 min.)
overig talig gedrag	0% (0 min.)
niet talig gedrag	4% (13 min.)
VIDEO	
luistert naar de juf	0% (0 min.)
luistert naar een kind	0% (0 min.)
luistert naar iets of iemand anders	0% (0 min.)
spreekt	0% (0 min.)
overig talig gedrag	1% (3 min.)
niet talig gedrag	0% (0 min.)
TOTAAL	100% (329 min.)

minuten naar de juf. De juf is in de kring dus veel aan het woord. De tijd dat de leerlingen zelf in de kring aan het woord zijn, is 3 minuten of 1%. Ook het percentage overig talig gedrag is in deze setting aan de lage kant. Bij deze gegevens moet worden opgemerkt dat de observator niet zeker weet of een kind echt luistert. De code luisteren is genoteerd als iemand aan het woord was en het leek of het

betreffende kind luisterde. Het kan natuurlijk voorkomen dat het kind helemaal niet opneemt wat er gezegd wordt en dus niet luistert.

Het percentage talig gedrag is in de *speelwerk*tijd lager dan in de kring, namelijk 12% of 40 minuten. In de speelwerktijd spreken de kinderen echter veel meer dan in de kring: gemiddeld 5% of 17 minuten van de tijd. De speelwerktijd is blijkbaar de setting om met andere leerlingen te communiceren. Het percentage dat de leerlingen luisteren is 3% of 10 minuten. Het percentage dat de leerlingen naar de juf luisteren is niet hoog. Reezigt et al. (1995) merken op dat veel leerlingen tijdens de werkles niet door de leerkracht worden begeleid en 'onbewaakt' spelen. De leerlingen hebben tijdens de speelwerktijd niet veel contact met de leerkracht. Aan het overig talig gedrag wordt in de speelwerktijd 13 minuten besteed. Een kind kan bezig zijn met lezen, schrijven, met een taelopdracht of met materiaal dat betrekking heeft op taal. Eerder zagen we al dat de leerlingen de speelwerktijd voor een groot deel doorbrengen met tekenen en knutselen en spelen met constructiemateriaal.

Binnen de setting *bewegen* neemt het niet talig gedrag een grote plaats in: 15% van de tijd of 50 minuten vertonen de leerlingen niet talig gedrag. In het talige gedrag dat de leerlingen vertonen, heeft het spreken het grootste aandeel, namelijk 3%. Opgemerkt moet worden dat tijdens het bewegen en vooral tijdens het buiten spelen het vaak moeilijk was om te observeren of een leerling sprak of luisterde. Spreken en luisteren zijn alleen gecodeerd indien de observator duidelijk zag dat het plaatsvond. Het is mogelijk dat de leerlingen meer gesproken en geluisterd hebben. De percentages voor spreken en luisteren zijn waarschijnlijk dan ook hoger dan hier is weergegeven.

Het percentage niet talig gedrag gedurende de setting *organisatie* ligt hoog, het is 14% of 46 minuten. De kinderen luisteren gedurende 7% van de tijd, waarvan 5% naar de juf. Ook binnen de setting *dagelijkse dingen* neemt het percentage niet talig gedrag een groot deel in, namelijk 4% of 13 minuten.

Binnen de setting *video* is het percentage niet talig gedrag zeer gering. Het percentage overig talig gedrag is relatief hoog (1% of 3 minuten). Dat komt omdat onder overig talig

gedrag ook het kijken (en luisteren) naar de video wordt verstaan.

Tijd besteed aan talig en niet talig gedrag binnen de activiteiten

Wat betreft de tijd die binnen de activiteiten aan talige gedragingen besteed wordt, beperken we ons hier tot de settings waarin het meeste talig gedrag voorkomt: de kring en de speelwerktijd.

Binnen de setting *kring* neemt 'vertellen door de kinderen' gemiddeld 17 minuten in beslag (zie Tabel 6). Een groot deel van deze tijd is er sprake van talig gedrag; er wordt voornamelijk geluisterd, 8 minuten naar de juf en 5 minuten naar kinderen. De leerlingen zelf zijn gemiddeld slechts 1 minuut aan het woord. 'Verjaardag vieren' duurt gemiddeld 10 minuten. Daarvan luisteren de leerlingen gedurende 2 minuten naar de juf. 'Voorlezen door de leerkracht' neemt gemiddeld 7 minuten in beslag, waarvan de leerlingen 6 minuten naar de juf luisteren.

Binnen de setting *speelwerktijd* wordt bij de activiteit 'tekenen/kleuren/knutselen/schilderen' die gemiddeld 36 minuten duurt, slechts 11 minuten talig gedrag vertoond. De leerlingen luisteren gedurende 4 minuten en spreken 6 minuten. Bij het 'spelen met constructiemateriaal' (duur 10 minuten) zijn de leerlingen 4 minuten talig bezig: ze spreken 2 minuten en luisteren 2 minuten.

Bij de activiteit 'spelen met ontwikkelingsmateriaal' wordt er weer meer talig gedrag vertoond. Van de 10 minuten zijn de leerlingen 7 minuten talig bezig. Een groot gedeelte van die tijd, 5 minuten, wordt in beslag genomen door overig talig gedrag. Dat gedrag bestaat dan voornamelijk uit het bezig zijn met taalontwikkelingsmateriaal en uit lezen, verwoorden bij plaatjes en rijmen. De leerlingen luisteren gemiddeld gedurende 1 minuut naar de juf. Gedurende 1 minuut wordt er gesproken.

Het talig gedrag neemt binnen de activiteit 'werken aan een lees-schrijfpdracht' 4 van de 7 minuten in beslag. Ook hier neemt overig talig gedrag een grote plaats in: het beslaat 3 minuten. In die tijd zijn de leerlingen bezig met het maken van een opdracht die met lezen en/of schrijven te maken heeft, dan wel met lezen of schrijven. De leerlingen luisteren vrijwel niet en spreken gedurende 1 minuut.

Binnen de activiteit 'spelen in de lees-schrijfhoek', die gemiddeld 7 minuten duurt, vertonen de leerlingen gedurende 5 minuten talig gedrag. Er wordt gedurende 1 minuut gesproken en 2 minuten geluisterd, waarvan 1 minuut naar iets of iemand anders (in de meeste gevallen een moeder) en bijna niet naar de juf. Het overig talig gedrag duurt 2 minuten; de leerlingen zijn dan bezig met lezen, schrijven en materialen die daarop betrekking hebben.

4 Conclusies en discussie

In ons onderzoek stonden twee vragen centraal. De eerste vraag luidde: 'Welke activiteiten voeren leerlingen en leerkrachten in groep 2 uit en hoeveel tijd nemen deze activiteiten in beslag?' Uit het observatie-onderzoek komt naar voren dat een dag in groep 2 is onder te verdelen in de settings kring, speelwerktijd, bewegen, dagelijkse dingen en organisatie.

De *speelwerktijd* is de setting die de meeste tijd beslaat. Binnen de speelwerktijd wordt veel tijd besteed aan tekenen, kleuren, knutselen en schilderen. Daarnaast spelen de leerlingen veel met constructie- en met ontwikkelingsmateriaal. Van de hoeken wordt weinig gebruik gemaakt.

Gedurende een schooldag wordt een aanzienlijk percentage van de tijd aan *organisatie* besteed. Opruimen, wachten en omkleden zijn de activiteiten die binnen deze setting de meeste tijd vragen.

Per dag worden er gemiddeld 69 minuten besteed aan *bewegen*. Als het weer het toelaat, gaat de leerkracht met de kinderen naar buiten. Buiten spelen beslaat dan ook 71% van de setting 'bewegen'. Als er binnen wordt gespeeld, is dat meestal met materiaal.

De geobserveerde leerlingen zitten gemiddeld gedurende een uur per dag in de *kring*. Activiteiten die we binnen de kring veel geobserveerd hebben, zijn het vertellen door de kinderen over hun belevenissen, het vieren van een verjaardag en het voorlezen door de leerkracht.

De *dagelijkse dingen* beslaan 8% van de dag. 's Morgens vindt de dagopening plaats en wordt er fruit gegeten, 's middags is er af en toe een dagsluiting.

Uit dit alles concluderen we dat er gedurende de dag veel psychomotorische en weinig cognitieve activiteiten plaatsvinden.

De tweede vraag was: 'Welke talige gedragingen voeren leerlingen en leerkrachten uit in groep 2 en hoeveel tijd vragen deze gedragingen?'

Uit de observaties komt naar voren dat de leerlingen gedurende 47% van de dag talig gedrag vertonen. Als in de basisschool het taalonderwijs centraal staat, dan is de verhouding talig en niet-talig gedrag scheef. De leerlingen zouden gedurende meer tijd talig gedrag moeten vertonen. De kinderen luisteren erg veel, namelijk 92 minuten van de 330. Er wordt daarentegen weinig gesproken, gemiddeld maar 12% van de dag (40 minuten). Het percentage overig talig gedrag is ook gering.

Uit de observaties blijkt dat de *speelwerk*tijd de setting is waarin veel gecommuniceerd wordt: de leerlingen praten gedurende 18% van de tijd. Er is echter maar weinig interactie tussen de leerkracht en de leerlingen en ook tussen leerlingen onderling. De leerkracht zou meer aandacht kunnen besteden aan interacties en dialogen met en tussen individuele leerlingen en groepjes van leerlingen (Van Oers & Janssen-Vos, 1992; Verhoeven & Aarmoutse, 1996).

Verder zijn we van mening dat er in de speelwerktijd meer speelwerksituaties gecreëerd kunnen worden, die gericht zijn op de taalontwikkeling (Inspectie van het onderwijs, 1994; Stevens & Harskamp, 1996). Gedurende het grootste gedeelte van de speelwerktijd wordt er door de leerlingen geknutseld en tijdens deze activiteit vertonen de leerlingen maar weinig talig gedrag. Ook kan er meer en beter gebruik gemaakt worden van de hoeken. In de hoeken kan het taalaanbod verbeterd worden (Janssen-Vos, 1992).

Binnen de settings *organisatie* en *dagelijkse dingen* vertonen de geobserveerde leerlingen weinig talig gedrag. De overgangsmomenten zoals wachten, omkleden, jassen aandoen en fruit eten kunnen meer uitgebuit worden door er taalactiviteiten in te verweven (Van Zeeland, 1992).

Binnen de setting *bewegen* is maar één keer een talige activiteit geobserveerd. Deze setting biedt wel mogelijkheden om activiteiten te

doen die betrekking hebben op taal. Door middel van kleine rollenspelen of oefeningen in begripontwikkeling kan de leerkracht ervoor zorgen dat ook in deze setting meer talig gedrag vertoond wordt (Harskamp, Pijl & Snippe, 1992).

De *kring* is de setting waarvan we verwachten dat er de meeste specifieke taalactiviteiten voorkomen. De kringactiviteiten die we het meest observeerden, waren vertellen door de kinderen, verjaardag vieren en voorlezen door de leerkracht. Er is echter een veel grotere variatie in taalactiviteiten binnen de kring mogelijk (Reezigt, Guldmond & Ros, 1995).

De leerlingen vertoonden gedurende 79% van de kringtijd talig gedrag. We constateerden echter dat de meeste tijd geluisterd wordt, vooral naar de juf. De leerling is maar gedurende een zeer geringe percentage van de tijd zelf aan het woord, namelijk 5%. Activiteiten in een kleinere kring zijn niet geobserveerd. In een kleinere kring is er echter sprake van veel meer interactie tussen de leerkracht en de leerlingen en de leerlingen onderling en zal de leerling zelf meer aan het woord kunnen zijn. Daarom is het belangrijk dat er regelmatig activiteiten gedaan worden met enkele leerlingen in een kleine kring (Damhuis, 1988; Gaffney & Anderson, 1991; Slavin, 1995; Van Elsäcker & Verhoeven, 1997).

Samenvattend kunnen we concluderen dat er in de praktijk van de kleutergroep maar weinig sprake is van interactie tussen de leerkracht en de leerlingen en tussen de leerlingen onderling, terwijl deze interacties juist van zeer groot belang zijn voor de taalontwikkeling. Tevens kan het taalaanbod uitgebreid en verbeterd worden.

Door middel van dit onderzoek hebben we geprobeerd een zo volledig mogelijk overzicht te geven van de praktijk van het (taal)onderwijs in groep 2, door op acht scholen telkens één leerling te volgen. Het onderzoek kent echter zijn beperkingen. Het zou de generaliseerbaarheid van de onderzoeksgegevens ten goede komen als meerdere leerlingen op meerdere scholen geobserveerd waren. Bovendien kan er pas een echt plaatje van een week onderwijs worden gemaakt, als er ook gedurende een hele week wordt geobserveerd. De gegevens die in

dit onderzoek gevonden zijn, dienen dan ook voorzichtig geïnterpreteerd te worden.

Naar aanleiding van de tweede onderzoeksvraag werd nagegaan welke taalactiviteiten er plaats vinden in groep 2. Het zou interessant zijn om ook de kwaliteit van de taaluitingen van de geobserveerde leerlingen te onderzoeken met behulp van de opnamen die gemaakt zijn. Dit gaat in de toekomst ook gebeuren.

We geven nog een laatste suggestie voor verder onderzoek. In dit onderzoek zijn alleen 'doorsnee' scholen geselecteerd. Het zou goed zijn om ook te observeren op onderwijsvoorrangsscholen en de activiteiten die daar plaats vinden te vergelijken met die op de 'doorsnee' scholen.

Literatuur

- Aarnoutse, C., Glopper, K. de, Litjens, P., Sijstra, J., & Vernooij, K. (1995). *Probleemidentificatie en aanzet voor een actieplan taal*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- ARBO (1990). *Spelen en jongleren. Advies over de vernieuwing van het onderwijs aan jonge kinderen*. Zeist: ARBO.
- Bus, A. (1990). Van voorlezen naar lezen. *Moer*, 7, 320-327.
- Bus, A., & Verhoeven, L. (1992). *Instructie in beginnend lezen*. Delft: Eburon.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Cohen de Lara, H. (1996). *Theorie en praktijk. Wat hebben onderwijsvisies onderbouwleerkrachten te bieden?* Hoevelaken: CPS.
- Commissie Evaluatie Basisonderwijs (1994). *Onderwijs aan jonge kinderen. Evaluatie van het basisonderwijs. Deelrapport 3*. De Meern: Inspectie van het basisonderwijs.
- Damhuis, R. (1988). *Tweede-taalverwerving in kleutergroepen. Een onderzoek naar de gelegenheid tot het leren van het Nederlands door Turkse en Marokkaanse kleuters*. SCO-rapport nr. 164. Amsterdam: SCO.
- Elsäcker, W. van, & Verhoeven, L. (1997). Kleuters leren meer van voorlezen in kleine groepen. *Pedagogische Studiën*, 74, 117-129.
- Gaffney, J.S., & Anderson, R.C. (1991). *Two-tiered scaffolding: Congruent processes of teaching and learning*. Technical report nr. 523. Campaign: Center for the Study of Reading.
- Gennip, H. van, & Ruijs, A. (1992). *Aan de basis. Onderzoek naar de pedagogisch-didactische kwaliteit van het onderwijs aan jonge kinderen*. Nijmegen: ITS.
- Harskamp, E.G., Pijl, Y.J., & Snippe, J. (1992). *Spel en spelstimulering in het primair onderwijs*. Groningen: RION.
- Inspectie van het Onderwijs (1992). *Goed onderwijs aan jonge kinderen. Een evaluatieve beschrijving van aspecten van het onderwijs aan jonge kinderen in de basisschool*. De Meern: Inspectie van het onderwijs.
- Inspectie van het Onderwijs (1994). *De toestand van het onderwijs aan jonge kinderen*. De Meern: Inspectie van het onderwijs.
- Janssen-Vos, F. (1992). *Basisontwikkeling*. Assen: Van Gorcum.
- Janssen-Vos, F. (1996). *Naar continuïteit en samenhang. Bouwstenen voor een actieprogramma onderbouw*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- Janssen-Vos, F., & Laevers, F. (Red.) (1996). *De opvang van en het onderwijs aan jonge kinderen. Een comparatieve studie betreffende de situatie in Vlaanderen en Nederland*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- Janssen-Vos, F., Pompert, B., & Vink, H. (1995). *Naar lezen, schrijven en rekenen* (2e druk). Assen: Van Gorcum.
- Jongerius, J., & Beernink, R. (1993). *Wat heb je vandaag gedaan op school. Onderwijs in groep 1 en 2 van de basisschool*. Enschede: SLO.
- Kuyk, J.J. van (1991). Een schets voor het onderwijs aan jonge kinderen. *Pedagogische Studiën*, 68, 378-389.
- Meer T.A.M. van der, Appelfhof, P.N., Doesschate, J.J. ten, & Kramer, G.B. (1986). *Het onderwijs in de klas tussen 9 en 12*. Utrecht: SAC.
- Netelenbos, T. (1995). *Een impuls voor het basisonderwijs. Beleidsplan naar aanleiding van het rapport van de Commissie Evaluatie Basisonderwijs*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- Oers, B. van, & Janssen-Vos, F. (Red.) (1992). *Visies op onderwijs aan jonge kinderen*. Assen/Maastricht: Van Gorcum.
- Reezigt, G. (1996). Onderzoek naar vormgeving en effecten van het onderwijs aan jonge kinderen in Nederland. *Tijdschrift voor Onderwijsresearch*, 21, 7-32.

- Reezigt, G., & Ros, A (1993). Kleuters kunnen vaak zelf kiezen wat ze willen doen. *Didactief*, 23 (6), 29-30.
- Reezigt, G.J., Guldemond, H., & Ros, A.A. (1995). *Effectief onderwijs aan jonge kinderen*. Groningen: RION.
- Schaffer, E.C., & Nesselrodt, P.S. (1992). *The development and testing of the Special Strategies Observation System*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.
- Slavin, R.E. (1995). *Cooperative learning*. Needham Heights, MA: Allyn & Bacon.
- Stevens, L.M., & Harskamp, E.G. (1996). Actuele opvattingen over onderwijs aan jonge kinderen. *Tijdschrift voor Onderwijsresearch*, 21, 81-94.
- Verhoeven, L. (1994). *Ontluikende geletterdheid: een overzicht van de vroege ontwikkeling van lezen en schrijven*. Lisse: Swets & Zeitlinger.
- Verhoeven, L., & Aarnoutse, C. (1996). Verbetering van het onderwijs Nederlands: een plan van aanpak. *Spiegel*, 14 (2), 53-69.
- Westerlaak, J. M. van, Kropman, J. A., & Collaris, J.W.M. (1975). *Beroepenklapper*. Nijmegen: ITS.
- Zeeland, R. van, (1992). *Taal bij kleuters*. In Nijmeegse werkgroep, *Taaldidactiek aan de basis*. (vierde druk) (pp. 161-199). Groningen: Wolters-Noordhoff.

Manuscript aanvaard 27-6-1997

Auteurs

D. Manders is projectmedewerkster bij de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen.

C. Aarnoutse is als universitair hoofddocent verbonden aan de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen.

Adres: Katholieke Universiteit Nijmegen, Vakgroep Onderwijskunde, Montessorilaan 3, 6525 HR Nijmegen.

Abstract

What do kindergarten pupils and teachers do?

D. Manders & C. Aarnoutse. *Pedagogische Studiën*, 1997, 74, 338-354.

The goal of this study was to investigate which activities and linguistic behaviors students (age 5 or 6) and teachers from kindergarten perform and what amount of time is spent on these activities and behaviors. On eight ordinary elementary schools one child was observed during two mornings and afternoons. The observation instrument consisted of four variables: time, setting, activity and linguistic behavior of the child. The data revealed that during a day many psychomotoric activities take place and relatively few cognitive activities. The most important conclusion is that in kindergarten's practice there is too little interaction between the teacher and the students and between the students themselves. Moreover, the language input in kindergarten can be extended and improved.

Onverwerkt verlies, gedesororganiseerde gehechtheid en beangstigend opvoedersgedrag: Een vooronderzoek*

C. Schuengel, M.H. van IJzendoorn, M.J. Bakermans-Kranenburg en M. Blom**

Samenvatting

Onverwerkt verlies bij ouders blijkt samen te gaan met gedesororganiseerde gehechtheid bij jonge kinderen (Van IJzendoorn, 1995). Een verklaring die is geopperd voor dit verband veronderstelt dat deze ouders zich beangstigend gedragen (Main & Hesse, 1990). We observeerden het gedrag van 16 moeders bij hen thuis. Onverwerkt verlies werd vastgesteld op basis van het gehechtheidsbiografisch interview (George, Kaplan & Main, 1985). Gedesororganiseerde gehechtheid stelden we vast in de Vreemde Situatie procedure. Beangstigend gedrag bleek voor te komen tijdens de thuisobservaties, en samen te hangen met onverwerkt verlies. We vonden echter vooralsnog geen samenhang tussen beangstigend gedrag en desorganisatie.

Inleiding

Het verlies van een geliefde persoon is voor veel mensen de meest ingrijpende ervaring in hun leven. Het is voor de achterblijvers na een sterfgeval vaak moeilijk om de gevoelens van rouw en verdriet te verdragen. Niet alleen blijft er een leegte achter, het sterfproces zelf is voor sommigen ook een uiterst beangstigende ervaring. Bij de meeste mensen lijkt de ervaring na enige tijd te zijn verwerkt, zonder negatieve

gevolgen. Niet iedereen slaagt er echter even goed in om de gevoelens die gepaard gaan met verlies te verwerken. Bij ouders blijken deze onverwerkte verlieservaringen te worden overgedragen op de kinderen (Main & Hesse, 1990; Ainsworth & Eichberg, 1991). Hoe dat gebeurt is echter nog onduidelijk. Met name is onduidelijk welk gedrag van de ouder verantwoordelijk is voor de overdracht. Deze studie gaat over de mogelijkheid om via observatie van het gedrag van moeders het mechanisme van overdracht in beeld te brengen. Hierbij gaan we uit van recente inzichten die zijn ontwikkeld vanuit de gehechtheidstheorie.

Bowlby (1989) heeft beschreven hoe kinderen in hun eerste levensjaar gehecht raken aan vertrouwde personen in hun omgeving, meestal de ouders. Deze gehechtheid uit zich in het zoeken van nabijheid van en contact met deze personen wanneer kinderen moe, ziek of bang zijn. Voor kinderen vormen de vertrouwde personen een veilige haven (Ainsworth, 1969), vanwaaruit ze de wereld kunnen verkennen. Ainsworth, Blehar, Waters en Wall (1978) hebben echter aangetoond dat er aanzienlijke kwalitatieve verschillen zijn in de manieren waarop kinderen hun ouders kunnen gebruiken als veilige haven, en dat die verschillen te maken hebben met de manier waarop ouders ingaan op de signalen van het kind.

Ainsworth en haar collega's ontwierpen een laboratoriumprocedure waarmee verschillen tussen één- tot anderhalfjarige kinderen in de kwaliteit van gehechtheidsrelaties aan het licht konden worden gebracht: de Vreemde Situatie. In deze procedure worden kinderen blootgesteld aan licht stressvolle omstandigheden, vooral doordat het kind enige minuten alleen gelaten wordt door de opvoeder – in een vreemde omgeving. De eerste keer gebeurt dat met een voor het kind onbekende persoon, de tweede keer is het kind helemaal alleen.

Main (1990) heeft de verschillen in kwali-

* Dit onderzoek is mede mogelijk gemaakt door een Pionierssubsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) aan M.H. van IJzendoorn (PGS 59-256).

** De auteurs zijn dank verschuldigd aan M. Main en E. Hesse van de Universiteit van Californië in Berkeley voor hun bijdrage aan het beoordelen van het oudergedrag. De auteurs danken T. Joels van de Universiteit van Haifa voor haar hulp bij de codering van de Vreemde Situatie procedure.

teit van gehechtheidsrelaties beschreven in termen van strategieën voor het omgaan met angst in aanwezigheid van de ouder. De strategie die veilig gehechte kinderen toepassen komt overeen met de instinctieve reactie op angst: het kind licht de ouder in over zijn gevoelens en onderneemt actie om contact of nabijheid tot stand te brengen. In de Vreemde Situatie zijn deze kinderen dan ook al snel getroost wanneer de gehechtheidsfiguur hen bij terugkeer oppakt, waarna ze de exploratie van de omgeving hervatten. De ouders van deze kinderen zijn over het algemeen sensitief: ze zijn gevoelig voor de signalen van hun kind, en reageren prompt en adequaat.

Voor onveilig gehechte kinderen is de situatie ingewikkelder. Ook zij zoeken een uitweg voor de angst die voortkomt uit de situatie, maar zij hebben ook rekening te houden met de ouder. De strategie van onveilig-vermijdend gehechte kinderen is om de eigen aandacht af te wenden van hun angstgevoelens. De ouder wordt genegeerd of vermeden. Dit zou te verklaren zijn uit het feit dat de ouders van deze kinderen vaak afwijzend reageren op de signalen van hun kind. Onveilig-ambivalent gehechte kinderen doen er daarentegen alles aan om de volle aandacht van de ouder te krijgen. Bij het minste teken van onraad zoeken zij contact en zijn dan zeer moeilijk te troosten. Vaak gaat het contactzoeken gepaard met uitingen van boosheid gericht op de ouder. Deze ouders zijn vaak beperkt beschikbaar voor hun kinderen, of zijn hierin inconsistent (Cassidy & Berlin, 1994). In de meeste studies worden twee op de drie kinderen geclassificeerd als veilig gehecht. Van de onveilige groepen is de onveilig-vermijdend gehechte groep meestal de grootste.

Main en Solomon (1990) ontdekten echter een vierde groep kinderen. Het enige wat deze kinderen gemeen hadden was dat hun gedrag tijdens de Vreemde Situatie niet goed paste in één van de hierboven geschetste patronen. Sommige kinderen reageerden tegenstrijdig op de terugkeer van hun opvoeder in de Vreemde Situatie, bijvoorbeeld wanneer vermijden en contactzoeken werden gecombineerd in het met afgewend gelaat naderen van de opvoeder. Andere kinderen vertoonden bizarre gedragingen, zoals het bevriezen van alle beweging in aanwezigheid van de opvoeder. Weer andere kinderen vertoonden tekenen van angst voor de

opvoeder op het moment dat ze die terug zagen komen. Meestal was bij kinderen die dergelijk gedrag vertoonden wel te bepalen of ze voor het overige konden worden geclassificeerd als veilig, dan wel onveilig-vermijdend of onveilig-ambivalent. Main en Solomon (1990) noemden deze kinderen gedesorganiseerd (type D – van Disorganized) in hun gehechtheidsrelatie met de opvoeder. Deze kinderen vertonen geen consistente strategie als oplossing voor angst in de aanwezigheid van hun opvoeder. Uit fysiologische metingen (Hertsgaard, Gunnar, Erickson & Nachmias, 1995; Spangler & Grossmann, 1993) bleek dat gedesorganiseerde kinderen na de Vreemde Situatie langer last hadden van stress-symptomen.

Follow-up onderzoek door Main en Cassidy (1988) en door Wartner, Grossmann, Fremmer-Bombik en Suess (1994) bij deze kinderen toonde aan dat zij op zesjarige leeftijd ook als gedesorganiseerd geclassificeerd konden worden, zij het dat desorganisatie zich op deze leeftijd anders manifesteert. Desorganisatie op éénjarige leeftijd voorspelde in één onderzoek gedragsproblemen (Goldberg, Gotowiec & Simmons, 1995) en in een ander onderzoek cognitieve problemen (Jacobsen, Edelstein, & Hoffmann, 1994) jaren later. Ander onderzoek vond een relatie met vijandig gedrag en agressie jegens klasgenoten (Lyons-Ruth, Alpern & Repacholi, 1993). Ook Wartner en haar collega's (1994) vonden op grond van observaties tijdens schooltijd dat gedesorganiseerde vijfjarigen minder competent waren in hun sociale gedrag. Ook de mentale en motorische ontwikkeling van gedesorganiseerde kinderen blijkt negatief te worden beïnvloed (Lyons-Ruth, Repacholi, McCleod & Silva, 1991).

Men stelde al snel vast dat desorganisatie vaak voorkwam bij kinderen die mishandeld of verwaarloosd werden (zie Cicchetti & Toth, 1995). Ook depressie bij de opvoeder bleek een rol te kunnen spelen (Dawson, Grofer Klinger, Panagiotides, Spieker & Frey, 1992; DeMulder & Radke-Yarrow, 1991). Maar in normale opvoedingssituaties – waar men kan aannemen dat dit soort problemen slechts een kleine groep betreft – komt deze classificatie toch nog in gemiddeld 15 procent van de gevallen voor (Van IJzendoorn, 1995). Wat kan hier de verklaring van zijn?

Rond dezelfde tijd dat Main en Solomon gedesororganiseerde gehechtheid ontdekten, ontwikkelden George, Kaplan en Main (1985) het Adult Attachment Interview (AAI), een interview voor volwassenen waarin de proefpersoon wordt uitgenodigd te vertellen over ervaringen die te maken hebben met gehechtheid. Dit zijn zowel ervaringen in de jeugd als ervaringen tijdens de volwassenheid. Main en Goldwyn (in druk) ontwikkelden voor dit interview een codeersysteem. Zij ontdekten namelijk dat de manier waarop volwassenen vertellen over hun gehechtheidservaringen een goede voorspeller is van de kwaliteit van de gehechtheidsrelatie met hun kinderen. Ouders van veilig gehechte kinderen bijvoorbeeld hadden uiteenlopende ervaringen, van positief tot negatief, maar ze hadden gemeen dat ze objectief maar toch met gevoel erover konden praten en zich bewust waren dat die ervaringen hen wellicht zouden kunnen hebben beïnvloed (Van IJzendoorn & Bakermans-Kranenburg, 1994).

Aparte vragen uit het AAI gaan over mogelijke traumatische ervaringen en bij de meeste mensen zijn dat sterfgevallen. Uit de antwoorden van sommige mensen bleek dat het verlies nog niet te boven was gekomen. Dat kon blijken uit hoe ze over het verlies spraken, bijvoorbeeld wanneer iemand middenin een zin stilvalt, maar het kon ook blijken uit hun gedachtenwereld, bijvoorbeeld wanneer iemand in de tegenwoordige tijd spreekt over de overledene alsof die nog steeds in leven is. Het bleek dat ouders met dergelijke tekenen van mentale desorganisatie of desoriëntatie gedesororganiseerde kinderen hadden (Main & Hesse, 1990). Aan het codeersysteem voor het AAI werd daarom de classificatie voor onverwerkt verlies (U – van Unresolved) toegevoegd. Deze classificatie wordt ook toegepast op andere traumatische ervaringen. Uit onderzoek bleek al snel dat er een belangrijk onderscheid is tussen ouders die blijik geven van onverwerkt verlies en ouders die, ondanks het feit dat ook bij hen ingrijpende verliezen hebben plaatsgevonden, niet als onverwerkt worden geclassificeerd. Verlieservaringen op zichzelf bleken geen voorspeller te zijn van gedesororganiseerde gehechtheidsrelaties met kinderen. Onverwerkt verlies zoals vastgesteld met het AAI echter wel (Ainsworth & Eichberg, 1991). Dit ver-

band is later een aantal malen gerepliceerd (zie Van IJzendoorn, 1995).

Main en Hesse (1990) kwamen met de volgende hypothese om het verband tussen onverwerkt verlies en gedesororganiseerde gehechtheid te verklaren: onverwerkt verlies brengt ouders ertoe beangstigend, angstig en/of dissociatief gedrag (hierna kortweg: beangstigend gedrag) te vertonen tegenover hun kinderen. Hierdoor worden de kinderen geconfronteerd met een onoplosbare paradox: de veilige haven in geval van angst – namelijk de opvoeder – is voor het kind tegelijkertijd de bron van angst. Deze paradox zorgt ervoor dat deze kinderen geen coherente strategie kunnen ontwikkelen om met angst om te gaan bij deze ouder, en dat uit zich in gedesororganiseerde gehechtheid. Een verklaring voor dit beangstigend gedrag zou volgens Main en Hesse kunnen liggen in gedissocieerde gedachten en gevoelens. Hierbij baseerden zij zich op Liotti (1992). Liotti legde verband tussen onverwerkt verlies en dissociatie, waarbij dissociatie het proces is waarbij traumatische geheugenfragmenten gedeeltelijk ontoegankelijk raken. Dissociatie resulteert echter niet in een complete verdwijning van de herinneringen aan de traumatische gebeurtenis en iemand kan onwillekeurig zich de gedissocieerde geheugenfragmenten bewust worden. Het kan beangstigend zijn om in een soort flashback opnieuw de ingrijpende ervaringen en emoties gewaar te worden.

Om Main en Hesse's (1990) gedurfde hypothese te toetsen, is het noodzakelijk dat beangstigend gedrag kan worden vastgesteld. Aangezien onverwerkt verlies en gedesororganiseerde gehechtheid in ongeveer één op de vijf of zes gevallen voorkomen in normale, niet-klinische steekproeven, moet beangstigend gedrag ook in dergelijke gezinnen aan te treffen zijn. Het sterkste bewijs voor het belang van beangstigend gedrag zou bovendien zijn te leveren wanneer dit gedrag spontaan optreedt in de opvoedingssituatie zelf. Wij deden daarom een vooronderzoek naar de mogelijkheid om in een niet-klinische steekproef beangstigend gedrag vast te stellen door middel van de observatie van ongestructureerde moeder-kind interactie in de thuissituatie. Daarnaast zochten we naar eerste aanwijzingen ter ondersteuning van de theorie rondom onverwerkt verlies en gedesororganiseerde gehechtheid.

1 Methode

1.1 Steekproef en procedure

Twee gemeenten in de nabijheid van Leiden verleenden medewerking aan ons onderzoek. Op basis van het bevolkingsregister stuurden we aan 157 moeders die ongeveer 6 maanden daarvoor een kind hadden gekregen een brief waarin ze werden uitgenodigd om mee te doen aan het onderzoek. Van 138 moeders werd een antwoord verkregen, waarvan er 78 hun medewerking toezegden (respons ratio: 50%). Een selectie van 13 moeders vond plaats op basis van telefonisch en schriftelijk verkregen informatie. Acht moeders werden geselecteerd vanwege hoge scores op de BLAAQ-U (Berkeley-Leiden Adult Attachment Questionnaire-U gedeelte), een vragenlijst die een voorspelling geeft van de classificatie onverwerkt trauma op grond van het AAI (Main, Van IJzendoorn & Hesse, 1993). Vijf moeders werden geselecteerd vanwege lage scores op deze lijst. Alle moeders hadden verlieservaringen, hoewel er verschillen waren in het aantal en de nabijheid van de personen die waren overleden. Video-opnamen van de moeder-kind interactie vonden plaats als de kinderen 10 maanden oud waren. Deze duurden 3 tot 4 uur. De moeders werden geïnstrueerd om de normale, dagelijkse activiteiten te ondernemen, mits deze binnenshuis plaatsvonden. Rond de eerste verjaardag van de kinderen kwamen de moeders naar de universiteit voor het AAI. Toen de kinderen 14 maanden oud waren vond op het instituut de Vreemde Situatie plaats. Eén moeder uit de groep die hoog scoorde op de BLAAQ-U is verder niet meegenomen in dit verslag aangezien ook andere familieleden aanwezig waren tijdens de thuisopnamen.

Het opleidingsniveau van deze moeders was minimaal MAVO/MBO. De twee moeders die buitenshuis werkten, deden dat voor 28 en 32 uur per week. De gemiddelde leeftijd van de moeders was 30,6 jaar ($sd=2,33$) op het moment dat ze het kind kregen waarmee ze aan het onderzoek meededen. Voor drie moeders was dit kind de eerste, de andere negen moeders hadden tussen de één en de drie oudere kinderen. Er waren 3 jongens en 9 meisjes in de studie.

Om meer beangstigend gedrag te kunnen observeren, breidden we de groep van 12 uit

met vier moeder-kindparen uit een nog lopend onderzoek over hetzelfde thema. Ook deze moeders rapporteerden alle belangrijke verliezen, en scoorden hoog op de BLAAQ-U. Van deze paren zijn alleen vragenlijstgegevens en thuisobservaties beschikbaar. In één geval is nog wel de Vreemde Situatie afgenomen.

1.2 Instrumenten

1.2.1 Vragenlijsten

Toen de kinderen tussen de zes en acht maanden oud waren, werd bij alle moeders een vragenlijst afgenomen om vast te stellen welke verlieservaringen ze hadden meegemaakt. In de vragenlijst worden 22 mogelijke personen in de familie-, vrienden- en kennissenkring genoemd, waarbij kan worden aangegeven of een dergelijke persoon was overleden, en zo ja, of de proefpersoon onder één dak had geleefd met deze persoon en hoe het contact was geweest met hem of haar.

Tegelijk werd ook een Nederlandse vertaling van de IBQ (Rothbart, 1981), de baby gedragsvragenlijst (BGV), afgenomen. Aan de hand van 59 items werden de volgende vier temperamentskenmerken gemeten: (a) glimlachen en lachen; (b) activiteitsniveau; (c) angst; en (d) troostbaarheid. Deze vier schalen zijn gebleken valide en betrouwbaar te zijn en stabiel over een periode van drie tot negen maanden (Rothbart, 1986). De kenmerken glimlachen en lachen, activiteitsniveau en angst konden worden gecombineerd tot één dimensie, namelijk reactiviteit (Rothbart, 1986).

Rond de eerste verjaardag namen we bij de geselecteerde moeders nog enkele vragenlijsten af, waaronder de Nederlandse vertaling van de 'Befindlichkeitsskala' (Bf-S) van Elsenga (1988) voor het meten van momentane depressie, en de BLAAQ-D (Main, Van IJzendoorn & Hesse, 1993) voor het meten van trekdepressie. Voor gegevens over de validiteit en betrouwbaarheid van deze schalen verwijzen we naar de betreffende auteurs.

1.2.2 Beangstigend gedrag

Aan de hand van informele observaties van gedrag van ouders tijdens Vreemde Situatie procedures en op basis van theoretische overwegingen (Main & Hesse, 1990; Hesse & Main, in prep.; Liotti, 1992; Main & Morgan,

1996) ontwikkelden Main en Hesse (1992) een codeersysteem voor gedrag van ouders dat kan leiden tot desorganisatie bij hun kinderen. Dit systeem is verder ontwikkeld en verfijnd in de loop van de huidige studie. De typen gedrag die worden onderscheiden zijn: a. beangstigend gedrag; b. angstig gedrag; c. andere indicaties dat een ouder in een enigszins dissociatieve toestand is geraakt. De onderscheiden angstige en dissociatieve gedragingen worden verondersteld beangstigend te zijn voor een baby wanneer ze vertoond worden door de moeder. De angstige gedragingen die in het systeem worden onderscheiden hebben voor het kind geen observeerbare oorzaak, en dat veroorzaakt indirect angst. Dissociatieve gedragingen kunnen ook beangstigend zijn, wanneer de moeder tijdelijk een andere en potentieel beangstigende werkelijkheid lijkt te beleven en daardoor uit haar gewone doen is.

Naast beangstigend, angstig en dissociatief gedrag onderscheiden we in het systeem ook onderdanig gedrag, geseksualiseerd gedrag, en niet direct beangstigend mogelijk dissociatief gedrag (bv. een moeder die dromerig met poppen speelt). Dat de gedragingen in deze laatste categorieën aparte scores krijgen is omdat ze niet direct angst hoeven op te roepen bij het kind. Ze zijn wel belangrijk, omdat ze mogelijk op andere manieren te maken kunnen hebben met desorganisatie.

Elk gedrag dat in één van de genoemde soorten beangstigend gedrag valt, krijgt een score van 1 tot 9 die de zwaarte van het gedrag aangeeft. Het codeersysteem geeft hiervoor een aantal aanwijzingen. Bijvoorbeeld: gedrag waarbij de moeder het kind aanvalt door het kind agressief beet te pakken en met haar gelaat en haar stem te bedreigen wordt op zichzelf al zwaar gerekend, maar wordt nog wat zwaarder gerekend als dat gebeurt wanneer het kind huult. Op basis van alle apart beoordeelde gedragingen komt men tot een eindscore, ook op een schaal van 1 tot 9. Hierbij wordt de hoogste score die voor een gedraging is toegekend als uitgangspunt genomen. Al naar gelang dergelijke ernstige gedragingen vaker voorkomen kan worden besloten de eindscore op te hogen. Eindscores hoger dan 5 leiden altijd tot de classificatie Fr (van Frightening, frightened, and/or dissociated). Bij een score van precies 5 moet de codeur zelf beslissen of

de moeder in de beangstigende of in de niet-beangstigende groep valt.

Een kwalitatieve screening van het materiaal werd ondernomen om te bepalen welke fragmenten voor codering in aanmerking kwamen. Elke opname werd gescreend voor beangstigend gedrag door twee personen, die beide vertrouwd waren met het codeersysteem voor beangstigend gedrag. Gedragingen werden geselecteerd om gecodeerd te worden wanneer ze leken op de gedragsbeschrijvingen in het codeersysteem, of wanneer het gedrag leek te beantwoorden aan de definitie van beangstigend gedrag. Verschillende nieuwe vormen van beangstigend gedrag kwamen hierdoor aan het licht. Hierdoor moest echter wel een aantal gevallen opnieuw worden gescreend.

De scores werden in consensus toegekend door de vier auteurs, Main en Hesse. Hier worden alleen de classificaties gerapporteerd. Er is aan dit onderzoek geen bepaling van de interbeoordelaarbetrouwbaarheid voorafgegaan. Twee van de huidige codeurs bereikten in een andere studie een betrouwbaarheid van .83 (Schuengel, Van IJzendoorn, Bakermans-Kranenburg & Blom, 1996). De codeurs waren blind voor de classificatie van de moeders als onverwerkt ten aanzien van een trauma of verlies, en voor de classificatie van de kinderen als gedesorganiseerd.

1.2.3 GehechtheidsBiografisch Interview

Zoals hierboven al is beschreven ontwikkelden George, Kaplan en Main (1985) een semi-structureerd interview, het Adult Attachment Interview (gehechtheidsbiografisch interview; AAI), gericht op de representatie van gehechtheidservaringen. Aan de hand van het codeersysteem van Main en Goldwyn (in druk) worden de verschillende manieren om over die ervaringen te spreken onderscheiden, om zo te komen tot een classificatie. Het coderen vindt plaats op grond van een letterlijke transcriptie van het interview. Uit het transcript wordt alle informatie weggelaten die mogelijk de geïnterviewde zou kunnen identificeren. Voor dit onderzoek zijn alleen de delen van het interview getranscribeerd die te maken hadden met traumatische ervaringen en verlies. Op deze transcripten werd het codeersysteem voor onverwerkt trauma en onverwerkt verlies toegepast. Dit codeersysteem onderscheidt de ver-

schillende vormen waarin onverwerkt verlies naar buiten kan komen, zoals in versprekingen of in de manier waarop de verlieservaring wordt beschreven of verklaard. De sterkte van de gevonden aanwijzingen wordt uitgedrukt op een schaal van 1 tot 9. Scores hoger dan 5 leiden altijd tot de classificatie onverwerkt verlies. Bij een score van precies vijf moet de codeur de keuze maken.

Bakermans-Kranenburg (1993) heeft aangetoond dat de codering van het AAI betrouwbaar kan geschieden, dat de classificatie stabiliteit vertoont wanneer het enkele maanden later opnieuw wordt afgenomen, en dat de uitkomst onafhankelijk is van de persoon van de interviewer. Daarnaast bleek het AAI te beschikken over discriminante validiteit: de uitkomst was onafhankelijk van intelligentie, sociale wenselijkheid en autobiografisch geheugen. De codering van onverwerkt verlies en onverwerkt trauma in dit onderzoek werd gedaan door de derde auteur. Zij nam deel aan twee trainingen in het codersysteem, gegeven door M. Main en E. Hesse. Zij rapporteerde een hoge intercodeurbetrouwbaarheid voor de classificatie van onverwerkt verlies: 88% overeenstemming over 16 gevallen (Bakermans-Kranenburg, 1993).

1.2.4 *Vreemde Situatie procedure*

De Vreemde Situatie procedure van Ainsworth en haar collega's (1978) – twee scheidingen en herenigingen van opvoeder en kind in een onbekende omgeving, alsmede een ontmoeting met een onbekende volwassene – is hierboven al enigszins beschreven. De kinderen kunnen bijna altijd worden geclassificeerd in één der drie groepen: veilig (B); onveilig-vermijnd (A); onveilig-ambivalent (C). Als het gedrag van het kind niet valt binnen één van deze patronen, kan het aan de categorie 'onclassificeerbaar' (CC – van Cannot Classify) worden toegewezen. Altijd wordt het kind geclassificeerd als gedesorganiseerd (D) als zijn score op de schaal van Main en Solomon (1990) voor dergelijk gedrag hoger is dan vijf. Bij een score van vijf moet de codeur een beslissing nemen. Ook al wordt de D-classificatie toegekend, daarnaast wordt het kind ook altijd ingedeeld in één van de groepen A, B, C of CC.

De classificatie van de kinderen en de beoordeling van gedesorganiseerde gehech-

heid werden verricht door codeurs die blind waren voor alle andere informatie over de moeder-kindparen. Alle codeurs waren getraind door Main en Hesse. Intercodeur overeenstemming was adequaat, met overeenstemming voor de D-classificatie tussen 83% en 96%.

2 Resultaten

2.1 *Verbanden tussen de variabelen*

In Tabel 1 zijn de resultaten samengevat van de verschillende metingen. Op basis van deze gegevens hebben we de verbanden uit het model van Main en Hesse (1990) statistisch getoetst. De toetsen die hieronder zijn beschreven zijn gebaseerd op verschillende aantallen personen vanwege het ontbreken van sommige gegevens. Vanwege de kleine aantallen zijn alle verbanden non-parametrisch getoetst. Voor het bepalen van de significantie van verbanden tussen twee dichotome variabelen is voor kleine steekproeven de éenzijdige Fisher's Exact Test toepasselijk (Verbeek & Kroonenberg, 1993). Aangezien het hier gaat om een verkennende studie met een beperkte omvang naar een aantal nauw omschreven hypothesen, hanteren we een alpha van 0,10 om daarmee het gevaar van type II fouten enigszins te verminderen. Om een indruk te krijgen van het belang van de bevindingen zijn effectgrootte indicatoren belangrijker dan p -waarden. Wij rapporteren daarom ook de ϕ -waarde voor het verband in kruistabellen, aangezien deze op dezelfde manier is te interpreteren als r . Definitieve conclusies kunnen vaak pas worden getrokken op grond van meta-analyse van meerdere studies van hetzelfde verband en daarbij tellen alleen effectgrootten, en geen p -waarden (Rosnow & Rosenthal, 1992).

In 8 van de 11 gevallen (73%) correspondeerde onverwerkt verlies bij de moeders met gedesorganiseerde gehechtheid bij de kinderen ($\phi=0,45$; Fisher's Exact Test: $p=0,20$). In 9 van de 12 gevallen (75%) correspondeerde onverwerkt verlies met beangstigend gedrag bij de moeders ($\phi=0,60$; Fisher's Exact Test: $p=0,07$). In slechts 6 van de 12 gevallen (50%) correspondeerde beangstigend oudergedrag met gedesorganiseerde gehechtheid bij de kinderen ($\phi=0,13$; Fisher's Exact Test: $p=0,59$).

Tabel 1
Uitkomsten per moeder-kind dyade

	AAI ^a be- angstigend gedrag		SSP ^b		Fr-gedrag ^c					reacti- viteit		leeftijd bij verlies/ verlies/ leeftijd	
	U	gedrag	D	ABC	I	II	III	IV	V	baby	sekse	bevalling	leeftijd
1	7,5	wel	6	B	+					-1,85	meisje	30	grootvader/27
2	7	wel	1	B	+			-		-1,81	meisje	30	vader/20
3	6	wel	5	B	+					-1,32	meisje	25	vader van vriendin/14-15
4	5	wel	7	A	+		-	-		-2,11	jongen	28	beste vriend/27
5	5,5	wel	5,5	B		+				-1,37	meisje	31	vader/22
6	1	niet	1	B	-					-2,04	jongen	31	vader schoonzus/--
7	4,5	niet	--	--						-1,71	jongen	30	nicht van moeder/26-27
8	2	niet	6	B						-2,21	meisje	32	vader/32
9	2	niet	1	A	-					-2,14	meisje	28	vriend/18-kind van zus/--
10	1	wel	1	B	-	+	-	-		-1,62	meisje	34	grootouders/12-16-20-31
11	4	wel	2	B	+					-1,60	meisje	31	vriend/18
12	1	wel	1	B	+					-2,40	meisje	30	grootouders/5-6
13	--	wel	--		+					-2,09	jongen	--	--
14	--	wel	--			+		-		-1,89	meisje	--	--
15	--	wel	--		+		-			-1,70	meisje	--	--
16	--	niet	1	B					+	-2,36	meisje	33	goede vriend/19-vader/34

Opmerking. Ontbrekende gegevens worden aangegeven met twee streepjes (--). Dyades 13 tot en met 16 zijn afkomstig uit een andere steekproef.

^a Adult Attachment Interview. Onderstreepte scores geven aan dat de moeder geïnterviewd werd als onverwerkt ten aanzien van een verlies (U).

^b Strange Situation Procedure. Onderstreepte scores geven aan dat het kind geïnterviewd werd als gedesorganiseerd (D). Kinderen konden daarnaast geïnterviewd worden als veilig (B), onveilig-vermijdend (A), en onveilig-afwerend gehecht (C).

^c Beangstigend gedrag. De categorieën voor Fr-gedrag zijn: I. beangstigend gedrag; II. angstig gedrag; III. gedissocieerd gedrag. Daarnaast worden onderscheiden: IV. timide/ onderdanig gedrag; V. geseksualiseerd gedrag. Een plus-teken geeft aan dat het gedrag binnen een bepaalde categorie voldoende was voor een Fr-score van vijf of hoger. Een minus-teken geeft aan dat het gedrag in deze categorie voldoende was voor een score tussen de één en de vijf. Een spatie geeft aan dat dit gedrag niet werd waargenomen.

In dit onderzoek hanteerden we voor het ouder-gedrag een tweedeling in beangstigend en niet beangstigend. Vervolgonderzoek zou daarnaast gebruik dienen te maken van schaal-scores, aangezien continue variabelen vaak robustere verbanden opleveren. Dit blijkt al als we voor onverwerkt verlies en voor gedesorganiseerde gehechtheid gebruik maken van schaal-scores. De Spearman rangorde correlatie tussen de schaal-scores voor onverwerkt verlies en desorganisatie was statistisch significant ($r_s=0,51$; $p=0,05$). Het verschil tussen beangstigende en niet beangstigende moeders in score op de schaal voor onverwerkt verlies was ook significant (Mann-Whitney $U=8,0$; $p=0,09$). Het verschil in de scores van hun kinderen op de schaal voor desorganisatie/desoriëntatie werd echter niet significant (Mann-Whitney $U=11,0$; $p=0,18$).

We hebben getest of de gevonden effecten samenhangen met verschillen in temperament van de kinderen. We kunnen de Kendall Tau-b

gebruiken als maat voor de samenhang tussen de verschillende variabelen in dit onderzoek (Verbeek & Kroonenberg, 1993). Dit stelt ons in staat om de Kendall partiële rangorde correlatie (Siegel, 1956) te berekenen. Controle voor verschillen in reactiviteit ($n=11$) leverde geen verandering van betekenis op voor de verbanden tussen U, Fr en D.

Op dezelfde manier hebben we getest of depressie een rol kon spelen. Voor geen van de verbanden tussen U, Fr en D was er een daling van betekenis wanneer werd gecontroleerd voor de scores op de twee depressie schalen (Bf-S en BLAAQ-D; $n=9$).

2.2 Beangstigend gedrag

We kozen voor een oververtegenwoordiging van proefpersonen met hoge scores op de BLAAQ-U in de verwachting dat beangstigend gedrag slechts sporadisch zou komen in normale steekproeven. Dit bleek niet het geval. Van de onderzoeksgroep van 12 moeders wer-

den 8 moeders geclassificeerd als Fr. Deze 8 waren afkomstig uit een steekproef van 78 moeders, hetgeen betekent dat van deze groep van 78 minstens 10% van de moeders beangstigend, angstig en/of dissociatief gedrag vertoont.

Tabel 1 geeft aan wat voor gedragingen werden waargenomen bij de moeders die geclassificeerd werden als Fr. Opvallend is dat vooral beangstigend gedrag de doorslag gaf: in 8 van de 12 gevallen. Drie moeders werden geclassificeerd op grond van angstig gedrag. Eén moeder vertoonde alleen geseksualiseerd gedrag (zie hieronder).

Door de screening werd een aantal beangstigende gedragingen ontdekt dat in het systeem moest worden opgenomen, aangezien ze niet omschreven waren, maar wel binnen de definitie passen. Snelle bewegingen naar het gezicht van het kind toe stonden omschreven, maar een langzame beweging met de hand naar het gezicht toe, waarna de hand over de ogen werd gedrukt moest ook als beangstigend worden beschouwd. Een moeder stond voortdurend in een houding die niet kon worden omschreven als aanvallend, maar eerder als dreigend. Twee moeders sperden hun ogen zeer wijd open tegenover het kind, in een uitdrukking van afkeer of angst.

Een moeder vertoonde een patroon van angstig gedrag, waarbij de individuele gedragingen volgens het systeem geen hoge score konden opleveren. Vanwege het herhaaldelijk weglopen uit de buurt van het kind en het om het kind heen bewegen met de grootst mogelijke afstand kreeg deze moeder echter een verhoogde score. Dezelfde moeder gaf verbaal aan bij bepaald gedrag van haar kind dat ze er bang van werd. Deze mogelijkheid moest ook in het

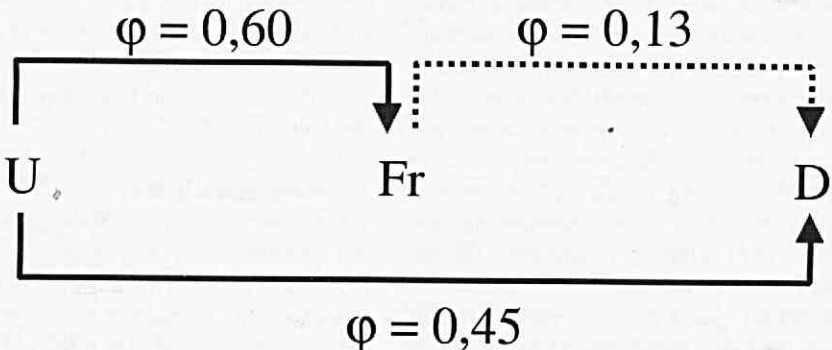
systeem worden opgenomen. Een andere moeder vertoonde overdreven schrikreacties wanneer haar kind viel.

Andere nieuwe gedragingen werden gevonden die kunnen worden geïnterpreteerd als uitingen van een dissociatieve toestand, zoals een onverklaarbare combinatie van knikken, slikken en kijken naar het kind, of opwaarts rollende ogen.

Een moeder kuste en beroerde met haar lippen langdurig het gezicht, de mond en het lichaam van haar kind, met een dromerige gelaatsuitdrukking. We concludeerden dat dit erotiserend gedrag conflicteerde met de expressie van gehechtheid door het kind en met verzorging en opvoeding door de ouder. Dit gedrag werd geschaard onder geseksualiseerd gedrag.

3 Discussie

De belangrijkste vraag, namelijk of het mogelijk is om door middel van naturalistische observatie beangstigend gedrag vast te stellen bij niet-klinische moeders, kon bevestigend worden beantwoord. Verder vonden we met onze bescheiden steekproef nu al enkele veelbelovende aanwijzingen voor het model voor de intergenerationele overdracht van onverwerkt verlies. In Figuur 1 zijn de belangrijkste resultaten samengevat. De hypothese dat onverwerkt verlies samengaat met beangstigend gedrag vindt enige ondersteuning. We vonden een sterk verband, zij het dat we vanwege de kleine aantallen en het lage significantieniveau nog moeten afwachten of de resultaten in ander onderzoek gerepliceerd worden. We hebben echter geen bevestiging dat moeders die derge-



Figuur 1. Het intergenerationele model en de in deze studie gevonden effectgrootten

lijk gedrag vertonen ook vaker kinderen hebben die gedesorganiseerd zijn. De kinderen lijken wel een indirecte invloed te ondervinden. Uit eerder onderzoek bleek al de samenhang tussen onverwerkt verlies en desorganisatie. Een meta-analyse leverde een gecombineerde effectgrootte r op van 0,31 (Van IJzendoorn, 1995). De in dit onderzoek gevonden effectgrootte van 0,45 mag dan vanwege de kleine aantallen niet statistisch significant zijn, deze wijkt niet in ongunstige zin af van het algemeen gevonden effect.

De rol van temperamentskenmerken van de kinderen bleek bescheiden. Bedacht moet worden dat de meting van temperamentskenmerken van de kinderen berust op rapportage door de moeders zelf. In onderzoek naar kinderen van depressieve ouders is bijvoorbeeld gebleken dat ouders met affectieve stoornissen de problemen bij hun kinderen overschatten (Fergusson, Lynskey & Horwood, 1993). In dit onderzoek leek depressie zoals gemeten met de Nederlandse versie van de Bf-S en de BLAAQ-D bij de moeder geen rol te spelen bij de verbanden tussen onverwerkt verlies, beangstigend gedrag en desorganisatie.

In het onderzoek naar de overdracht van de representatie van gehechtheid bij ouders op de kwaliteit van de gehechtheidsrelatie met het kind is de gangbare verklaring voor deze overdracht dat de representatie van gehechtheid de sensitiviteit van de opvoeders beïnvloedt. Deze verschillen in sensitiviteit zouden de verschillen in de kwaliteit van de gehechtheidsrelaties van de kinderen verklaren. De Wolff (1996) toont aan in haar meta-analyse dat er een verband bestaat tussen sensitiviteit van de ouder en kwaliteit van gehechtheid van het kind, maar dat dit verband niet groot genoeg is om de sterke intergenerationele overdracht van gehechtheid te verklaren. Daarom wordt er gesproken van de 'transmission gap' (Van IJzendoorn, 1995). Een dergelijke transmission gap lijkt ook in dit onderzoek aan de orde. Ook hier is de 'zwakste schakel' in het model de beïnvloeding van de kwaliteit van de gehechtheidsrelatie door het gedrag van de moeder.

Mogelijk is er sprake van drempel-effecten. Desorganisatie wordt in de Vreemde Situatie vastgesteld als blijkt dat het kind geen uitweg heeft voor de stress die is opgeroepen door de situatie. Dit zou komen doordat de moeder zelf

een bron van angst is voor het kind (Main & Hesse, 1990). Wanneer echter de ervaringen met de moeder mild beangstigend waren, kunnen sommige kinderen wellicht hun angst voor hun moeder opzij zetten en een veilig patroon van gehechtheidsgedrag vertonen. Het was bijvoorbeeld opvallend dat twee kinderen tijdens de thuisobservaties, waar de moeders beangstigend gedrag vertoonden, wél gedesorganiseerd gedrag vertoonden, maar tijdens de Vreemde Situatie niet. Dit zou er echter ook op kunnen duiden dat de Vreemde Situatie voor sommige kinderen niet de optimale context is om desorganisatie vast te stellen, wanneer deze kinderen bij andere bronnen van stress en in een andere omgeving wel als gedesorganiseerd zouden worden getypeerd.

Aan de codering van beangstigend gedrag ging nu een screening van de relevante fragmenten vooraf. Het bleek dat een screening ook noodzakelijk is om te voorkomen dat de codeurs 'besmet' worden met informatie over desorganisatie bij de kinderen. Twee kinderen reageerden op beangstigend gedrag met gedesorganiseerd gedrag, ernstig genoeg om hen als gedesorganiseerd te kunnen classificeren – zou het gedrag zijn vertoond tijdens de Vreemde Situatie.

Een belangrijke vraag in dit onderzoek was of de methode van naturalistische observatie geschikt is voor het valide vaststellen van beangstigend gedrag. Een bedreiging voor de validiteit zou kunnen zijn dat moeders met onverwerkt verlies anders reageren op de aanwezigheid van de videocamera dan moeders zonder onverwerkt verlies. Personen met onverwerkt verlies zouden in het algemeen angstiger kunnen zijn en daardoor meer angst hebben voor camera's. Om de situatie voor de moeders minder bedreigend te maken richtten de onderzoekers de camera, wanneer moeder en kind niet samen in beeld konden worden gebracht, voornamelijk op het kind. Moeders met angst voor de camera zouden er daardoor toe hebben kunnen komen om een zo groot mogelijke afstand te bewaren tot het kind, zodat ze zelf zo min mogelijk gefilmd zouden worden. Dit zou dan ten onrechte kunnen worden geïnterpreteerd als angstig gedrag.

In dit onderzoek vertoonden drie moeders voornamelijk angstig gedrag in tegenstelling

tot het direct beangstigende gedrag. Bij één moeder werd geen AAI afgenomen, maar ze scoorde hoog op de BLAAQ-U. Een andere moeder werd inderdaad geclassificeerd als onverwerkt op het AAI, maar de derde moeder werd niet geclassificeerd als onverwerkt. Alleen deze moeder vertoonde gedrag zoals hierboven beschreven: afstand houden van het kind, haar lichaam zo ver mogelijk van het kind vandaan houden, zich terugtrekken wanneer het kind naderde. De interpretatie van haar gedrag was dan ook in eerste instantie dat er sprake was van angst voor de camera. Maar het gedrag verminderde niet naarmate de moeder langer aan de camera werd blootgesteld, en ze vertoonde het gedrag op verschillende manieren, en telkens als een reactie op de toenadering door het kind en niet louter de nabijheid van het kind. Eén van de conclusies die hieruit kan worden getrokken is dat de relatief lange duur van de thuisopnamen (twee bezoeken van elk anderhalf à twee uur) van cruciaal belang is. Niet alleen stelt dit de proefpersonen in staat om aan de aanwezigheid van de camera te wennen, het geeft ook meer gelegenheid om een bepaald type gedrag in verschillende verschijningsvormen te observeren.

In aanmerking nemend dat beangstigend gedrag vaak voorkwam in de onderzoeksgroep, kunnen we concluderen dat de aanwezigheid van de camera en de camerapersoon moeders niet weerhoudt dergelijk gedrag te vertonen. Dit spoort met de veronderstelling van Main en Hesse (1990) dat het hier gaat om gedrag dat aan de bewuste controle ontsnapt. Het lijkt ook waarschijnlijk dat de dagelijkse omgeving van het eigen huis hiervoor een betere context is dan de laboratoriumsituatie (Hesse & Main, in prep.). Zo heeft Liotti (1992) een geval beschreven van een moeder met onverwerkt verlies die extreem reageerde wanneer haar kind in de buurt kwam van gebroken glas. Dit illustreert dat juist een grote variëteit van prikkels en voorwerpen gewenst is als het erom gaat vast te stellen of iemand beangstigend gedrag vertoont tegenover zijn of haar kind. De laboratoriumsituatie is hiervoor te beperkt.

Deze studie leert dat beangstigend gedrag van opvoeders een belangrijke nieuwe gedragscategorie is, die mogelijk meer licht kan werpen op de problematische start van sommige

kinderen. Naturalistische observatie lijkt geschikt te zijn voor het betrouwbaar vaststellen van dit gedrag. Inmiddels is in een grotere studie bij 85 moeders en kinderen, gebruikmakend van dezelfde methode, vastgesteld dat beangstigend gedrag inderdaad gedesorganiseerde gehechtheid voorspelt (Schuengel, Van IJzendoorn, Bakermans-Kranenburg & Blom, 1997). Tevens werd een belangrijk interactie-effect gevonden tussen onverwerkt verlies en veilige gehechtheid van de moeder zelf: veilige moeders met onverwerkt verlies scoorden laag op de schaal voor beangstigend gedrag, onveilige moeders met onverwerkt verlies scoorden hoog. Moeders zonder onverwerkt verlies hadden scores hier tussenin, waarbij er geen verschillen waren tussen veilige en onveilige moeders. De uitkomsten van dit zowel als het huidige onderzoek wijzen erop dat beangstigend gedrag verweven is met onverwerkte verliezen in het verleden van de moeders.

Literatuurlijst

- Ainsworth, M.D.S. (1969). Object relations, dependency, and attachment: A theoretical review of the infant-mother relationship. *Child Development, 40*, 969-1025.
- Ainsworth, M.D.S., Blehar, M.C., Waters, E., & Wall, S. (1978). *Patterns of attachment*. Hillsdale, NJ: Lawrence Erlbaum.
- Ainsworth, M.D.S., & Eichberg, C. (1991). Effects on infant-mother attachment of mother's unresolved loss of an attachment figure, or other traumatic experience. In C.M. Parkes, J. Stevenson-Hinde & P. Marris (Eds.), *Attachment across the life cycle* (pp. 160-183). London: Routledge Tavistock.
- Bakermans-Kranenburg, M.J. (1993). *Het gehechtheidsbiografisch interview: Betrouwbaarheid en discriminante validiteit*. Ongepubliceerd proefschrift, Rijksuniversiteit Leiden.
- Bowlby, J. (1989). *Attachment* (2nd ed.). London: Penguin.
- Cassidy, J., & Berlin, L.J. (1994). The insecure/ambivalent pattern of attachment: Theory and research. *Child Development, 65*, 971-991.
- Cicchetti, D., & Toth, S.L. (1995). A developmental psychopathology perspective on child abuse and neglect. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 541-565.

- Dawson, G., Grofer Klinger, L., Panagiotides, H., Spieker, S., & Frey, K. (1992). Infants of mothers with depressive symptoms: Electroencephalographic and behavioral findings related to attachment status. *Development and Psychopathology, 4*, 67-80.
- DeMulder, E.K., & Radke-Yarrow, M. (1991). Attachment with affectively ill and well mothers: Concurrent behavioral correlates. *Development and Psychopathology, 3*, 227-242.
- Elsenga, S. (1988). Een zelfbeoordelingschaal voor depressie. *Tijdschrift voor Psychiatrie, 30*, 114-126.
- Fergusson, D.M., Lynskey, M.T., & Horwood, L.J. (1993). The effect of maternal depression on maternal ratings of child behavior. *Journal of Abnormal Child Psychology, 21*, 245-269.
- George, C., Kaplan, N., & Main, M. (1985). *Adult attachment interview*. Unpublished manuscript.
- Goldberg, S., Gotowiec, A., & Simmons, R.J. (1995). Infant-mother attachment and behavior problems in healthy and chronically ill preschoolers. *Development and Psychopathology, 7*, 267-282.
- Hertsgaard, L., Gunnar, M., Erickson, M.F., & Nachmias, M. (1995). Adrenocortical responses to the strange situation in infants with disorganized/disoriented attachment relationships. *Child Development, 66*, 1100-1106.
- Hesse, E., & Main, M. (in prep.). *Frightened behavior in traumatized but non-maltreating parents: Potential risk factor with respect to anxiety*. University of California at Berkeley.
- Ijzendoorn, M.H. van (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychological Bulletin, 117*, 387-403.
- Ijzendoorn, M.H. van, & Bakermans-Kranenburg, M. J. (1994). Intergenerationele overdracht van gehechtheid: De stand van zaken in het psychometrisch, klinisch en pedagogisch onderzoek. *Kind en Adolescent, 15*, 1-24.
- Jacobsen, T., Edelstein, W., & Hofmann, V. (1994). A longitudinal study of the relation between representation of attachment in childhood and cognitive functioning in childhood and adolescence. *Developmental Psychology, 30*, 112-124.
- Liotti, G. (1992). Disorganized disoriented attachment in the etiology of the dissociative disorders. *Dissociation, 5*, 196-204.
- Lyons-Ruth, K., Alpern, L., & Repacholi, B. (1993). Disorganized infant attachment classification and maternal psychosocial problems as predictors of hostile-aggressive behavior in the preschool classroom. *Child Development, 64*, 572-585.
- Lyons-Ruth, K., Repacholi, B., McCleod, S., & Silva, E. (1991). Disorganized attachment behavior in infancy: Short-term stability, maternal and infant correlates, and risk-related subtypes. *Development and Psychopathology, 3*, 377-396.
- Main, M. (1990). Cross-cultural studies of attachment organization: Recent studies, changing methodologies, and the concept of conditional strategies. *Human Development, 33*, 48-61.
- Main, M., & Cassidy, J. (1988). Categories of response to reunion with the parent at age 6: Predictable from infant attachment classifications and stable over a 1-month period. *Developmental Psychology, 24*, 415-426.
- Main, M., & Goldwyn, R. (in druk). *Adult attachment scoring and classification systems*. Unpublished manuscript, University of California at Berkeley.
- Main, M., & Hesse, E. (1990). Parents' unresolved traumatic experiences are related to infant disorganized attachment status: Is frightened and/or frightening parental behavior the linking mechanism? In M.T. Greenberg, D. Cicchetti & E. M. Cummings (Eds.), *Attachment in the preschool years: Theory, research, and intervention* (pp. 161-182). Chicago and London: The University of Chicago Press.
- Main, M., & Hesse, E. (1992). *Frightening, frightened, timid, dissociated or disorganized behavior on the part of the parent: A coding system for use with videotaped parent-infant interactions in the home or laboratory setting*. Unpublished manuscript, University of California at Berkeley.
- Main, M., & Morgan, H. (1996). Disorganization and disorientation in infant strange situation behavior: Phenotypic resemblance to dissociative states? In L. Michelson & W. Ray (Eds.), *Handbook of dissociation: Theoretical, empirical and clinical perspectives* (pp. 107-138). New York: Plenum Publishing.
- Main, M., & Solomon, J. (1990). Procedures for identifying infants as disorganized/disoriented during the Ainsworth Strange Situation. In M.T. Greenberg, D. Cicchetti & E. M. Cummings (Eds.), *Attachment in the preschool years: Theory, research, and intervention* (pp. 121-160). Chicago and London: The University of Chicago Press.

- Main, M., IJzendoorn, M.H. van, & Hesse, E. (1993, March). *Adolescent attachment organization: Findings from the BLAAQ self-report inventory, and relations to absorption and dissociation*. Paper presented at the 60th meeting of the Society for Research in Child Development, New Orleans, LA.
- Rosnow, R.J., & Rosenthal, R. (1992). Statistical procedures and the justification of knowledge in psychological science. In A.E. Kazdin (Ed.), *Methodological issues & strategies in clinical research* (pp. 295-314). Washington, DC: APA.
- Rothbart, M.K. (1981). Measurement of temperament in infancy. *Child Development*, 52, 569-578.
- Rothbart, M.K. (1986). Longitudinal observation of infant temperament. *Developmental Psychology*, 22, 356-365.
- Schuengel, C., IJzendoorn, M.H. van, Bakermans-Kranenburg, M.J., & Blom, M. (1996). *Observation procedures frightening, frightened, and/or dissociated maternal behavior*. Ongepubliceerd manuscript, Rijksuniversiteit Leiden.
- Schuengel, C., IJzendoorn, M.H. van, Bakermans-Kranenburg, M.J., & Blom, M. (1997, April). Frightening, frightened and/or dissociated behavior, unresolved loss and infant disorganization. In D. B. Jacobvitz (Chair), *Caregiving correlates and longitudinal outcomes of disorganized attachment in infants*. Symposium conducted at the Biennial Meeting of the Society for Research in Child Development, Washington, DC.
- Siegel, S. (1956). *Nonparametric statistics: For the behavioral sciences*. New York: McGraw-Hill.
- Spangler, G., & Grossmann, K.E. (1993). Biobehavioral organization in securely and insecurely attached infants. *Child Development*, 64, 1439-1450.
- Verbeek, A., & Kroonenberg, P.M. (1993). *Fisher 3.10: Testing independence in rxc tables*. Groningen: iec ProGamma.
- Wartner, U.G., Grossmann, K., Fremmer-Bombik, E., & Suess, G. (1994). Attachment patterns at age six in South Germany: Predictability from infancy and implications for preschool behavior. *Child Development*, 65, 1014-1027.
- Wolff, M.S. de (1996). *Maternal interactive behavior and infant attachment: A meta-analysis*. Ongepubliceerd proefschrift, Rijksuniversiteit Leiden.

Auteurs

C. Schuengel is als Universitair Docent verbonden aan de Vakgroep Algemene Pedagogiek van de Rijksuniversiteit Leiden.

M.H. van IJzendoorn is als Hoogleraar verbonden aan de Vakgroep Algemene Pedagogiek van de Rijksuniversiteit Leiden.

M.J. Bakermans-Kranenburg is als Universitair Docent verbonden aan de Vakgroep Algemene Pedagogiek van de Rijksuniversiteit Leiden.

M. Blom is studente aan de Vakgroep Algemene Pedagogiek van de Rijksuniversiteit Leiden.

Adres: Vakgroep Algemene Pedagogiek, RU Leiden, Wassenaarseweg 52, Postbus 9555, 2300 RB Leiden

Abstract

Unresolved loss, disorganized attachment, and frightening parental behavior: A pilot-study

C. Schuengel, M.H. van IJzendoorn, M.J. Bakermans-Kranenburg & M. Blom. Pedagogische Studiën, 1997, 74, 355-366.

Unresolved loss in parents appears to be associated with disorganized infant attachment (Van IJzendoorn, 1995). One explanation presumes that these parents show frightening behavior (Main & Hesse, 1990). We observed the home behavior of 16 mothers with their infants. Unresolved loss was determined with the Adult Attachment Interview (George, Kaplan & Main, 1985). Disorganized attachment was determined in the Strange Situation. We found frightening behavior appearing during the home observations, and the behavior was associated with unresolved loss. We did not find an association between frightening behavior and disorganized attachment.

De Onderwijsresearchdagen

1997 in Leuven,

21-23 mei 1997

Inleiding (J.N. Streumer, TO, Universiteit Twente, Enschede)

Van 21 tot en met 23 mei vond onder auspiciën van de Vereniging voor Onderwijsresearch (VOR) en het Vlaams Forum voor Onderwijs-onderzoek (VFO) de vierentwintigste aflevering plaats van de Onderwijsresearchdagen (ORD). De VFO is de Vlaamse evenknie van de VOR en bestaat pas sinds 1995.

Het centrale thema van deze Onderwijsresearchdagen was "Leren en onderwijzen van leervaardigheden doorheen het curriculum". Alleen deze titel al verraaft, niet inhoudelijk maar taalkundig, dat de Onderwijsresearchdagen deze keer in Vlaanderen plaats vonden; om precies te zijn in Leuven.

In zijn inleiding verantwoordde Erik de Corte, voorzitter van de organisatiecommissie, de keuze van dit centrale thema: op dit moment geniet het thema op internationaal niveau ruime belangstelling in kringen van onderzoekers, het beleid en de praktijk van het onderwijs. Het belang van cognitieve en meta-cognitieve vaardigheden als hefboven van leren wordt in toenemende mate beklemtoond, aldus De Corte.

De vrees bestond dat, als gevolg van de enorme concurrentiestrijd tussen nationale en internationale conferenties om (weer) in de gunst te komen van deelnemers en het feit dat de ORD dit keer in Vlaanderen werd georganiseerd, het aantal deelnemers zou dalen. Dit bleek niet het geval! Even wat 'statistiek'. In 1994 (Utrecht) bedroeg het aantal deelnemers ongeveer 460; in 1995 (Groningen) daalde het aantal deelnemers tot ongeveer 420; in 1996 (Tilburg) werden ongeveer 460 deelnemers genoteerd. In Leuven konden, afgaande op de deelnemerslijst per 14 mei, 432 deelnemers worden geteld. Dit aantal zou, naar ik uit informele bron heb vernomen, tijdens de conferentie zijn opgelopen tot ongeveer 500. Er is dus

geen sprake van een structureel dalende tendens wat betreft de bezoekersaantallen.

Hoe zat het met de verdeling Nederlanders en Vlamingen (en overige nationaliteiten) in Leuven? Weer afgaande op de geregistreerde deelnemers per 14 mei, werd de ORD 1997 bezocht door 245 Nederlanders, 181 Vlamingen, drie Amerikanen, twee deelnemers uit het Verenigd Koninkrijk en twee Zuid-Afrikanen.

De presentaties (paperlezingen, symposia, postersessies en fora) waren gegroepeerd in een negental thema's, die in deze kroniek worden besproken, te weten: leren en instructie, methodologie en evaluatie, vakdidactiek, onderwijs en samenleving, curriculum, lerarenopleiding en leraarsgedrag, hoger onderwijs, beleid en organisatie in het onderwijs en tusslotte beroeps-, bedrijfsopleidingen en volwasseneneducatie (willekeurige volgorde). Op de vastgestelde deadline voor het indienen van paper- en symposiumvoorstellen waren onvoldoende voorstellen ingediend, maar na het uitvoeren van enige druk en het verschuiven van de uiterste inleverdatum beschikte de organisatiecommissie alsnog over voldoende aanbod: 160 paper- en symposiumpresentaties, 6 fora, 12 posters en 3 demonstraties.

De Corte ging hierop kort in door te stellen dat de enorme druk, vooral in Nederland, om in internationale tijdschriften te publiceren mogelijk debet kon zijn aan het aanvankelijk te lage aantal papervoorstellen.

Verloop, voorzitter VOR, ging hier in zijn toespraak ook op in. De discussie over het aantal toegelaten Nederlandstalige publicaties is de laatste tijd actueel, mede als gevolg van de door het Pedagogisch Tijdschrift aangezwengelde discussie over "de herwaardering van het Nederlands als wetenschappelijke taal". Deze actie heeft inmiddels een vervolg gehad door het organiseren van een bijeenkomst over deze kwestie door het Pedagogisch Tijdschrift in Utrecht, waarvoor alle (hoofd)redacteurs van Nederlandstalige wetenschappelijke tijdschriften waren uitgenodigd. Dat heeft geresulteerd in een concretisering van het oorspronkelijke door het Pedagogisch Tijdschrift opgesteld 'manifest', waarin wordt gesteld dat iedereen met een onderzoeksaanstelling in een periode van vijf jaar zes publicaties in gerefereerde tijdschriften zal moeten realiseren, waarvan er maximaal drie in het Nederlands geschreven

mogen worden. Inmiddels heeft het managementteam van het Interuniversitair Centrum voor Onderwijskundig Onderzoek (ICO) zich over deze uitspraak gebogen en geconcludeerd dat één publicatie in het Nederlands de limiet is. We weten als onderwijsonderzoekers nu waar we ons aan te houden hebben, of zouden we toch nog worden gehoord?

In deze kroniek zijn de indrukken van een aantal ORD-gangers gegroepeerd en beschreven rond de hiervoor genoemde negen thema's. Uiteraard geven de verslagen geen volledig beeld van datgene wat in de thema's aan de orde is gesteld. De persoonlijke belangstelling van de auteurs is richtinggevend geweest bij de selectie en bespreking van de onderwerpen.

Tenslotte, ondertussen is in Twente de organisatiecommissie ORD 1998 van start gegaan om de vijftiengste aflevering van de ORD voor te bereiden. De Twentenaren zullen er ongetwijfeld een bijzondere ORD 1998 van maken.

Leren en Instructie (E.R. Savelsbergh, Technische Universiteit Eindhoven en Universiteit Twente Enschede)

Het thema 'Leren en Instructie' was dit jaar ingevuld met vier symposia, een papersessie en een postersessie, waar in totaal 21 bijdragen gepresenteerd werden. Daarnaast was er binnen het thema een forum georganiseerd. Terugkerende onderwerpen waren constructief en ontdekkend leren, de krachtige leeromgeving waarin dit ontdekken plaats moet vinden, en tot slot 'leren leren' en de leerstijl die je nodig hebt om in zo'n krachtige leeromgeving te overleven.

Het eerste symposium binnen het thema 'Leren en instructie' was de sessie rond 'Elementen van krachtige leeromgevingen voor realistisch rekenonderwijs', waar medewerkers van het Utrechtse Freudenthal Instituut hun onderzoek presenteerden. Gravemeijer presenteerde een paper over het gebruik van een rekenrek als instructiemiddel bij het leren rekenen. Hij liet zien hoe het rekenrek eerst gebruikt werd als 'model van' het op te lossen probleem, dat wil zeggen als onmisbare representatie waarmee een probleem tastbaar gemaakt wordt, en hoe in de loop van het leerproces de rol van het rekenrek verschuift naar

die van 'model voor', dat wil zeggen een metafoor voor het oplossen van een bepaald type problemen. Feijs beschreef in haar presentatie de ontwikkeling van het meetkunde-curriculum voor de Amerikaanse 'Middle School', een groot project waarbij het Nederlandse 'wiskunde A' onderwijs duidelijk inspiratiebron is geweest. Het materiaal dat Feijs presenteerde had betrekking op het onderwerp 'hoeken', gepresenteerd aan de hand van kijklijnen in de Grand Canyon. Net als in de presentatie van Gravemeijer was ook hier duidelijk dat het proces van ontdekkend leren voortdurende subtiele sturing behoeft die impliciet geboden kan worden door de keuze van de juiste contexten. De context zal, om op de juiste manier aan te sluiten bij de ervaringswereld van de lerende, redelijk bekend moeten zijn maar niet overbekend. Van den Heuvel-Panhuizen vervolgde het symposium met een bijdrage over de aspecten van realistisch rekenonderwijs die vooral van belang zijn voor meisjes. Uit haar onderzoek bleek dat aansprekende contexten maar een deel van het verhaal zijn: ook het type vaardigheid waarop een beroep wordt gedaan bepaalt de prestaties van meisjes. De laatste presentatie in deze sessie was van Van den Boer, die vertelde hoe allochtone leerlingen met realistisch rekenonderwijs uit de voeten kunnen. Aanleiding voor dit onderzoek was de bevinding dat de prestaties van allochtone leerlingen bij alle vakken, met uitzondering van wiskunde, inlopen op die van autochtone leerlingen. Zij vond in haar onderzoek dat allochtone leerlingen bij het oplossen van contextopgaven even vaak goede antwoorden vinden als autochtone leerlingen, maar dat allochtone leerlingen vaker niets opschrijven als ze het niet helemaal weten. Verder bleken binnen de allochtone groep de leerlingen van Turkse afkomst er in positieve zin uit te springen. De discussiant (Verschaffel, KUL) stelde vooral vragen over de theoretische onderbouwing van het realistisch rekenonderwijs: er zijn wat algemene uitgangspunten en er is een praktijk, maar de uitgangspunten geven nog weinig houvast voor de praktijk. Zo wilde Verschaffel weten waarom de leerling altijd moet geloven dat hij de wiskunde zelf ontdekt, en of het gebruik van het rekenrek volgt uit algemene principes.

Van het symposium 'Metacognitie en het

oplossen van wiskundige problemen in de bovenbouw van de basisschool' heb ik alleen de eerste presentatie gezien; daarna ben ik 'overgestoken' naar het parallel aan dit symposium geprogrammeerde forum. In deze eerste presentatie (Veenman en Van Dam, RUL) werd een training in metacognitieve vaardigheden gepresenteerd die toegepast was in een groep 10- tot 12-jarigen bij het leren oplossen van redactiesommen. Er werd een significante prestatieverbetering gerapporteerd ten opzichte van de controlegroep met gewone instructie. Helaas was alleen zeer nabije transfer gemeten waardoor de uitkomst weinig zegt over de werkelijke verbetering van metacognitieve vaardigheden.

Tijdens het forum werd goed duidelijk wat de breedte is van het thema leren en instructie. Het forum droeg de titel 'The outcomes of constructive learning environments'. Waar De Jong (UT) een technische invulling gaf aan het begrip leeromgeving door het auteursstelsel voor simulatieleeromgevingen 'SimQuest' te presenteren, koos Resnick een heel andere invulling door het over de klas als leeromgeving te hebben. Ook Scheerens (UT) maakte deel uit van het panel; zijn onderzoek richt zich op het schoolniveau en hij opperde dan ook de gedachte dat er ook op schoolniveau iets als een leeromgeving bestaat. Net als de term leeromgeving kreeg ook de term constructivisme verschillende betekenissen: het woord kan staan voor een beschrijvende theorie van menselijk leren of voor een prescriptief model voor het inrichten van onderwijs. Mede door deze terminologische verschillen kwam de discussie niet echt op gang.

Terwijl in het onderwijskundig onderzoek het constructivisme leeft en soms aanleiding geeft tot verhitte discussie, blijkt het constructivistisch gedachtegoed in de praktijk nog ver te zoeken: in het symposium rond leerconcepties, leerstijlen en leerstrategieën in het secundair onderwijs presenteerden Waeytens, Vandenberghe en Lens (KUL) een onderzoek onder leraren in het secundair onderwijs (bij de vakken wiskunde en Nederlands) waaruit bleek dat zo'n 70% van de onderzochte leraren een beperkte opvatting had van 'leren leren', gekoppeld aan een leeropvatting die vooral gericht was op informatieoverdracht. Deze leraren zien als hun voornaamste taak vak-

inhoudelijke informatie over te dragen, en geven alleen als er problemen zijn wat tips over leren leren. Tijdens hetzelfde symposium presenteerden De Jager en Reezigt (RUG) een onderzoek naar de relaties tussen metacognitieve vaardigheden en de prestatie bij Nederlands en wiskunde. De gebruikte data kwamen uit een grootschalig survey-onderzoek waardoor zowel leerlingeffecten als school- en leerkrachtheffecten gemeten konden worden. In dit onderzoek kon een negatief verband tussen metacognitie en wiskundeprestaties worden aangetoond. Tussen metacognitie en prestaties bij Nederlands werd geen verband gevonden. Er werden verschillende verklaringen geopperd voor de – toch wel verrassende – negatieve relatie. Helaas waren de gebruikte data niet geschikt om in detail uit te zoeken waardoor deze relatie veroorzaakt wordt. Een tweede opvallende bevinding uit het onderzoek van De Jager en Reezigt was dat, in vergelijking met de vakinhoudelijke prestaties, de metacognitieve vaardigheden van leerlingen veel meer individueel bepaald zijn en veel minder afhangen van school en docent. Lamberigts en Verhoeven (KUN) presenteerden de resultaten van Verhoevens afstudeeronderzoek, waarbij een aangepaste versie van de 'inventaris leerstijlen' van Vermunt (1992) werd gebruikt bij leerlingen in 4 HAVO/VWO. De analyse van de data leverde een wat andere indeling van leerstijlen op dan in het bekende origineel van Vermunt. Met name de betekenisgerichte en de toepassingsgerichte leerstijl bleken niet te scheiden. Dit zou wellicht verklaard kunnen worden doordat het verschil tussen beide leerstijlen in het algemeen voortgezet onderwijs ook 'in het echt' nog niet zo duidelijk is. Verhoeven had ook geprobeerd onderwijsbehoefte te meten en deze te relateren aan de gemeten leerstijlen. Dit onderdeel van het onderzoek bleek helaas moeilijk te interpreteren, enerzijds door de manier waarop de schalen geconstrueerd waren, anderzijds doordat niet duidelijk is of bijvoorbeeld iemand die behoefte heeft aan frontaal onderwijs ook werkelijk baat heeft bij frontaal onderwijs. Vermunt (RUL) trad in deze sessie op als discussiant. Hij toonde zich kritisch over de voorgestelde nieuwe indeling. Zijn voornaamste bezwaren waren: ten eerste dat de nieuwe indeling een teruggang betekende naar de oude tweedeling tussen diepe en

oppervlakkige stijlen en ten tweede dat de inventaris leerstijlen niet zonder meer vertaald kan worden van het hoger onderwijs naar andere onderwijstypen. In dezelfde sessie presenteerden Klatter, Lodewijks en Aarnoutse (KUN/KUB) over de constructie van een meet-instrument leerconcepties voor 12- tot 14-jarigen waarbij in ieder geval het laatste bezwaar omzeild werd doordat de leerstijlen gereconceptualiseerd waren met behulp van interviews alvorens er een vragenlijst geconstrueerd werd die toegespitst was op de terminologie en opvattingen van deze leerlingen. Hier kwam de vraag aan de orde wat eigenlijk de toegevoegde waarde van een grootschalig vragenlijstonderzoek nog was bij het beantwoorden van de onderzoeksvraag. Een meer algemene opmerking die ook op het onderzoek van Verhoeven en op het onderzoek van De Jager en Reezigt sloeg was dat vragenlijsten geen inzicht geven in het werkelijk studiedrag maar alleen in de opvattingen van studenten; er is daarom dringend behoefte aan validatie-onderzoek waarin het werkelijke gedrag gemeten wordt.

Van de posters die gepresenteerd werden tijdens de postersessie hoorden er twee bij het thema leren en instructie. Een van de twee posters presenteerde ik zelf, waardoor ik de poster van Mets – over ‘de rol van schoolse interactie bij mattheuseffecten in leesvaardigheid’ – niet gezien heb. Mijn poster beschreef de ontwikkeling van een computerondersteunde leeromgeving voor probleemoplossen in elektrostatica-onderwijs in de onderbouw van de studie natuurkunde (Savelsbergh, De Jong en Ferguson-Hessler, TUE/UT). Deze leeromgeving biedt studenten ondersteuning bij het oplossen van vergelijkingen en bij het tekenen van grafieken. Doel van het programma is meer nadruk te leggen op het begrijpen en interpreteren van het probleem. Resultaten van een eerste experiment wijzen erop dat het werken met het pakket positief gewaardeerd wordt door studenten. Er kon echter nog geen leerwinst t.o.v. de traditionele methode worden aangetoond.

De overlap tussen het thema leren en instructie en het gast-thema van dit jaar: ‘vakdidactiek’, was opvallend en bij een deel van de sessies, zoals bijvoorbeeld het symposium over elementen van krachtige leeromgevingen, leek de indeling bij een van de twee thema’s

willekeurig. In de laatste presentatieronde was er binnen het thema vakdidactiek een symposium over ‘model in de beta-didactiek’ dat nauw aansloot bij het onderwerp krachtige leeromgevingen en vooral ook bij mijn eigen belangstelling, zodat ik de laatste twee parallel geprogrammeerde onderdelen van leren en instructie (een symposium over ‘computers in het onderwijs’ en een papersessie) gemist heb.

Literatuur

Vermunt, J.D.H.M. (1992). *Leerstijlen en het sturen van leerprocessen in het hoger onderwijs: naar procesgerichte instructie in zelfstandig leren*. Lisse: Swets & Zeitlinger

Methodologie en Evaluatie (H.J. Vos, TO, Universiteit Twente, Enschede)

Rond het thema ‘Methodologie en Evaluatie’ hadden de coördinatoren Onghena (KUL) en Sanders (CITO) dit jaar twee papersessies geprogrammeerd, waarin een breed scala aan onderwerpen was samengebracht. De eerste papersessie was gegroepeerd rondom het subthema “Evaluatie van het onderwijs”, terwijl de tweede papersessie ging over ‘Toetsing van het onderwijs’. Beide papersessies werden druk bezocht en er was genoeg tijd voor discussie.

In de presentatie van Colla (Vlekho, Brussel) werd gedemonstreerd hoe dynamische kengetallen, ontleend aan technieken die zowel in de theorie van manpowerplanning als bij het gebruik van sterftetafels in demografische populatievoorspellingsmodellen veelvuldig worden gebruikt, ook heel goed bruikbaar zijn voor het bepalen van de efficiëntie in het onderwijs. Op basis van deze dynamische kengetallen kunnen opleidingen o.a. worden vergeleken op een aantal relevante aspecten, zoals de kans dat een student zijn doctoraal diploma behaalt in de nominale studietijd, met één jaar vertraging zijn doctoraal diploma behaalt, ooit het doctoraal diploma behaalt, het percentage ingezette middelen verbruikt door de studenten die het doctoraal diploma niet behalen, enz. Aange-toond werd dat de berekende kengetallen sterk beïnvloed werden door de slaagkansen in de propaedeuse. Een opvallend resultaat was dat er op de universiteiten in Vlaanderen grote regionale verschillen bestaan; in West-Vlaanderen liggen de slaagkansen bijvoorbeeld beduidend hoger

dan in Belgisch-Limburg. Opvallend waren de lage slagingspercentages voor modieuze studierichtingen als psychologie, hetgeen werd toegeschreven aan de grote toestroom van minder gemotiveerde studenten.

Van der Tuin (GION) toonde in 'Verschillen tussen meer en minder effectieve docenten binnen scholen' aan dat, naast verschillen in leerling- en schoolkenmerken, ook docentkenmerken van invloed zijn op verschillen in leerprestaties. Als theoretisch kader werd het onderwijsactiviteitsmodel van Creemers gebruikt, waarbinnen docentgedrag wordt gelokaliseerd op klasniveau. Nadat voor 17 scholen uit het nationale cohort leerlingen onderzoek (VOCL'93) met behulp van multilevelanalyse was gecorrigeerd voor leerling- en schoolkenmerken, bleef er nog 8% onverklaarde variantie als docentvariabelen op klasniveau over. Getracht werd deze resterende variantie te verklaren door in lesobservaties naar het lesgedrag van de docent en de interactie tussen docent en leerling te kijken, waarbij ook de actieve betrokkenheid van leerlingen werd gescoord. Geregistreerd werd bijvoorbeeld hoe de lestijd verdeeld werd, hoe vaak leraren bij elkaar kwamen, of leraren afspraken maakten over gezamenlijke toetsen, of leraren veel overleg met collega's hadden om over het onderwijs van gedachten te wisselen, manier van omgaan met elkaar (formeel/informeel), of de leerling centraal stond of dat kennisoverdracht centraal stond, of leerlingen individueel werkten of dat ze samenwerkten. Uit de resultaten kon voorzichtig worden geconcludeerd dat heldere afspraken en een adequate tijdsbesteding op zowel klas- als schoolniveau een positieve invloed lijken te hebben op de leerprestaties.

De Corte, Verschaffel en Knoors (KUL) deden verslag over de 'Bruikbaarheid van het Nederlands instrumentarium voor periodieke peilingen van het onderwijsniveau (PPON) voor evaluatie van de nieuwe eindtermen wiskunde in het basisonderwijs in Vlaanderen'. Om de bruikbaarheid van de PPON-items na te gaan, werden deze door de drie onderzoekers gecategoriseerd volgens de Vlaamse eindtermen. De belangrijkste algemene conclusie die hieruit getrokken werd was dat sommige eindtermen zeer goed vertegenwoordigd waren in het PPON-instrumentarium. Een groot aantal

eindtermen bleef echter ook onderbelicht, vooral de eindtermen die hogere denkvaardigheden vereisen zoals probleemoplossende vaardigheden, het reflecteren op het eigen probleemoplossingsproces en het ontwikkelen van een kritische houding ten aanzien van cijfermateriaal. Deze eindtermen zijn echter ook moeilijk te vatten in schriftelijke korte-antwoord items, waarvan het PPON-instrumentarium zich bedient. Geconcludeerd werd dan ook dat, naast korte-antwoord vragen met het grote voordeel van hun relatief eenvoudige constructie, afname en verwerking, andere toetsvormen noodzakelijk zijn om de hele variëteit aan eindtermen te kunnen evalueren.

Straetmans en Eggen (CITO) rapporteerden over een 'Vergelijkend onderzoek naar de nauwkeurigheid en efficiëntie van drie toetsprocedures voor een plaatsingstoets in de volwasseneneducatie'. Vanwege de grote spreiding in de vaardigheid van instromende cursisten bij de volwasseneneducatie (VE) werd voor (beperkt) computergestuurd adaptief toetsen (CAT) gekozen om instromende cursisten voor het onderdeel rekenen/wiskunde toe te wijzen aan één van de drie niveaus waarop het onderwijs verzorgd werd. De plaatsingstoets was een zogenoemde tweefasen toets, waar elke kandidaat in de eerste fase dezelfde toets te maken kreeg en in de tweede fase werd toegewezen aan één van de drie niveaus op basis van de toetsprestatie in de eerste fase. De zak/slaag beslissing werd vervolgens genomen op basis van de totale toets. Om toetsprocedure-effecten (o.a. computervrees en weinig of geen ervaring hebben met computertoetsen) na te gaan werd, naast een CAT en een papieren plaatsingstoets (PBT), ook een gecomputeriseerde versie van de PBT aangeboden, de CBT (computer based test). CAT bleek 25% minder items nodig te hebben om met dezelfde nauwkeurigheid als de PBT en CBT plaatsingsbeslissingen te nemen. Verder werden in een experimenteel design 90 proefpersonen at random aan verschillende condities toegewezen. Uit de analyses bleek dat er geen sprake was van een toetsprocedure-effect maar wel van een volgorde-effect (procedure eerste en tweede af te nemen toets).

Vos (UT) liet zien hoe dynamische programmeringstechnieken konden worden gebruikt om 'Optimaal sequentiële beslisregels

voor het 3-actie beheersingsprobleem' af te leiden binnen het theoretisch kader van de sequentieel Bayesiaanse besliskunde. Om de kans op een correct antwoord te modelleren, bij een gegeven ware score, werd uitgegaan van een binomiale verdeling. Voor de verliesfunctie werd zowel een drempel als een lineaire verliesfunctie onderzocht. Tenslotte werd voor de a priori ware score een beta verdeling geassumeerd. Aan de hand van empirische gegevens voor concept-leren bij leerling-verpleegkundigen werd gedemonstreerd hoe na ieder aangeboden item één van de drie classificatiebeslissingen (t.w. slagen, partieel beheersen en zakken) of een continueerbeslissing kon worden genomen. Uit de resultaten bleek dat voor duidelijke 'beheersers', 'partiële beheersers' en 'niet beheersers' reeds na enkele items een beslissing kon worden genomen. Leerling-verpleegkundigen wier prestaties echter in de buurt van de grensscores lagen moesten relatief veel items beantwoorden alvorens met voldoende accuraatheid een classificatiebeslissing kon worden genomen.

Eggen (CITO) ging in op het verschil tussen statistisch schatten en statistisch toetsen bij CAT. Traditioneel is CAT gericht op het zo efficiënt mogelijk schatten van de vaardigheid van een leerling onder een itemresponstheorie model, zodat uit een gecalibreerde itembank voor elke leerling op grond van de resultaten op de al afgenomen items het meest informatieve item kan worden aangeboden. Aan de hand van simulatiestudies werd echter aangetoond dat voor het classificeren van leerlingen in een beperkt aantal categorieën (bijv. geslaagd/gezakt) schattingsalgoritmes niet het meest geëigend zijn maar men beter zijn toevlucht kan nemen tot statistisch toetsen. Voor de statistische toets werd de uit de literatuur (o.a. bij kwaliteitscontrole) bekende Sequential Probability Ratio Test (SPRT) van Wald genomen, welke beduidend betere resultaten liet zien dan traditionele itemselectie methoden gericht op het efficiënt schatten van de vaardigheidsparameter. Het gemiddeld aantal benodigde items bleek lager te liggen, terwijl het percentage correcte beslissingen hoger bleek te liggen.

Tenslotte deed Sanders (CITO) verslag van een studie naar 'Het samenstellen van parallelle toetsen' met behulp van mathematische programmeringsmodellen, die voldoen aan de psy-

chometrische, inhoudelijke en praktische specificaties van toetsconstructeurs. De voorgestelde methode ging terug op de uit de literatuur bekende matching random subsets methode van Gulliksen voor het samenstellen van parallelle toetsen, waarbij paren van items worden gekoppeld op basis van hun moeilijkheidsgraad en item-testcorrelatie. In deze methode wordt vervolgens ieder item van elk gekoppeld paar of drietal at random toegewezen aan een toets, bijvoorbeeld aan een voortoets en een natoets bij een pretest-posttest design. Op het blote oog ('timmermans oog') worden bij Gulliksen's oorspronkelijke methode punten in een grafiek met als coördinaten de moeilijkheidsgraad en item-testcorrelatie zo goed mogelijk aan elkaar gekoppeld, waardoor uitkomsten afhankelijk kunnen zijn van de toetsconstructeur. Bij gebruik van mathematische programmeringstechnieken als integer programming kunnen echter formele criteria, zoals het minimaliseren van de afstanden tussen de moeilijkheidsgraad en de item-testcorrelatie, als doelfunctie worden gehanteerd om items aan elkaar te koppelen. Nieuw in Sanders' benadering was dat gebruik werd gemaakt van zogenoemde Greedy algoritmes om items aan elkaar te koppelen, waarmee op een efficiënte manier zowel zwak parallelle als sterk parallelle toetsen konden worden samengesteld.

Vakdidactiek (A.E. van der Valk en M.J. Vollebregt, Centrum voor Didactiek van Wiskunde en Natuurwetenschappen, Universiteit Utrecht).

Het thema 'Vakdidactiek' bestond uit een papersessie, een postersessie en twee symposia. Deze boden de gelegenheid om van gedachten te wisselen over onderwerpen die verschillende schoolvakken met elkaar gemeen hebben. We beperken ons tot de symposia die rond praktijkkennis van leraren en modellen werden georganiseerd.

Het symposium 'Vakdidactisch onderzoek vanuit het perspectief van de praktijkkennis van leraren' was een ontmoetingsplaats voor vakdidactici en algemeen onderwijskundigen.

Beijaard (RUL) schetste in zijn inleiding op dit symposium waarom het onderzoek naar het denken van leraren zich richt op de praktijkkennis van leraren. Hij gaf aan wat de aard van

die praktijkkennis is: persoonsgebonden, geïntegreerde kennis afkomstig uit verschillende bronnen, dicht bij de praktijk en dus gebruiksklaar. Als essentieel onderdeel van praktijkkennis ziet hij de vakdidactische kennis of zoals Shulman (1986) het noemt, pedagogical content knowledge (PCK). Deze kennis ontstaat door transformatie van vakkennis in schoolvak kennis. Als fundamentele problemen in het onderzoek naar praktijkkennis ziet hij onder andere de vraag naar de methodologie: hoe kom je tot kwalitatief goede (valide, betrouwbare) uitkomsten?

Van der Valk en Broekman (UU) reflecteerden op de methode waarop zij onderzoek hebben gedaan naar de aanvankelijke vakdidactische praktijkkennis van studenten aan het begin van de lerarenopleiding. Zij hebben studenten die net met de lerarenopleiding beginnen in een contextrijke omgeving (een leslokaal met veel wiskunde-spullen, zoals een kubus, kartonnen driehoeken en schaar en papier, maar zonder wiskundeboek) geplaatst. Ze gaven hun de opdracht een les voor een brugklas over oppervlakte voor te bereiden. Daarna lieten zij hen over hun les vertellen en tot slot vroegen ze door op onderwerpen als 'waar heb jij deze kennis opgedaan?' Zij vonden dat deze methode bij studenten veel praktijkkennis die vakdidactisch van aard was naar boven haalde, kennis uit hun 'praktijk als levenslang lerende'. Dankzij de aangeboden omgeving en opdracht kon deze kennis gearticuleerd worden. Daarnaast signaleerden zij ook een leereffect in die zin dat de studenten hun praktijkkennis verder ontwikkelden.

Van Driel en Beijaard (RUL) onderzochten de praktijkkennis over modellen van ervaren docenten in één van de natuurwetenschappen die een omscholingscursus voor het nieuwe twee-fase-vak Algemene Natuur Wetenschappen (ANW) volgden. In het huidige onderwijs in de exacte vakken worden weliswaar veel modellen gebruikt, maar wordt in het algemeen niet op het modellenbegrip als zodanig ingegaan. De expliciete opdracht voor modellen in het nieuwe vak ANW, dat in de tweede fase ingevoerd zal worden, vereist evenwel dat de vakleraren hun kennis op dat gebied uitbouwen. Van Driel c.s. vond dat docenten sterk verschillen in de manier waarop zij het modellenbegrip in het natuurwetenschappelijk onder-

wijs gebruiken. Die verschillen lijken verband te houden met de visie op leren en onderwijzen van de docenten.

Alblas (LUW) heeft geprobeerd praktijkkennis van leraren te gebruiken om de didactiek van Natuur- en Milieu Educatie (NME) te verbeteren. Hij stelde dat deze verwaarloosd is door het accent op het overdragen van de 'milieuboodschap'. In plaats daarvan zou NME naar 'Bildung' moeten streven.

Hulshof (RUL) rapporteerde over onderzoek naar het gebruik van analogieën (een belangrijke component van PCK volgens Shulman) door docenten Nederlands. Zijn doel was o.a. het begrip PCK te verhelderen door vakspecifieke praktijkvoorbeelden van de component 'analogie' te zoeken. Hij vond er een aantal voor het vakonderdeel tekstbegrip, zowel in lessen als in leerboeken. Maar het aantal was kleiner dan verwacht en blijkt in de les sterk persoonsgebonden. Hulshof was de enige alfa die aan dit symposium bijdroeg. Hij verwees veelvuldig naar resultaten van onderzoek in de exacte vakken. Dat geeft aan dat het onderwerp 'praktijkkennis' inderdaad uitgaat boven de traditionele grenzen tussen de alfa, bèta, en gamma vakken.

Terwel (VU, UvA) zette als discussiant een aantal kritische kanttekeningen bij de onderzoeken. De meest prikkelende daarvan was zijn waarschuwing dat de gepresenteerde onderzoeken zouden kunnen bijdragen aan het dualisme in het onderwijs. De onderzoekers, allemaal bezig met het eerstegraads gebied van het VO, zijn gericht op 'Bildung', de 'Ausbildung' ofwel beroepsvoorbereiding overlatend aan tweedegraders in het MAVO/LBO gebied. Hij pleitte ervoor toch vooral ook aandacht aan het tweede gebied te besteden, en de nadruk te leggen op de proceskant van het onderwijs. Helaas was er weinig tijd meer over om hierop in te gaan, maar in de wandelgangen is de discussie voortgezet.

Het andere vakdidactische symposium werd verzorgd door onderzoekers van het Centrum voor Didactiek van Wiskunde en Natuurwetenschappen (CDB) van de UU. Daarin stond een overkoepelend begrip in de β -vakken centraal: Modelvorming in de β -didactiek.

Lijnse opende dit symposium met een schets van de problematiek: uit onderzoek blijkt dat leerlingen minder dan verwacht in

staat zijn om onderwezen modellen te begrijpen en hiermee te werken. Het blijft daarbij vooralsnog onduidelijk in hoeverre zij de conceptuele inhoud niet begrijpen, dan wel te weinig beschikken over geschikte vaardigheden. In het CDB wordt sinds enige tijd geprobeerd om middels ontwikkelingsonderzoek meer greep te krijgen op deze problematiek. In dergelijk onderzoek wordt nieuw onderwijs ontworpen waarbij de verwachting is dat leerlingen dan een stap verder komen in het leren begrijpen en werken met modellen. Het onderwijs- en leerproces, dat aan de hand van dit ontwerp ontstaat, wordt nauwkeurig gevolgd en de analyse hiervan moet leiden tot bijstelling van het onderwijs. Zo levert het onderzoek enerzijds prototypes van lesmateriaal op, waarvan empirisch is aangetoond dat succesvolle implementatie mogelijk is, en anderzijds is er een opbrengst voor vakdidactische theorievorming en nascholing.

Doorman liet aan de hand van een concreet wiskundig voorbeeld zien hoe grafieken voor leerlingen kunnen veranderen van een 'model voor' integratie- en differentiatieproblemen. Hierbij wordt de grafiek voor leerlingen meer en meer een op zichzelf staand object dat kan gaan functioneren als 'model voor' wiskundig redeneren. Voorwaarde daarbij is dat de ervaring en kennis van leerlingen als uitgangspunt wordt genomen en dat zij gelegenheid krijgen om de wiskunde, onder begeleiding van de docent, heruit te vinden.

Van Hoeve-Brouwer stelde in haar presentatie de rol van modellen in chemie schoolboeken centraal. Uit haar analyse kwam naar voren dat modellen vaak niet als zodanig expliciet benoemd worden en dat er regelmatig een logische samenhang ontbreekt in de overgang van het ene naar het andere model. Vervolgens lichtte zij de verschillende manieren toe waarop een model in het onderwijs geïntroduceerd kan worden en noemde ze een aantal modelkenmerken, m.b.t. functie, aard en werkwijze. Tot slot werd dit alles gekoppeld aan enkele criteria voor het ontwerpen van lesmateriaal.

Vollebregt en Janssen lieten in hun bijdragen zien hoe onderwijs voor de introductie en ontwikkeling van een model, in respectievelijk het natuurkunde- en biologie-onderwijs, concreet kan worden vormgegeven. Beiden werkten vanuit een probleemstellende benadering.

Dat wil zeggen dat leerlingen door middel van geschikt gekozen activiteiten worden gestimuleerd om zelf bepaalde belangrijke inhoudelijke vragen/problemen te gaan zien en formuleren. In volgende activiteiten worden zij in de gelegenheid gesteld om kennis te ontwikkelen waarmee zij deze problemen kunnen oplossen, waarbij de oplossing tevens weer nieuwe vragen oproept. Vollebregt werkte dit uit in een lessenserie ter introductie van een deeltjesmodel voor het beter begrijpen van gaswetten en faseveranderingen. Hierin werd gepoogd om leerlingen goede redenen in handen te geven voor op het eerste gezicht vreemde hypothesen en om hun inzicht te geven in de aard en functie van deeltjesmodellen in het algemeen.

Janssen liet zien hoe hij het probleemstellende uitgangspunt had verwerkt in een lessenserie over het immuunsysteem. In de hiervoor gebruikte ontwerptheorie speelden functionaliteit en optimaliteit een belangrijke rol.

Van Oers (VU) had als discussiant de taak om enige kritische kanttekeningen te maken. Hij vroeg zich af in hoeverre een 'model van...' ook werkelijk een eigen model van leerlingen wordt en aan welke eisen dit moet voldoen om later als 'model voor' te kunnen functioneren. Met betrekking tot probleemstellend onderwijs stelde hij dat meer mogelijkheid tot differentiatie zou kunnen voorkomen dat leerlingen het ontwerpen onderwijs te veel als een keurslijf gaan ervaren. Tevens meende hij dat de rol van de docent tijdens het leerproces niet duidelijk genoeg naar voren was gekomen. Ten slotte pleitte hij voor een longitudinale benadering van leren modelleren. Helaas was er geen tijd meer om over deze prikkelende opmerkingen te discussiëren.

Wat het aantal deelnemers aan de sessies betreft is het thema 'Vakdidactiek' een kwetsbaar thema. Het keert niet jaarlijks terug in het programma, met als gevolg dat zich (nog) geen groep onderzoekers heeft gevormd die gewend is rond dat thema op de ORD samen te komen. De bijdragen aan dit thema hebben echter duidelijk laten zien dat het een blijvende plaats op de ORD verdient. Daar komt bij dat in de onderwijskunde steeds meer wordt gepleit voor een vakspecifieke uitwerking van algemene onderwijskundige onderwerpen, bijvoorbeeld van de algemene vaardigheden in het 'studie-

huis'. Het thema 'vakdidactiek' op de ORD kan vakdidactici en onderwijskundigen de gelegenheid bieden nader tot elkaar te komen rond deze nieuwe problematiek.

Onderwijs en Samenleving (J.F.M. de Fraiture, SCO, Universiteit van Amsterdam)

'Onderwijs en Samenleving' als jaarlijks terugkerend thema op de ORD omvatte dit jaar vijf symposia, drie papersessies en een postersessie. De subthema's die hier aan bod kwamen waren onderwijsbeleid, de brede/ de verlengde schooldag, onderwijs/ongelijkheid en arbeidsmarkt, leerplichtverlenging, PRIMA-cohort-onderzoek en bijzondere groepen in het onderwijs.

In deze bijdrage zal ik verslag doen van de door mij bezochte symposia en papersessies. Het was jammer dat de onderwerpen in de bezochte papersessies erg uiteen liepen, zodat er na afloop geen discussie op gang kwam, zoals dat aan het einde van symposia, waarin de onderwerpen van het onderzoek meer op elkaar aansloten, wel gebeurde.

Roeleveld (UvA) begon de presentatie van zijn onderzoek naar effecten van leerrechtbeperkingen in het Nederlandse onderwijs met de uitspraak: "Een rechte stok is niet te vinden". Hiermee gaf hij aan dat er in de praktijk veel inefficiëntie bestaat in de theoretisch ontwikkelde schoolloopbanen. In zijn onderzoek wordt door middel van simulaties en scenario's gezocht naar een antwoord op de vraag hoe leerwegen effectiever kunnen worden gemaakt. Beperking van het recht op door de overheid bekostigd onderwijs is één van de voorgestelde mogelijke oplossingen. Uit het onderzoek blijkt dat eventuele maatregelen, bijvoorbeeld leerrechtbeperkingen, niet linea recta leiden tot efficiëntere leerwegen, zonder nadelige bijeffecten. Wanneer het voor vwo-ers financieel aantrekkelijker wordt gemaakt om het wetenschappelijk onderwijs te verkiezen boven het hoger beroepsonderwijs, zou dit kunnen leiden tot een groep WO-studenten, die minder gemotiveerd is en vervolgens een nieuwe groep 'students at risk' vormt.

Het symposium over de brede school/ de verlengde schooldag was zeer interessant en gezien het centrale thema 'Onderwijs en samenleving' ook erg van toepassing op de

huidige maatschappelijke discussie. Voor werkende ouders is de mogelijkheid om hun kinderen de hele dag naar school te laten gaan erg aantrekkelijk, maar wat zijn de gevolgen voor de overige participanten?

Van Erp (UvA) evalueerde een experiment Verlengde Schooldag op 19 scholen in Rotterdam, Den Haag en Utrecht. Het voornaamste doel van dit experiment is verbetering van de schoolprestaties en de schoolloopbanen. Uit het evaluatieonderzoek bleek dat de verlengde schooldag resulteert in meer zelfvertrouwen en exploratief gedrag bij leerlingen.

Veen (UvA) presenteerde haar onderzoek naar de relatie ouders-basisschool. Het uitgangspunt van dit onderzoek is dat in principe alle ouders betrokken zijn bij de school van hun kinderen. Dit komt in de praktijk echter niet altijd tot uitdrukking. Om de betrokkenheid van ouders bij de school te bevorderen zouden scholen meer onderwijs ondersteunend gedrag moeten vertonen, zodat er een positief klimaat ontstaat dat ouders stimuleert. Deze verbreding van het aanbod van de school zou naar mijn idee ook positieve invloed kunnen hebben op de verlengde schooldag in die zin, dat ouders een bijdrage zouden kunnen leveren aan de invulling van de verlengde schooldag.

Van Eerde (EUR) presenteerde tenslotte haar onderzoek naar de Brede School, dat het RISBO in opdracht van het Fonds Achterstandsbestrijding Onderwijs uitvoert. Het hoofddoel van het project Brede School is het terugdringen/opheffen van maatschappelijke achterstand van kinderen en jongeren in achterstandssituaties. Dit komt onder andere tot uitdrukking in de verschillende thema's die worden onderscheiden: schoolklimaat, vrije tijd, zorgverbreding, ouders en schoolomgeving.

De discussie aan het einde ging met name over de kwaliteit van de diverse activiteiten. Tevens werd een vraag gesteld over de eisen waaraan een leerkracht zou moeten voldoen. Ik zou hier zelf aan willen toevoegen, dat er naar mijn mening zeer hoge eisen worden gesteld aan de leerkrachten. Bovendien hebben beide initiatieven de nodige financiële consequenties, te meer als er hoge eisen worden gesteld aan de kwaliteit van onderwijs.

In het symposium 'Leerlingen en PRIMA' presenteerde De Fraiture (UvA) een onderzoek naar etnische diversiteit in afstand tussen taal-

en rekenscores tussen jongens en meisjes. De belangrijkste conclusie was dat de afstand tussen taal- en rekenscores voor een belangrijk deel verklaard wordt door geslacht en etnische groep, en in geringe mate door de opleiding van de ouders. Opvallend was dat niet alle etnische groepen afwijken van het algemeen Nederlandse scorepatroon. Wanneer we praten over afstanden tussen taal- en rekenscores tussen jongens en meisjes praten we over cultuurverschillen.

Roeveld (UvA) ging in het kader van Weer Samen Naar School in op de problematiek aangaande onderzoeksdesigns om leerlingen in het speciaal onderwijs te kunnen vergelijken met leerlingen in het basisonderwijs. Het PRIMA-onderzoek, waarin voor het eerst ook op landelijke schaal gegevens zijn verzameld op scholen voor speciaal onderwijs, biedt mogelijkheden om verschillende aspecten van de ontwikkeling van leerlingen op basisscholen en scholen voor speciaal onderwijs met elkaar te vergelijken. De vraag die uiteindelijk beantwoord moet worden is: Welke situatie het beste is voor 'risicoleerlingen': integratie binnen het basisonderwijs of segregatie binnen het speciaal onderwijs.

Van de Grift en Tabak (Inspectie) presenteerden een onderzoek naar de kwaliteit van het onderwijs aan achterstandsleerlingen in het basisonderwijs. Het blijkt dat schoolprestaties afhankelijk zijn van de schoolbevolking. Zo presteren Turkse en Marokkaanse leerlingen het best op een school waar het 1.25-leerlingendeel overheerst. Ook wat betreft de wijze waarop Nederlands-als-tweede taal aangeboden wordt verschilt per school. De kwaliteit van dit vak is gemiddeld het laagst, weinig differentiatie in instructie, op scholen waar het 1.9-leerlingendeel overheerst. Goede scholen met een hoog percentage allochtone leerlingen worden gekenmerkt door leerkrachten, die hun leerlingen motiveren en stimuleren en nauwkeurig hun vorderingen bijhouden.

Tijdens de papersessie 'Bijzondere groepen in het onderwijs' ging Scheepstra (RUG) in op het functioneren van kinderen met Downs syndroom in het regulier basisonderwijs. De stijging van deze groep kinderen in het basisonderwijs is een weerspiegeling van de maatschappelijke ontwikkeling om gehandicapten meer te integreren in de maatschappij. De voor-

naamste reden om kinderen met Downs syndroom te plaatsen in het regulier onderwijs is voor zowel ouders als leerkrachten het sociale aspect; zowel voor het gehandicapte kind zelf als voor medeleerlingen is het leerzaam om met elkaar om te gaan.

Van Welzenis (KUL) pleitte naar aanleiding van haar onderzoek naar het zelfconcept en toekomstperspectief van jongens met negatieve schoolervaringen voor meer aansluiting tussen verschillende leefwerelden. Wanneer de discrepantie tussen het schools en het vrije tijds concept minder wordt, wordt ook de kans op een delinquent gedragspatroon verminderd. Leerlingen die zich op het schoolse vlak niet kunnen bewijzen, kunnen dit op andere vlakken doen, bijvoorbeeld in sport en hoeven geen uitvlucht te zoeken in het criminele circuit.

Kools (LUW) nam sekse als criterium voor te onderscheiden groepen in het onderwijs. Zij voerde in het middelbaar agrarisch onderwijs een onderzoek uit om na te gaan of mannelijke leerlingen in een 'meisjesinrichting' met een meerderheid aan vrouwelijke klasgenoten zich gemakkelijker een plaats in de klas verwerven dan vrouwelijke leerlingen in een 'jongensinrichting' met een meerderheid aan mannelijke klasgenoten. Ze concludeerde dat voor onderwijssituaties geldt dat meisjes in jongensrichtingen minder goede ervaringen hebben dan jongens in meisjesrichtingen.

Curriculum (J.I.A. Visscher, TO, Universiteit Twente, Enschede)

Binnen het thema Curriculum werden dit jaar vier papersessies en één symposium verzorgd. In de presentaties werd ingegaan op de volgende onderwerpen: nieuwe ontwikkelingen in het curriculum onderzoek; onderzoek m.b.t. het curriculum van het basis- en het voortgezet onderwijs; de relatie tussen curriculumtheorie en curriculumpraktijk; curriculumontwikkeling: theorie en praktijk; curriculum en vakdidactiek. De presentaties waren doorgaans goed tot uitstekend verzorgd. In deze terugblik zal kort worden stilgestaan bij deze subthema's, met uitzondering van het eerste subthema dat helaas door dubbele programmering niet bezocht kon worden.

De papersessie 'Onderzoek m.b.t. het curriculum van het basis- en het voortgezet onder-

wijs' bestond uit drie interessante paperpresentaties, die qua inhoud echter niet direct op elkaar aansloten. In de eerste presentatie werden curriculaire bevindingen gepresenteerd op basis van zes case studies op scholen die met studiehuis-varianten werkten. Riemersma en Veugelers (UvA) rapporteerden een aantal verschillen tussen 'gewone scholen' en scholen waar het studiehuisconcept in meer of mindere mate vormkrijgt. Zo hebben scholen waar het studiehuisconcept is geïntegreerd niet alleen aandacht voor kennis maar ook voor persoonlijke vorming van de leerlingen, hebben leerlingen inzicht in het leerplan en wordt er in grotere blokken gewerkt (80 i.p.v. 50 minuten); evaluatie is in de planning opgenomen. De onderzoekers concludeerden dat op deze scholen de docent nog steeds een belangrijke sturende rol heeft en dat het leerboek nog steeds als belangrijkste leermiddel centraal staat. Zij kwamen tot de conclusie dat het studiehuisconcept op deze scholen nog sterk technocratisch vormkrijgt. Dat wil zeggen dat er nog veel vaste trajecten worden gepland, waarbij leerlingen af en toe eigen inbreng hebben. De bevindingen ontlokten veel informatieve vragen uit de zaal. De volgende twee presentaties kwamen voor rekening van Kuiper (UT) en Bos (UT) en gingen beide over de resultaten van het TIMSS-onderzoek (Third International Math & Science Study). Kuiper ging in zijn presentatie in op de scores op een internationale wiskundetoets die in 45 landen is afgenomen. In het onderzoek is gekeken naar de verschillen tussen verschillende schooltypes, tussen meisjes en jongens, en tussen verschillende leerjaren. Ook werd onderzocht in hoeverre de onderwerpen van de toets door de docenten in de les waren behandeld. Nederland doorstond de vergelijking met andere landen doorgaans goed: prestaties van de Nederlandse leerlingen horen thuis in de top 10. Bos ging op zoek naar voorspellers van verschillen in wiskundeprestaties in acht verschillende landen. In sommige landen werd een direct verband gevonden tussen klassegrootte en wiskundeprestaties. Ook de opleiding van de ouders heeft een positieve invloed op wiskundeprestaties. Bos vond een indirecte invloed van sekse en verwachting van de moeder op de houding van de leerling ten aanzien van wiskunde. Reacties uit de zaal ontlokten Bos de opmerking (met

enige terughoudendheid) dat motiverende factoren veel en factoren op schoolniveau weinig invloed hebben op wiskundeprestaties.

Het symposium 'De relatie tussen curriculumtheorie en curriculumpraktijk' werd goed bezocht. In het symposium stonden twee vraagstellingen centraal:

- Hoe kunnen curriculumpraktijken gekarakteriseerd worden en met welke modellen kan de ontwikkelpraktijk eventueel worden afgebeeld?
- Welke karakteristieken kan een praktijktheorie voor curriculumontwikkeling hebben en in hoeverre kunnen curriculummodellen daar deel van uitmaken?

In de eerste presentatie van Visscher (UT) werd op basis van interviews met 23 ontwerpers uit uiteenlopende contexten van onderwijs en training een viertal ontwerp-aanpakken getypeerd. In deze typering werd vooral gefocust op de onderliggende rationaliteit van de aanpakken, zodat deze in de praktijk herkenbaar bleven. Looy (SLO) presenteerde een praktijktheorie die is ontwikkeld binnen de context van de SLO. Deze theorie is gebaseerd op vier concepten: 1) leerplanniveaus en curriculaire verschijningsvormen, 2) leerplankundige componenten, 3) projectmatig werken, en 4) ontwikkeltrajecten. De praktijktheorie heeft verschillende functies: een gemeenschappelijke taal voor ontwerpers; communicatiemedium en beslissingskader; theoretisch kader voor beginnende ontwerpers; opsporingsmechanisme voor onvolkomenheden in een aanpak; en ontwikkeling van de ontwerp-professie. Moen (SLO) demonstreerde aan de hand van een concreet voorbeeld uit welke componenten lesmateriaal bestaat en welke functies deze componenten hebben, met als doel te komen tot een handelingsondersteunend instrument voor leerplanontwikkelaars. Door de stiptheid van de voorzitter was er na afloop van de presentaties voldoende ruimte voor de discussiant en de zaal om vragen te stellen en commentaar te leveren. Discussiant Terwel sprak zijn waardering uit over de presentaties, maar had moeite met het feit dat de beschreven ontwikkelprocessen niet gekoppeld waren aan een specifieke vakinhoud. Ook vond hij het jammer dat de drie presentaties vanuit het perspectief van de curriculumontwikkelaar waren geformuleerd, waardoor de indruk werd gewekt dat de

docent als mogelijke mede-ontwikkelaar uit het oog was verloren. Zowel sprekers als mensen uit de zaal onderstreepten echter het belang van de docent als (mede)-ontwikkelaar van lesmateriaal. De discussie moest vanwege het verstrijken van de tijd vroegtijdig worden afgekap.

Binnen de papersessie 'Curriculumontwikkeling: theorie en praktijk' werd ingegaan op het 'sociale karakter' van het ontwerp- en/of revisieproces. Deketelaere en Kelchtermans (KUL) verwoordden op een heldere manier hun ervaringen in een ontwikkeltraject waarbij docenten en onderwijskundigen samen concrete curriculummaterialen ontwikkelden ten behoeve van het bevorderen van de gelijkheid van onderwijskansen van jongens en meisjes in het secundair onderwijs. De onderzoekers concludeerden op basis van hun ervaringen dat de verschillen in kennis van de participanten in het ontwikkeltraject enerzijds kunnen leiden tot spanningen en mogelijke conflicten en tegenstellingen, maar anderzijds ook een groot potentieel bieden van complementaire competentie en effectiviteit. De onderzoekers rapporteerden dat samenwerking van docenten in een ontwikkeltraject een positieve invloed heeft op hun professionele voldoening en taakperceptie. Docenten werden zich bewust van hun eigen onderwijstheorie en spraken de intentie uit om na het project meer activiteiten op school te gaan verrichten. De tweede presentatie kwam voor rekening van Smit (SLO). Zij beschreef strategische beslissingen in revisietrajecten voor kerndoelen en examenprogramma's met als doel ervaringsgegevens te verzamelen voor toekomstige revisietrajecten. Smit en collega Boersma onderscheidden vijf strategische variabelen: aard van de revisie; aard van de (externe) communicatie, gehanteerd ontwikkelmodel, mate en wijze van aansturing door projectmanagement, mate en wijze van beïnvloeding. Smit pleitte ervoor om beslissingen omtrent veranderingen altijd gezamenlijk met de opdrachtgever te nemen. Van Oers (VU) stelde in zijn presentatie de gedeelde opvatting van de maakbaarheid van kant en klare curricula ter discussie. Aan de hand van een case study beschrijving in groep 1 en 2 van de basisschool liet hij zien welke rol leerlingen en leerkracht hebben in de opbouw van een curriculum. Curriculumontwikkeling in de klas

berust op conceptgeleide interpretatie en doelgerichte beslissingen van de leerkracht op grond van interacties met leerlingen, ondersteuners en curriculummateriaal. Dit leidde tot een discussie in de zaal over de rol en relatie tussen leerkracht en curriculumontwikkelaar en de wenselijkheid en mogelijkheid van docenten om een belangrijke rol in het curriculumontwikkelproces te spelen.

In de papersessie 'Curriculum en vakdidactiek' werd ingegaan op conceptvorming en vakdidactische thema's binnen drie inhoudsgebieden: lichamelijke opvoeding, geschiedenis, menswetenschappen. Het was jammer dat er zo weinig belangstelling voor deze sessie was. Door het accent op de inhoud en vakdidactiek was het voor niet-didactici moeilijk de presentaties op hun waarde te schatten alhoewel er door de presentatoren voldoende ruimte werd gegeven voor informatieve vragen. Van Assche (KUL) kwam in zijn heldere betoog tot een nieuw concept van lichamelijke opvoeding en beoordeelde de bestaanskracht van dit concept volgens verschillende criteria: past het binnen opvattingen over basisvorming; wordt het gesteund door literatuur; past het binnen bestaande leerplannen; waardenonderzoek, etc. Van Assche kwam tot de conclusie dat het concept bestaanskracht heeft, maar nog verdere uitontwikkeling behoeft. Goegebeur (VUB) ging in zijn presentatie in op het concept 'historisch besef' binnen het vak geschiedenis. Hij beschreef historisch besef vanuit o.a. verschillende dimensies (cognitief, affectief, evaluatief) en inhouden (tijdsbesef, werkelijkheidsbesef, historiciteit, identiteit, politiek, socio-economisch besef, moreel besef). Goegebeur constateerde dat historisch besef een waardegericht concept is geworden. De Gols (LP Menswetenschappen, VVKSO) exploreerde de mogelijkheden van 'menschwetenschappen' als specifieke studierichting op basis van een studie van 62 scholen die menswetenschappen geven. Het bleek belangrijk om verschillende vakken in samenhang aan te bieden (één thema vanuit verschillende invalshoeken als psychologie, sociologie, muziek, media, etc. belichten). Om als school in aanmerking te kunnen komen voor het inrichten van een dergelijke studierichting moet deze aan twee specifieke voorwaarden voldoen: mobiliteit van leerkrachten en vakdidactische competentie.

De Onderwijs Research Dagen hebben onder meer als belangrijk doel de onderzoeker op de hoogte te brengen van de laatste stand van zaken binnen zijn/haar domein. In onderstaande impressie zal dan ook vooral worden ingaan op de trends, binnen Lerarenopleiding en Leraarsgedrag, zoals die door mij zijn waargenomen, voorzien van enig commentaar. De onderstaande beschrijving is daarom niet chronologisch en zeker niet volledig.

Een veel gehoord geluid tijdens de ORD, niet alleen binnen de divisie 'Lerarenopleiding en Leraarsgedrag', was de angst voor een dalend belang van de ORD en andere meer lokale fora. Om dit idee te controleren ben ik de twee programmagidsen, die van vorig jaar in Tilburg en die van dit jaar te Leuven, eens ingedoken en aan het tellen geslagen. Een aantal dingen vielen op.

Allereerst blijkt er een toename te zijn van het aantal divisie-gerelateerde bijeenkomsten (10 in 1997 tegenover zes in 1996) en het aantal subbijdragen daarbinnen (16 in 1996 voor de gehele divisie, tegenover 29 in 1997). Deze toename uit zich niet alleen in absolute zin, maar ook in relatieve zin. Bij een stijging van het totale aantal bijdragen van 174 in 1996 naar 191 in 1997 is het aandeel van de divisie binnen het congres gegroeid van ongeveer 10% naar 15%. In 1996 waren er geen Vlaamse bijdragen binnen de divisie, ditmaal konden er drie genoteerd worden, hetgeen ongeveer 10% van het divisie-aandeel is. Over wat men van de hoogte van de genoemde aantallen moet vinden laat ik mij niet uit, duidelijk is dat de genoemde angst voor de divisie 'Lerarenopleiding en Leraarsgedrag' (gelukkig) vooralsnog niet gegrond blijkt.

Ten tweede kenmerkte deze ORD zich door een grotere verscheidenheid aan onderwerpen. Net als op de vorige ORD (Tilburg, 1996) was er ook dit jaar weer aandacht, in de vorm van een symposium, voor de veranderende rol van de docent onder invloed van het streven naar meer zelfstandigheid voor de leerlingen. De aandacht voor dit onderwerp was echter van meer bescheiden aard dan in het voorgaande jaar. Ook de tweede trendsetter van 1996, professionaliteit, was dit jaar weer veelvuldig

op het menu terug te vinden. Daarnaast was er dit jaar uitgebreide aandacht voor het thema nascholing, waarbij een speciale plaats gereserveerd bleek voor mentoring. Tenslotte stonden cognities en denkprocessen van docenten hoog op de onderzoekersagenda. De hiervoor genoemde onderwerpen sprongen het meest duidelijk naar voren, maar opgemerkt moet worden dat zij slechts een greep waren uit het totale divisie-aanbod op de ORD.

De grotere verscheidenheid en het groeiend aantal bijdragen mogen verheugend genoemd worden. We kunnen constateren dat de divisie 'Lerarenopleiding en Leraarsgedrag' een erg actieve is. We moeten echter niet te vroeg in euforie uitbarsten, denk ik. Ten eerste zeggen deze gegevens natuurlijk weinig over hoe het er op de ORD in Twente aan toe zal gaan. Daarnaast kunnen de hoge aantallen deelnemers, bezoekers en bijdragen natuurlijk ook hebben samengehangen met de komst van een tweetal publiekstrekkingen voor de divisie, namelijk Feiman-Nemser en Little.

Mentoring was gedurende het onderzoeksjaar 1996-1997 het centrale thema van de divisie en stond reeds centraal in de diverse divisiebijeenkomsten die gedurende het jaar gehouden werden. De lezing van Feiman-Nemser was bedoeld als de climax waarin dit thema-jaar zijn einde zou vinden. Binnen haar lezing ging Feiman-Nemser in op twee vormen van mentoring, 'educative mentoring', waarbij professionele ontwikkeling nagestreefd wordt, en 'conventional mentoring', gericht op het verlenen van steun. Vervolgens gaf zij enkele kenmerken aan die men kan gebruiken om mentoring te beschrijven en te onderzoeken. Aan de hand van drie 'vignets' (beschrijvingen van de situatie tussen mentor en lerende) ging zij in op wat volgens haar de beste manier van mentoring was. Haar verhaal was doorspekt met voorbeelden en in die zin ook erg illustratief, maar inhoudelijk nogal zwak, vond ik. Het was vooral een beschrijving, geen wetenschappelijke beschouwing. Daarbij wekte zij sterk de indruk dat veel van 'haar vondsten' rechtstreeks vanuit de praktijk kwamen, waarbij haar theorie (en ook de rest van haar verhaal) gebaseerd bleek op slechts drie cases. Het leek erop alsof zij vooral gezien wilde worden als iemand die in de praktijk staat, maar vooral niet als een vooraanstaand wetenschapper. Jammer,

want de ORD is in de eerste plaats toch een onderzoekscongres.

Het onderzoek naar (na)scholing van docenten liet ook enkele opvallende zaken zien. Allereerst lijkt er onder onderzoekers steeds meer belang te worden gehecht aan de overgang van aanbod-gerichte nascholing naar meer vraag-gerichte nascholing. Het onderzoek van Pompe en Kwakman (KUN) viel binnen dat kader, evenals het symposium 'Nascholing in Vlaanderen en Nederland' (Mahieu, Van Couwenberghe, Kerkhoff, Raaijmakers, Biesta en Van Dieten, allen IVA, KUB).

Onderzoeken naar de effecten van nascholing, zowel dat van de Nijmeegse Lerarenopleiding UNILO/KUN: Mathijssen, Bergen, Derksen, Lamberigts, Pelkmans en Engelen als dat in samenwerking met landelijke Studiecentra (Veenman, Beems, Gerrits en Op de Weegh) lieten zien dat de resultaten van nascholingscursussen erg beperkt zijn en in eerste instantie vooral optreden in de vorm van toegenomen tevredenheid onder de deelnemers. Voor langdurige en meer constante effecten lijkt de inzet van coaching en/of mentoring van groot belang, alsmede een voldoende lange duur van de betreffende cursus. Verder blijken doelen en meetinstrumenten nogal eens niet op elkaar aan te sluiten. De discussiant besprak de mogelijkheid om kennis vanuit andere divisies nadrukkelijker in te brengen, bijvoorbeeld meer ontwerpgerichte noties.

Veel aandacht was er ook voor cognities en denkprocessen van docenten. Naast de papersessie van meer ervaren onderzoekers en opleiders viel het symposium van beginnende onderzoekers: Longayroux (ICLON/RUL), Van Veen (UNILO/KUN), Den Brok en Zanting (ICLON/RUL) op door zijn opzet: het uitengaan in subgroepen om vervolgens plenair de resultaten van iedere subgroep te bespreken bleek de interactiviteit tussen presentatoren en publiek sterk te vergroten en zorgde voor veel feedback naar de onderzoekers toe. Het symposium was daarom in alle opzichten een groot succes: wellicht een suggestie voor de andere deelnemers?

Het lijkt erop, dat onderzoek op het gebied van cognities te maken krijgt met een steeds groter wordende diversiteit aan subthematieken en onderzoeksmethoden. Een mogelijk probleem voor de toekomst zal de integratie

van de resultaten van al deze onderzoeken zijn. Voor de minder ervaren onderzoekers speelt daarnaast ook het probleem van terreinafbakening, zo bleek tijdens de discussies.

Tenslotte was er op deze ORD veel aandacht voor het concept professionaliteit. Little (USA) besteedde aandacht aan de condities die nodig zijn voor professionalisering. Allereerst werd kort de geschiedenis m.b.t. vernieuwen doorgenomen. Deze kan grofweg uiteengelegd worden in twee periodes, de periode 1988-1992 waarin vernieuwingen afzonderlijk ingevoerd worden, en de periode 1990 tot heden, waarbij vernieuwen meer schoolbreed aangepakt wordt (restructuring). Als belangrijkste vernieuwingen in de VS werden, vergelijkbaar met Nederland, genoemd: vakoverstijgende curricula, meer en breder docerrepertoire, samenwerking tussen docenten en heterogene klassen. Er blijkt nog een aantal belangrijke problemen te bestaan als het gaat om implementatie van vernieuwingen; zo is er volgens Little te weinig aandacht voor de eisen die vernieuwingen aan de deelnemers, met name docenten, opleggen, blijkt er voortdurend een tendentie te bestaan om terug te vallen op traditionele strategieën, zijn de verwachtingen van betrokkenen dikwijls te hoog gespannen, verwacht men veel te snel resultaat en wordt de professionalisering van docenten (met betrekking tot bestaande vernieuwingen) dikwijls door 'nieuwe' projecten verstoord. De lezing was erg herkenbaar, hetgeen tegelijk een zwakte was. Het was bekende kost, waar je na een uur wel genoeg van had.

Naast Little's lezing was er ook een papersessie rondom professionaliteit van docenten, met, onder andere aandacht voor de professionele oriëntatie van docenten en samenwerkingsverbanden tussen docenten van het voortgezet onderwijs.

Nieuw was de aandacht voor de professionaliteit van de lerarenopleider: een belangrijk onderzoeksthema voor de toekomst?

De ORD heeft een overwegend positieve indruk op me achter gelaten. Haar belangrijkste kenmerk was in mijn ogen verscheidenheid. Deze verscheidenheid uitte zich zowel in de onderzoeksthematieken als in de deelnemers. Vreemd genoeg heb ik de publiekstrekking als het meest teleurstellend ervaren, zowel inhoudelijk als qua presentatievorm. Hiermee wil ik

maar zeggen dat er in mijn ogen geen direct verband bestaat tussen de kwaliteit van een congres en de (onderzoeks)ervaringen van haar deelnemers. Een geluid dat men in de wandelingen immers ook steeds meer hoort, is dat gevestigde onderzoekers zich steeds minder zouden laten zien (en vooral horen) op de lokale congressen, zodat deze dreigen te onttaarden in promovendi-bijeenkomsten. Hoewel deze trend zich nog niet heeft gemanifesteerd binnen de divisie Lerarenopleiding en Leraarsgedrag, het aandeel in de bijdragen dat door AIO's/OIO's en studenten is geleverd is met 25% stabiel gebleven ten opzichte van eerdere jaren en ligt daarbij zeker niet te hoog, en voor zover aanwezig zeker niet heeft geleid tot een daling in kwaliteit, moet voor een homogenisering van deelnemers wel gewaakt worden.

Hoger Onderwijs (C. Terlouw, OC, Universiteit Twente, Enschede)

Hoger Onderwijs, voor het eerst zonder Volwasseneneducatie, was dit jaar vertegenwoordigd met een keur van communicatievormen: twee fora, twee symposia, een postersessie, twee papersessies, en demonstraties in het verlengde van één van de papersessies. Tevens was er nog een plenaire lezing van Goodyear waarin het hoger onderwijs centraal stond.

Met steekwoorden als 'nieuw', 'toekomstgericht', en 'activerend onderwijs' kunnen de verschillende bijdragen worden gekarakteriseerd. In de fora stonden met name uitwerkingen centraal als gevolg van de verschuiving van docentgecentreerd naar studentgecentreerd onderwijs. De symposia richtten zich op de gehele studieloopbaan van de student: van de aankomende student, in het bijzonder het aspect studiekeuzemotivatie, tot en met de afgestudeerde waarvoor de arbeidsmarkt in het zicht komt. In de postersessie, terecht is er apart tijd voor ingeruimd, en in één van de papersessies waren allerlei nieuwe onderwijsbenaderingen en -vormen aan de orde; de demonstraties sloten hier bij twee papers aan. De andere papersessie betrof het beleidsvoorbereidend onderzoek in het HO.

Helaas kon ik door de parallele programmering en eigen optredens slechts aanwezig zijn bij de fora, één van de papersessies, en de postersessie. De bezochte fora, papersessies en

postersessie worden hier besproken, gevolgd door een korte algemene conclusie.

In het forum 'De impact van nieuwe assessmentvormen op leren en onderwijs' stond de beoordelingscultuur in het HO centraal. In dit zeer levendige forum werden twee papers gepresenteerd. (Dochy en Moerkerke, beiden OU). Na deze presentaties reageerden enkele panelleden (Dyck, Pilot en Straetmans) waarbij de reactie van de laatste eveneens een volledige paperlezing opleverde. Daarna vond een felle discussie plaats met de bezoekers van het forum.

Zoals gezegd, beide paperlezingen betroffen de beoordelingscultuur in het hoger onderwijs. Dochy schilderde het gehele palet van mogelijke beoordelingsvormen dat nodig is om in te spelen op een aantal trends in de onderwijsontwikkeling van het hoger onderwijs: (i) het onderwijs staat dicht bij de werksituatie, (ii) er is meer aandacht voor het verwerven van complexe vaardigheden dan voor kennisverwerving, (iii) toetsen heeft in toenemende mate een functie tijdens het totale leerproces (volgen en ondersteuning) dan na afloop van het leerproces (summatief toetsen), en (iv) het inschakelen van studenten in het toetsproces. Dit alles maakt het nodig om van een testcultuur naar een toetscultuur te evolueren. Moerkerke gaf hier vervolgens in toegespitste zin een concrete illustratie van door in het bijzonder in te gaan op het curriculum als uitgangspunt voor het toetsen van complexe vaardigheden, en het gebruik daarin van portfolio's. Moerkerke pleit voor een curriculaire toetsingslijn naast een modulaire of cursus-toetsingslijn. In de eerste toetsingslijn gaat het met name om de individuele voortgang in de beheersing van een complexe vaardigheid vast te stellen (bijvoorbeeld: de vaardigheid in een team te werken).

Een discussiepunt dat hierbij dank zij de reactie/paperlezing van Straetmans (CITO, Arnhem) in alle levendigheid naar voren kwam, was in hoeverre we bij deze alternatieve assessmentvormen nu nog uit kunnen komen met de gebruikelijke testpsychologie waarin met name de betrouwbaarheid van de meting van een individuele prestatie centraal staat. De 'voorstanders' hiervan wezen op het belang van het overeind houden van de klassieke kwaliteitsnormen van toetsen en tentamens: er worden immers belangrijke beslissingen over de

verdere studieloopbaan op grond van de uitkomsten genomen. De 'tegenstanders' vroegen om één andere of een aangepaste toetsmethodologie, omdat de klassieke werkwijzen zo langzamerhand een obstakel gaan vormen voor inhoudelijke onderwijsvernieuwing.

In het tweede forum stond het begin van dit jaar uitgekomen handboek voor docenten 'Onderwijskunde Hoger Onderwijs' centraal. Eén van de redacteurs, Hans van Hout (UvA) hield een korte inleiding waarin het centrale thema 'Zelfstandig leren in een consistente omgeving' werd uiteen gezet. In toenemende mate voeren faculteiten in het hoger onderwijs een beleid uit waarin het 'zelfstandig leren' en in het bijzonder de ontwikkeling daarvan cruciaal is. Daartoe is een consistente omgeving nodig waarin de rol van de docent kenmerkend is veranderd. De aandacht van docenten moet nu vooral uitgaan naar (i) werkvormen waarin momenten voor 'zelfstandig leren zijn ingebouwd, (ii) onderwijsontwikkeling waarin tegelijkertijd rekening wordt gehouden met het les-, cursus- en curriculumniveau, (iii) specifieke werkvormen die bij uitstek 'zelfstandig leren' kunnen stimuleren: vormen van probleemgeoriënteerd leren (PGL) en projectonderwijs (PO), en (iv) een onderwijsorganisatie gericht op grotere professionaliteit en kwaliteitszorg.

De discussiant Pilot (IVLOS, UU) had naast lof – eindelijk eens wat zaken systematisch op een rij vanuit een duidelijke visie op het hoger onderwijs – ook kritiek op de pretenties van de auteurs. De kritiek en de discussie spitste zich toe op de inschatting van de vaardigheden en de mogelijkheden van docenten. Het boek zal zeker zijn nut hebben voor docenten die op enigerlei wijze bij onderwijsontwikkelingsprojecten zijn betrokken en/of een docentencursus volgen bij één van de RWO-centra; een individuele bestudering van een 'losse' docent vraagt teveel naar het oordeel van sommige aanwezigen, ook al gelet op de theoretische toonzetting. Wat dat betreft had het boek eigenlijk nog één slag extra nodig gehad om nog meer aan te sluiten bij de concrete problemen van docenten.

In de papersessie 'Nieuwe onderwijsbenaderingen in het hoger onderwijs' kwam een viertal papers aan de orde waarin op enigerlei wijze een relatie bestond met nieuwe onder-

wijsvormen: een simulatie, courseware, projectonderwijs, en probleemgestuurd leren.

Coetsier en Lievens (UG) rapporteerden over het ontwikkelen van een video-simulatie voor de evaluatie en selectie van kandidaat medische studenten. De achterliggende gedachte is dat studenten beter kunnen worden geselecteerd als ook een aantal representatieve, realistische situaties, waarover zij vragen moeten beantwoorden, onderdeel uitmaakt van de toelatingsproef. Coetsiers en Lievens vertelden op welke wijze zij stapsgewijs de videofragmenten van dokter-patiënt gesprekken en een teamvergadering hadden ontwikkeld en gevalideerd. De discussie spitste zich toe op twee vragen: (a) hoe kunnen beginnende studenten een goed antwoord geven op dergelijke vragen als ze het nog moeten leren?; en (b) zou dit op zich mooie materiaal niet veel beter gewoon in het onderwijs kunnen worden gebruikt, inclusief als voortgangstoets (zie Dochy en Moerkerke hiervoor), in plaats van in een toelatingstoets? Naast deze videosimulatie presenteerden Lievens en Coetsier in een demonstratiesessie ook andere videosimulaties: één voor personeelselectie en één voor screening en training van beoordelaars.

Proost en Elen (KUL) presenteerden onderzoek en ontwikkeling (althans de eerste stap) van een 'Electronic Performance Support System' (EPSS), een elektronisch hulpmiddel voor docenten dat hen moet aanmoedigen en ondersteunen bij het gebruik van technologieën in het onderwijs. Het ging hier in het bijzonder om DigIT (Digitale Informatie Technologie), een computergebaseerd informatiesysteem (gegevensbank) in hypertextformaat, ter ondersteuning van het gebruik, ontwikkelen, en evalueren van didactische programmatuur. DigIT is in een aantal versies ontworpen waarbij in de werkwijze het betrekken van de doelgroep van docenten als cruciaal werd gezien. Bij wijze van veldtest zal de nu bestaande voorlopige versie via het lokale universitaire netwerk in Leuven (KULnet) vrijblijvend ter beschikking worden gesteld (zie <http://www.cs.kuleuven.ac.be/~koeinh/digit>). In de discussie werden twijfels geuit of een dergelijke vrijblijvende opzet nu wel tot resultaat zal leiden. Voorts vroeg men zich af in hoeverre docenten bereid zijn, en de tijd kunnen vinden, om op een dergelijke rationele wijze de vorm-

geving van het onderwijs te ontwerpen. Ondanks de wellicht kritische opmerkingen werd de ontwikkeling van een dergelijk 'job aid' voor docenten toch zeer interessant gevonden. In een demonstratie werd het programma DigIT ook getoond. De gegevensbank bevat verschillende soorten van computergebruik in een onderwijssetting, ontwikkelingsstappen, criteria voor evaluatie van software, programma-voorbeelden, en verwijzingen naar personen, literatuur en websites.

Perez Salgado, Terlouw en Kramers-Pals (UT) rapporteerden de resultaten van een exploratief onderzoek naar de invloed van geslacht op leereffecten en welbevinden in groepsprocessen in projectonderwijs in chemische technologie. De eerste onderzoeksvraag is of gedifferentieerd tutorgedrag, waarbij de tutor inspeelt op coöperatief/competitief/individueel (groeps)gedrag effect heeft op de aard van het groepsgedrag en de leerresultaten van de vrouwelijke en mannelijke gesocialiseerde studenten (gender). De tweede en derde onderzoeksvraag is of coöperatief leergedrag een positief effect heeft op respectievelijk de leerresultaten en het welbevinden van de vrouwelijke en mannelijke gesocialiseerde studenten. De resultaten betreffen vooralsnog een onderdeel van de eerste vraag, en zijn samengevat als volgt: (1) het groepsgedrag is coöperatiever naarmate het aandeel vrouwen (0% - 33%) hoger is; er is hier geen duidelijke relatie met gender; (2) de groepsprocessen in de projectgroepen zijn competitiever naarmate het aandeel mannelijke studenten (67% - 100%) hoger is, en wanneer de overall gender-typeering 'mannelijk' is; en (3) in aanwezigheid van de tutor gedragen de projectgroepen zich coöperatiever. In alle gevallen was de tutor zelf overwegend coöperatief. De discussie spitste zich met name toe op de plaats en rol van meisjes in technische studies en hun perceptie daarvan, en de mogelijkheden van een docent(e)/tutor om hier invloed uit te oefenen. De discussie hierover is zeker nog niet afgerond.

Van den Hurk, Wolfhagen, Dolmans en Van der Vleuten (UM) gingen na in hoeverre de door studenten geformuleerde leerdoelen in Probleemgestuurd Onderwijs (PGO) nu ook daadwerkelijk een leidraad vormden bij de zelfstudie, en in welke mate dit positief inwerkte op de bestede tijd aan zelfstudie en op de

leerprestaties. Daartoe werd een vragenlijst geconstrueerd en voorgelegd aan 725 studenten Geneeskunde uit de eerste vier jaar van de universiteit van Maastricht. De resultaten: (i) vooral in het eerste jaar laten studenten zich leiden door de geformuleerde leerdoelen, in de latere jaren neemt het af; (ii) er zijn geen verschillen in tijdbesteding, met die nuance, dat dit wel het geval was bij leerdoeloverstijgende studieactiviteiten; en (iii) er zijn ook geen verschillen in scores op de bloktoets. In de discussie werd er op gewezen dat eigenlijk toch wel een vrij centraal leerstuk van PGO, een zelfgeformuleerd leerdoel als leidraad voor zelfstudie, ter discussie wordt gesteld. De aanwezigen kwamen er in de beschikbare discussietijd niet uit wat dit nu betekende voor PGO in het algemeen, en de stap van het 'leerdoel formuleren' in het bijzonder.

Er werden twee posters gepresenteerd. Oosterhuis-Geers (UT) stelde twee wijzen van integratie van studievaardigheden aan de orde: (a) integratie van studievaardigheden in het traditionele onderwijs, en (b) integratie van managementvaardigheden in projectonderwijs. In het laatste geval vormen de 'klassieke studievaardigheden' een deelverzameling van de managementvaardigheden. De poster van Oosterhuis-Geers richtte zich met name op de tweede vorm van integratie waarbij één en ander concreet is toegepast in de studierichting BedrijfsInformatieTechnologie (BIT) van de Universiteit Twente. Op de poster en in de gesprekken kwamen aan de orde de gehanteerde onderwijstheorie, de opzet van het onderwijs, en de positieve resultaten van het onderwijs. Van belang is de rol van de begeleiding tijdens het groepswerk: studenten die democratisch en stimulerend worden begeleid, leren meer managementvaardigheden aan dan studenten bij wie sprake was van een 'laissez faire' begeleiding.

Van Dijk, Van den Berg en Van Keulen (TUD) rapporteerden per poster een exploratief onderzoek onder TU-Delft docenten naar activering van studenten tijdens hoorcolleges: de wijze van activering, de frequentie, de tevredenheid van docenten over de huidige manier van college geven en de mogelijke oorzaken van de eventuele discrepanties. Er is een vragenlijst opgesteld en afgenomen onder 1206 TU-docenten. De resultaten dienen als basis

voor aanbevelingen voor de ontwikkeling van een professionaliseringsmodel dat in docenten-trainingen kan worden gebruikt. De poster-sessie werd druk bezocht, en er was sprake van een levendige directe discussie tussen de aanwezigen.

Tenslotte een aantal korte conclusies. Alle bijeenkomsten waren druk bezet, en er was veelal voldoende tijd voor discussie. Het was jammer dat de bijeenkomsten voor het thema hoger onderwijs (als enige) parallel waren geprogrammeerd, hetgeen het bijwonen van een aantal interessante presentaties verhinderde. Verheugend was het grote aantal Vlamingen op het congres. Wat mij betreft had er zelfs wel meer gebruik gemaakt van kunnen worden door Vlaamse en Nederlandse onderzoeksprojecten systematisch naast elkaar te zetten (bijvoorbeeld de beide symposia over studieloopbanen) Ik heb de indruk dat we in het hoger onderwijs niet veel van elkaars werk en achtergronden weten. Helaas blijft de conclusie van de kroniekschrijver van 1996 (C. U. Sikke-ma) nog overeind: Iedereen is met aparte onderzoeken bezig waardoor een onsamenhangend beeld ontstaat; er ontbreekt een grote lijn. Wellicht een opdracht voor de volgende ORD.

Beleid en Organisatie in het Onderwijs (R. Maslowski, TO, Universiteit Twente, Enschede)

Rond het thema 'Beleid en Organisatie in het Onderwijs' had de organisatiecommissie dit jaar een symposium, een forum en drie papersessies geprogrammeerd, waarin uiteenlopende onderwerpen aan bod kwamen. Het symposium had als thema 'Lokaal schoolbeleid en interne machtsverhoudingen'. In het forum stond de verhouding tussen onderwijsonderzoek, onderwijsbeleid en praktijk in Nederland en Vlaanderen centraal. In de papersessies werd verslag gedaan van onderzoek op het terrein van schoolcultuur, schoolverandering en beleidseffectiviteit.

De papersessie 'Beschrijvingsinstrumenten schoolcultuur' stond in het teken van vragenlijsten voor het meten van schoolcultuur. Maslowski (UT) deed verslag van een analyse van instrumenten voor het meten van organisatiecultuur in het basis- en voortgezet onderwijs. Daarin kwam naar voren dat deze aanzienlijk van elkaar verschillen waar het de operationali-

sering van cultuurdimensies en de wijze van formulering betreft. Opvallend was echter dat, desondanks, de professionaliteit van leerkrachten, en de consequenties hiervan voor autonomie en samenwerking, in nagenoeg alle instrumenten een belangrijke plaats bleek in te nemen. De professionaliteit van leerkrachten stond ook centraal in het instrument dat Van Petegem (UA) presenteerde. Hij ontwikkelde elf woordparen, zoals 'leerlinggericht versus leefstofgericht' en 'collegiaal versus rivaliserend', aan de hand waarvan scholen hun eigen profiel in kaart kunnen brengen. Hendriks en Bosker (UT) presenteerden een instrumentarium, 'School en klaskenmerken', dat in het kader van zelfevaluatie door basisscholen kan worden gebruikt. Het instrumentarium bestaat uit drie afzonderlijke vragenlijsten voor respectievelijk de schoolleiding, de leerkrachten en de leerlingen. Met behulp van de vragenlijsten kan het onderwijsaanbod van een school in kaart worden gebracht, en kunnen scholen inzicht krijgen in een aantal relevante procesfactoren op school- en klasniveau, zoals de mate van samenwerking tussen leerkrachten en de effectieve leertijd in de klas.

In de papersessie 'Schoolverandering' deed Klerks (UU) verslag van haar onderzoek naar veranderingscapaciteiten van scholen met het oog op de invoering van de basisvorming. Klerks gaf aan dat de invoering van de basisvorming aanleiding heeft gegeven tot zowel onderwijsinhoudelijke als organisatorische veranderingen in scholen. Uit onderzoek blijkt dat het fundament voor de basisvorming weliswaar is gelegd, maar dat het verder vormgeven van organisatorische en met name onderwijsinhoudelijke veranderingen nog een proces van jaren vergt. Voor een succesvol verloop van dit veranderingsproces is het van belang, zo gaf Klerks aan, dat scholen een duidelijke visie hebben ten aanzien van de te bereiken doelen, het veranderingsproces structureren, en dat leerkrachten en leden van de schoolleiding een lerende houding aannemen. Maes (KUL) maakte een vergelijkbaar onderscheid in zijn onderzoek naar het vermogen van basisscholen om doeltreffend leerfingen met onderwijsachterstanden te helpen. Hij ontwikkelde een vragenlijst, 'Zorg voor leerlingen met onderwijsleerproblemen', die aangeeft in hoeverre een school in staat is problemen bij leerlingen

te signaleren, deze doeltreffend te analyseren of te diagnostiseren, en de geanalyseerde problemen te verhelpen. Daarbij wordt zowel aandacht gegeven aan de achterliggende visie van de leerkracht, het schoolteam en de schoolleider, als aan concrete activiteiten die in school worden genomen om leerlingen met leerproblemen te helpen. Blees-Booij (ROA, Amsterdam) deed verslag van een onderzoek naar de maatregelen die gemeenten hebben genomen met het oog op de aanstaande decentralisatie van het onderwijsachterstandsbeleid. Uit het onderzoek bleek dat een relatief groot aantal gemeenten reeds begonnen is een achterstandsplan op te stellen. Niettemin moest in vrijwel alle gemeenten nog veel werk worden verricht, ten aanzien van zowel de inhoudelijke uitwerking van de plannen, als het overleg met schoolbesturen en instanties binnen en buiten de gemeentelijke organisatie.

Het onderzoek dat in de sessie 'Beleids-effectiviteit' werd gepresenteerd sloot hier nauw op aan. Pijl (RUG) deed verslag van een onderzoek naar het effect van het WSNS-beleid op de deelname aan LOM en MLK. Uit het onderzoek bleek dat de deelname, instroom en verblijfsduur voor de schooltypen LOM en MLK sinds de invoering van het WSNS-beleid niet verder is gegroeid, of minder is gegroeid dan in de jaren ervoor. Niettemin kunnen deze ontwikkelingen niet geheel worden toegeschreven aan het WSNS-beleid. De reden daarvoor is, zo gaf Pijl aan, dat de groei van het LOM al voor de implementatie van het WSNS-beleid was afgenomen en in het MLK de eerste tekenen van een ombuiging tegelijk met de eerste concrete maatregelen in WSNS-verband zichtbaar werden. Vanuit de zaal werd geopperd dat voor de invoering van het WSNS-beleid het overheidsbeleid ook al gericht was op terugdringing van het aantal leerlingen in het speciaal onderwijs. Dat zou kunnen betekenen dat het geheel aan beleidsprogramma's op dit terrein wel degelijk haar vruchten heeft afgeworpen.

Hellings en Sels (KUL) presenteerden aansluitend een onderzoek naar de loopbanen van leerkrachten in het Vlaams secundair onderwijs. Daarin kwam naar voren dat er op veel scholen sprake is van taakintegratie. Leerkrachten krijgen in toenemende mate naast hun onderwijzende taak ook beleidsmatige taken

toebeedeeld. Tegelijkertijd bleek dit echter niet te leiden tot meer tevredenheid onder leerkrachten over hun werk op school. Hellings en Sels opperden dat de reden hiervoor waarschijnlijk ligt in de taakverzwaring die veelal met taakverbreding gepaard gaat. Leerkrachten die niet-lesgebonden taken op zich namen bleken namelijk slechts in een beperkt aantal gevallen een gedeeltelijke vrijstelling van hun lestaak te krijgen.

Luyten (UT) presenteerde tot slot een onderzoek naar verschillen in wiskundeprestaties tussen parallelklassen, dat hij samen met De Jong (RUG) had verricht. Uit hun analyse kwam naar voren dat de verschillen tussen klassen en leerkrachten aanzienlijk kleiner waren dan de verschillen tussen scholen. In veel opzichten bleek de autonomie van de leerkrachten ingeperkt te worden. De inhoud en de doelen van het onderwijs, evenals de inhoud, de frequentie en beoordeling van proefwerken werden gezamenlijk vastgesteld. Luyten en De Jong concludeerden dat deze afstemming waarschijnlijk leerkrachten van parallelklassen aanzet om eenzelfde niveau te bereiken met hun leerlingen. Tegelijkertijd bleek echter dat tussen leerkrachten van parallelklassen, wat betreft hun stijl van lesgeven, wel grote verschillen bestonden. Dat wekt de indruk dat de leerstof die aangeboden wordt van groter belang is voor de prestaties van leerlingen dan de manier waarop deze wordt aangeboden. Al met al een interessante vraag voor vervolgonderzoek, en voor een presentatie op de volgende ORD.

Beroeps- en bedrijfsopleidingen en volwasseneducatie (S.J. van Zolingen, TO, Universiteit Twente, Enschede)

De organisatie had de presentaties en symposia ondergebracht in drie papersessies, één postersessie en een forum waarin werd gediscussieerd over de relatie tussen onderzoek en praktijk van bedrijfsopleidingen.

De eerste papersessie 'Beroepsonderwijs' was erg overladen met vijf papersessies die in twee uur gepresenteerd en bediscussieerd moesten worden. Dit laatste schoot er dan ook helemaal bij in en dit verminderde de toegevoegde waarde voor de deelnemers en presentatoren aanzienlijk. Papers op het gebied van de

volwasseneneducatie waren duidelijk ondervertegenwoordigd.

De papers van Onstenk (UvA) en Van Zolingen, Streumer en Nijhof (UT) betroffen beide de wijze waarop het beroeps onderwijs in kan spelen op de vele en snelle veranderingen die vanuit het arbeidssysteem op het beroeps onderwijs afkomen. In beide papers werd de rol die kernproblemen kunnen spelen in het beroeps onderwijs nader onderzocht. Van Zolingen e.a. deden een empirisch onderzoek waarin kernproblemen werden geoperationaliseerd in de autoreparatiesector voor de beroepen servicemanager en diagnosetechnicus. Zij interviewde 30 beroepsbeoefenaren in 20 garages. Van Zolingen e.a. legden vooral nadruk op de spanningsvelden zoals tijd versus kwaliteit; belangen van het bedrijf versus belangen van de klant; korte termijn denken versus lange termijn denken waarin deze kernproblemen zich afspelen en de sleutelkwalificaties waarover de servicemanager en de diagnosetechnicus moeten beschikken om ze op te lossen. Daarnaast waren de paperauteurs van mening dat de gedachte om kernproblemen op te nemen in de eindtermen van het beroeps onderwijs, om beroepen meer te 'contextualiseren', bijvoorbeeld door herinrichting van de huidige eindtermen tot de reële mogelijkheden behoort. Onstenk ging dieper in op de benutting van kernproblemen om het beroeps onderwijs te verbeteren. Het gaat daarbij niet direct om het oplossen van, maar om het leren omgaan met kernproblemen. Per beroep kan een karakteristieke set problemen worden onderscheiden die is opgebouwd uit een mengsel van: productieproblemen, regelproblemen, organisatieproblemen, lidmaatschapsproblemen, sociaal culturele problemen en samenwerkingproblemen. Onstenk onderzocht op basis van eindtermendocumenten en literatuur o.a. de zorgsector. De zorgsector maakt grote veranderingen door. Een typisch kernprobleem bleek dat door tijdsdruk en verzakelijking de verplegenden en verzorgenden in een dilemma terechtkomen waarin hun uitgangspunt om de belangen van de cliënt centraal te stellen soms haaks staan op de tijdsdruk waarin ze dagelijks moeten functioneren. Onstenk benadrukte dat ondanks het feit dat onderzoek heeft aangetoond, dat er een reëel spanningsveld bestaat tussen doelmatigheid en een humaan verzor-

gingsmodel, dit kernprobleem niet expliciet wordt gethematiseerd in de bestaande eindtermen.

Frietman (KUN) presenteerde onderzoek (uitgevoerd in samenwerking met Onstenk), waarin het lage rendement (50%) van het leerlingwezen in verband werd gebracht met regionale netwerken in het leerlingwezen. Frietman en Onstenk onderzochten de vraag hoe het leerlingwezen op regionaal niveau is ingebed in netwerken van bedrijfsleven, consultant, RBA en streekscholen en welke effecten dit heeft op de kwaliteitsbewaking en het rendement van het leerlingwezen. Zij concludeerden dat de betekenis van regionale netwerken evident is en met name tot uiting komt in de kwaliteit van de begeleidingsstructuur i.c. het samenspel van de betrokken actoren; de kwaliteit van de verschillende typen netwerken afhankelijk is van de kwaliteit van de afzonderlijke en gezamenlijke actoren; puur leerlinggebonden netwerken kwetsbaarder zijn dan andere typen netwerken; leerlingoverstijgende netwerken continuïteit bieden (mits actief), maar gemakkelijk een sluimerend bestaan kunnen leiden; geïnstitutionaliseerde netwerken (goa's) in het algemeen uitzicht bieden op een hoog rendement, maar er ook sprake is van strenge selectie 'aan de poort'. Een essentiële uitkomst was dat de feitelijke rol en effectiviteit van regionale netwerken niet of zeer moeilijk is aan te tonen. Dit roept de vraag op naar de zin van dit onderzoek.

Van Batenburg (RUG) presenteerde onderzoeksresultaten over de loopbaan van voortijdige schoolverlaters na het MBO (VSV'ers). Hij deed dit aan de hand van een eindeloze reeks tabellen die hij vlot behandelde zonder lang stil te staan bij verklaringen van feiten en cijfers, waarmee hij zijn toehoorders mijns inziens tekort deed. Van Batenburg onderzocht het MBO-cohort dat in 1987 zijn opleiding aanving. Bij de VSV'ers nam hij twee vragenlijsten af, één ruim een jaar en één ruim drie jaar na schoolverlaten. De belangrijkste genoemde reden voor schoolverlaten was zittenblijven. De zittenblijvers ondervinden waarschijnlijk leerproblemen kunnen in een negatieve spiraal terecht komen, aldus een door Van Batenburg aangehaalde verklaring van Finn (1994). Een andere belangrijke reden was ontevredenheid over de sfeer op school. Ook dit

feit past binnen de conceptualisering van Finn, waarbij de school niet past binnen het zelfbeeld van VSV'ers, aldus Van Batenburg. Werkloosheid komt nauwelijks voor bij de VSV'ers (slechts 8%). De arbeidsmarkt biedt VSV'ers uit het MBO blijkbaar voldoende mogelijkheden tot mobiliteit: 42% van de VSV'ers heeft ruim drie jaar na schoolverlaten werk op middelbaar niveau gevonden, hetgeen hoger is dan het niveau van hun vooropleiding. Hierbij valt op de VSV'ers die langer in het onderwijsstelsel zijn gebleven (van opleiding veranderd) op een hoger niveau werken. Dit is een leuke uitkomst die de human capital theorie bevestigt. Een verontrustend punt is de lagere mobiliteit van de vrouwelijke VSV'ers. Van Batenburg beveelt aan voortijdig schoolverlaten tegen te gaan door 'at risk' leerlingen binnen ROC's tijdig in een ander traject te plaatsen om te voorkomen dat ze blijven zitten.

Tenslotte rapporteerde Sandra van de (UA), van oorsprong een linguïst, over leren lezen in de basiseducatie. Hij was een buitenbeentje in dit gezelschap: de enige Belg in dit thema en de enige met een onderwerp uit de volwasseneneducatie. Hij presenteerde sober en legde helder en duidelijk uit dat het er niet om gaat welke leesmethode wordt toegepast in de basiseducatie, maar dat de juiste vraag is 'wanneer moet je welke methode toepassen (vorm vs betekenis)?' In deze keuze speelt volgens Sandra een combinatie van factoren zoals, affectie, cognitie, angst en motivatie een doorslaggevende rol.

De tweede papersessie 'Bedrijfsopleidingen' was eveneens met vijf presentaties overladen en dit bleef niet zonder gevolgen. De gezamenlijke lezing van Warmerdam (KUN) en Thijssen (UU) viel voor de helft in het water, mede doordat de voorzitter zich aan het tijdschema hield. De toehoorders misten hierdoor de presentatie van de resultaten door Thijssen en dat was erg jammer. Een voordeel was dat als gevolg van deze strakke tijdsbewaking er wel tijd was om te discussiëren aan het einde van de presentaties.

Gielen (STOAS, Wageningen) vertelde over de achtergrond van onderzoek naar het faciliteren van sectorale innovatie. Zij concludeerde op basis van drie case studies in de glas-tuinbouw, de installatietechniek en de broodbakkerij, dat het reguliere beroepsonderwijs in

het innovatieproces van een bedrijfstak geen grote rol speelt. Zij is van mening dat ROC's het beste kunnen aanhaken bij bestaande netwerken; dat ROC's hun kennisbronnen gebruiken om de eigen kennis te verbeteren en niet om kennisontwikkeling in bedrijven te faciliteren en dat ROC's alleen een continue rol spelen in de kennisinfrastructuur als ze hun eigen kennis managen.

Vermeulen, Streumer en Van der Klink (UT) onderzochten ontwikkelingen op het gebied van human resource development (HRD) tussen 1992 en 1996 aan de hand van trendartikelen in een aantal geselecteerde HRD tijdschriften. Deze artikelen waren vooral van beschouwende aard. Zij concludeerden dat het bij organisatorische ontwikkelingen met name ging om: veranderingsmanagement, vormgeving van organisaties, nadruk op teams en globalisering. Bij technologische ontwikkelingen stonden centraal: de informatiesnelweg, computernetwerken, multimediatoepassingen. Ontwikkelingen in opleiding en ontwikkeling die de aandacht trokken waren: de eigen verantwoordelijkheid van werknemers voor leren en loopbaanontwikkeling en just-in-time leren. Er was vanuit de toehoorders kritiek op de gehanteerde selectiecriteria voor de tijdschriften. Zij meenden dat door de gevolgde selectieprocedure bepaalde technologische en organisatorische ontwikkelingen onderbelicht bleven.

Warmerdam (KUN) en Thijssen (UU) vertelden over de achtergronden van onderzoek bij zes lokale banken naar organisatieveranderingen en opleidingsbeleid. De banken waren autonoom en bleken met name afhankelijk van hun voorgeschiedenis en de door hen gehanteerde management concepten (schaalgrootte, marktpositie en technologie bleken minder belangrijk) verschillende keuzes te maken in organisatie- en personeelsstrategieën. Spanningsvelden die in het opleidingsbeleid ontstonden door deze veranderingen waren volgens Warmerdam en Thijssen: opleiden van het hele personeel versus doelgroepen; het aanleren van brede basiskwalificatie versus specialistische functiegerichte kwalificering; standaardopleidingen versus kwalificering op maat; aandacht voor personeelsontwikkeling versus functiegerichte kwalificering.

Van der Heijden (UT) presenteerde een meetinstrument om individuele deskundigheid

in organisaties te kunnen traceren en evalueren. Het doel van haar onderzoek was het belichten van manieren om obsoletie van de individuele medewerker te voorkomen en mogelijkheden aan te dragen om de ontwikkeling van professionele competentie gedurende de gehele loopbaan te stimuleren. Het meetinstrument bevatte vijf verschillende aspecten of dimensies: sociale erkenning en status, kennisdimensie, metacognitieve kennisdimensie, vaardigheidsdimensie, groei en flexibiliteit, waarop experts hoog scoren. Op basis van een pilot-study met 163 items bij 233 respondenten, werkzaam op verschillende domeinen, is een vragenlijst met 78 items ontworpen. Deze is afgenomen bij 1000 respondenten werkzaam op middelbare en hogere posities. Van der Heijden verzamelde zowel gegevens bij werknemers (zelfbeoordeling) als bij de direct-leidinggevende. Vervolgens werd de multiscale-multimethode analyse toegepast om te bepalen in welke mate de schalen valide zijn voor het representeren van de vijf verschillende concepten. De resultaten bleken bevredigend.

Wognum (UT) deed onderzoek naar de invloed van contextkenmerken en opleidingsklimaat op keuzes binnen strategisch opleiden in arbeidsorganisaties. Zij presenteerde de resultaten van een survey onderzoek bij 107 voor opleiding verantwoordelijke functionarissen, werkzaam bij een grotere onderneming in de industrie of in de dienstverlening. Wognum onderscheidde in haar studie de vakgebieden automatisering en sociale vaardigheden. Dit onderscheid bleek nuttig en noodzakelijk. Organisaties blijken indien ze voor een bepaald gebied opleiden dit organisatiebreed te doen met name op het gebied van automatiseringsopleidingen. De opleidingsprogramma's zijn gericht op vertegenwoordigers van verschillende functieniveaus. De nadruk ligt op vernieuwingsgerichte opleidingen met name op het gebied van automatiseringsopleidingen. Een goede tweede vormen de verbeteringsgerichte opleidingen op het gebied van sociale vaardigheden, die veelal aan teams worden gegeven. De opleidingsfunctionarissen zijn redelijk tevreden over de effectiviteit van de opleidingen. In vervolgonderzoek zullen ook leidinggevende hierover aan het woord worden gelaten.

De derde papersessie 'Leren op de werk-

plek' was met slechts vier presentaties een verademing.

Bastiaens (UT) presenteerde de effectevaluatie van het EPSS (Electronic Performance Support System) COBRA (ComputerOndersteunende BeRoepen- en TaakAnalyse) voor opleiders/ontwikkelaars. COBRA heeft tot doel de opleider in een organisatie te ondersteunen bij het uitvoeren van een beroepen- en/of functieanalyse. Met behulp van een quasi-experimenteel design waaraan een experimentele groep (COBRA) en een controlegroep (klassikale instructie) deelnamen die beiden een opdracht (zelfstandige functieanalyse) en een leertest (vragen over reproductieve kennis van beroepen- en functieanalyse) uitvoerden en een vragenlijst invulden is onderzocht of COBRA betere werk- en leerresultaten opleverde. De resultaten waren niet eenduidig. De werkresultaten van de experimentele groep waren significant beter dan die van de controlegroep. De leerresultaten van de experimentele groep bleven echter achter bij die van de controlegroep. De experimentele groep bleek zeer weinig gebruik te maken van de informatie- en leercomponent van COBRA.

Van der Klink (UT) presenteerde resultaten van een studie naar de effectiviteit van opleiden op de werkplek. Hij verzamelde gegevens bij 68 cursisten in zeven regionale call-centres van de PTT over een werkplekopleiding voor telemarketeers. Naast het verstrekken van informatie is het de taak van telemarketeers de klanten te stimuleren tot de aanschaf van diensten en producten. Bij aanvang, afsluiting en 12-18 weken na het volgen van de opleiding werd de effectiviteit gemeten aan de hand van progressie in verkoop (gemeten met de variabelen: verkoopgesprek, klachtenafhandeling, gespreksvaardigheden) met behulp van vragenlijsten die werden ingevuld door de opleider, de manager en de cursist zelf. De opleiding voor telemarketeers bleek inderdaad tot verbetering in het verkoopgedrag van de telemarketeers te leiden en na afronding van de opleiding werd geen regressie in dit gedrag geobserveerd. Wel zijn opleiders van mening dat er een grotere leerwinst is opgetreden dan de cursisten zelf. Het is mogelijk dat sociale wenselijkheid een rol speelt in het oordeel van de opleiders. Ook is het mogelijk dat cursisten bij aanvang van de cursus hun eigen 'kunnen' te hoog inschatten.

Feedback van de manager en van collega's bleken een positieve invloed te hebben op de leerwinst. Een interessant resultaat was dat vooropleiding en zelfvertrouwen een positief effect hebben op transfer en opleidingsvariabelen, zoals kwaliteit en uitvoering, niet.

Poell (KUN) presenteerde een studie waarin hij de invloed van de strategieën van actoren vergeleek met de invloed van het type werk op leerprojecten. Hij omschrijft een leerproject als een instrument om systematisch informele incidentele en formele leeractiviteiten van een groep werknemers rond een probleem dat zich in hun werk voordoet weer te geven. Poell voerde een multiple case study uit die bestond uit vier leerprojecten, Lab (in een chemisch laboratorium), Care (op een verpleegafdeling), General (bij de algemene vakken) en Dutch (bij het vak Nederlands als tweede taal). Hij concludeerde dat de 'actorstrategieën' meer invloed hebben op een leerproject dan het type werk en dat de verschillende actoren, zoals managers, trainers en werknemers, verschillende strategieën hanteren op grond van uiteenlopende belangen.

Naast de paperpresentaties werd een forum gehouden waarin de relatie tussen onderzoek en praktijk van bedrijfsopleidingen werd geproblematiseerd door Baart, Thijssen, Mulder en Lowyck. De traditionele kloof tussen onderzoek en praktijk werd als eerste genoemd. De wetenschap is met name geïnteresseerd in de groei van theoretische kennis en de praktijk heeft meer belangstelling voor onder-

zoek dat resultaten oplevert die direct op de praktijk kunnen worden toegepast. De forumleden constateren tegenwoordig echter een versnelde vervreemding die onder andere veroorzaakt wordt doordat alleen publiceren in internationale tijdschriften beloond wordt, bedrijven vooral geld steken in hun primaire proces en HRM activiteiten gedecentraliseerd worden naar de lijn. Daarnaast neemt de professionalisering van bedrijfsopleiders toe. Een ander punt is dat de hoeveelheid geld voor het eerste geldstroom onderzoek afneemt waardoor er meer contractonderzoek wordt gedaan. Dit heeft het grote nadeel dat het onderzoek sterk gefragmenteerd wordt. Ook wordt er veel te weinig onderzoek gedaan naar het proces van opleiden. Alleen organisatieadviesbureaus doen hier iets aan, maar anderen hebben hier geen direct belang bij dus gebeurt het niet. De commitment van bedrijven is minimaal. Toch heeft het wetenschappelijk onderzoek op dit gebied, dat overigens nog heel jong is, de praktijk wel degelijk wat te bieden, maar onnodig complex taalgebruik schrikt de praktijk af.

Aan deze kroniek werd meegewerkt door:

P. den Brok, J.F.M. de Fraiture, R. Maslowski, E.R. Savelsberg, J.N. Streumer, C. Terlouw, A.E. van der Valk, J.I.A. Visscher, M.J. Vollebregt, H.J. Vos, S.J. van Zolingen

De eindredactie werd verzorgd door

J.N. Streumer

Boekbesprekingen

B.A. Tabachnick & L.S. Fidell

Using multivariate statistics (3rd ed.)

Harper Collins, New York 1996, 880 pagina's inclusief appendices, literatuur, index en data-disk, ISBN 0 673 99414 7

J.F. Hair, Jr., R.E. Anderson, R.L. Tatham & W.C. Black

Multivariate data analysis with readings (4th ed.)

Prentice-Hall, London 1995, 757 pagina's inclusief appendices, literatuur en index, ISBN 0 13 180969 5

J. Stevens

Applied multivariate statistics for the social sciences (2nd ed.)

Lawrence Erlbaum Associates, Hillsdale, NJ 1992, 629 pagina's inclusief appendices, literatuur en index, ISBN 0 8058 1197 4

"Welk boek over multivariate technieken zou U aanraden aan promovendi en onderzoekers in de sociale wetenschappen, in aanmerking nemend dat het ideale boek enerzijds statistisch correct en up-to-date moet zijn, en anderzijds voor inhoudelijk georiënteerde onderzoekers goed leesbaar moet zijn en bruikbare adviezen moet bevatten?"

Een vraag van deze strekking heb ik ongeveer een jaar geleden toegezonden aan de Internet discussielijst StatList, een discussielijst voor een bonte verzameling statistici, met als gemeenschappelijk kenmerk dat zij zich allen bezig houden met statistisch advieswerk. Zoals gebruikelijk bij dit soort vragen was het resultaat een lange *thread* (discussielijn) waarin een groot aantal boeken werd genoemd en besproken. De drie bovenstaande boeken werden veel en in zeer positieve zin genoemd. Een goede reden om ze hier in een vergelijkende recensie te bespreken.

De sleutelwoorden in mijn vraag zijn *up-to-date* en *bruikbaar*. Er is bepaald geen gebrek aan inleidende boeken over multivariate technieken; de markt is kennelijk groot genoeg dat iedere zichzelf respecterende uitgever er minstens één in portefeuille wil hebben. Veel van deze boeken beperken zich echter tot het

bespreken van een beperkt aantal klassieke multivariate technieken (een bekend rijtje is variantie analyse, regressie analyse, factor analyse, discriminant analyse en misschien multidimensioneel schalen). De moderne multivariate analyse omvat echter een veel groter scala aan analysetechnieken, zoals iedereen kan weten die wel eens de dikke stapel SPSS handleidingen doorgebladerd heeft. Tot het arsenaal aan methoden behoren tegenwoordig ook methoden als loglineaire modellen, logistische regressie, survival analyse, structurele modellen, en niet lineaire varianten op basis van homogeniteitsanalyse (HOMALS e.d.). Een goed overzichtsboek zou in ieder geval een flink aantal van dit soort methoden moeten behandelen. Daarnaast is er een delicaat evenwicht tussen statistische volledigheid en praktische bruikbaarheid. Een gedetailleerde uitleg van de statistische achtergrond van de meeste multivariate technieken is niet mogelijk zonder matrixformules en ingewikkelde afleidingen. Anderzijds is een globale verbale omschrijving onvoldoende voor gebruikers die een verantwoorde keuze willen maken tussen de vele geboden opties, of zorgvuldig willen nagaan in hoeverre aan de statistische aannamen is voldaan. Een goed overzichtsboek moet tussen deze uitersten in laveren.

Om met de conclusie te beginnen: de hier besproken boeken zijn niet voor niets door het statistisch forum naar voren gebracht. Alle drie de boeken zijn goed leesbaar voor inhoudelijk georiënteerde onderzoekers, en bevatten veel informatie over een breed scala aan technieken. Met 'goed leesbaar' bedoel ik overigens niet 'gemakkelijk leesbaar.' Door de grote hoeveelheid informatie is de schrijfstijl doorgaans compact, zodat deze boeken niet snel lezen. De lezer moet bereid zijn zich in te spannen. Omdat multivariate analyses zonder computer ondenkbaar zijn, is het bovendien verstandig na te gaan hoe de betreffende technieken in de beschikbare programmatuur geïmplementeerd is.

Het meest recente en meest omvangrijke boek is dat van Tabachnick en Fidell. Zij behandelen variantie analyse, regressie analyse, canonische correlatie, discriminant analyse, loglineaire modellen, logistische regressie, factor

analyse en structurele modellen. De behandeling van de variantie analyse omvat multivariate variantieanalyse (MANOVA), covariantie analyse (MANCOVA) en herhaalde metingen. Voorafgaand aan de hoofdstukken over de verschillende technieken worden de diverse technieken kort ingeleid en in hun onderling verband geplaatst en wordt gedetailleerd besproken hoe de gegevens voor de analyses gescreend kunnen worden op non-normale verdelingen en uitbijters. De auteurs bespreken de mogelijkheden van verschillende computerprogramma's zoals SPSS, SYSTAT, SAS en BMDP en in het structurele modellen hoofdstuk EQS, LISREL, SYSTAT en SAS. Daarbij bespreken zij stukken computeruitvoer en laten zij in de vorm van voorbeeldteksten zien hoe de resultaten in een manuscript in APA stijl opgeschreven kunnen worden. Elke techniek wordt voorzien van een check-list met stappen die in de analyse moeten worden afgelopen. Bij het boek wordt een disk geleverd met de gebruikte voorbeeld data.

Het boek van Hair et al. bespreekt regressie analyse, discriminant analyse, logistische regressie (kort), variantie analyse, canonische correlatie, factor analyse, cluster analyse, multidimensioneel schalen, conjuncte analyse en structurele modellen. De variantie analyse omvat MANOVA en MANCOVA; analyse van herhaalde metingen wordt slechts kort genoemd. Ook hier wordt voorafgaand aan de bespreking van de diverse technieken een korte inleiding gegeven in de verzameling technieken en de benodigde data screening. Het boek is duidelijk bedoeld voor het onderwijs. Elk hoofdstuk begint met een opsomming van sleutelbegrippen die aan de orde zullen komen en eindigt met een samenvatting, opgaven en referenties. Achter elke besproken techniek volgt een hoofdstuk met een herdruk van een bestaand onderzoeksartikel waarin de betreffende techniek wordt toegepast. De onderwerpen van de toegepaste artikelen zijn eclectisch, maar komen doorgaans uit het marktonderzoek.

Het boek van Stevens bespreekt regressie analyse, variantie analyse (MANOVA, MANCOVA en herhaalde metingen), discriminant analyse, componenten analyse, canonische correlatie en loglineaire modellen. Het geheel wordt voorafgegaan door hoofdstukken over

multivariate gegevens en matrix algebra. De analyses maken gebruik van de programma's SPSSX en SAS en er worden voorbeeldsetups gegeven en computeruitvoer besproken. De meeste hoofdstukken eindigen met een aantal opgaven en het boek bevat een appendix met de antwoorden op een selectie van de opgaven.

Zoals uit de samenvattingen blijkt, verschillen de boeken enigszins in het aantal technieken die zij bespreken. Daarnaast verschillen de boeken ook in de mate van statistisch detail en presentatie. Het boek van Tabachnick en Fidell volgt een strak stramien. Alle technieken worden besproken volgens het schema: algemeen doel, soort onderzoeksvragen, beperkingen, fundamentele vergelijkingen, belangrijke punten, vergelijking van programma's en voorbeelden. Per subsectie worden de relevante punten ook weer volgens een strak schema besproken. Dat maakt het tot een goed toegankelijk naslagboek, maar het is soms vermoeiend lezen. De verschillende computerprogramma's worden uitvoerig besproken, met hele lappen uitvoer in de tekst. Gegeven de onleesbaarheid van veel computerprogramma's is dat geen slecht idee, maar het nadeel is opnieuw vermoeiend lezen en snelle veroudering. (Bij wijze van voorbeeld: T&F bespreken de DOS versies SPSS 5, SAS 6, SYSTAT 6 en BMDP 7. SPSS maakt inmiddels geen nieuwe DOS versies meer en is bij Windows aangeland bij versie 7.5, en SYSTAT en BMDP zijn door SPSS opgekocht. Ook SAS richt zich tegenwoordig vooral op Windows). Tabachnick en Fidell geven tamelijk veel statistische achtergrond. De meeste formules worden gegeven in matrixvorm (een appendix geeft een 'skimpy introduction to matrix algebra'); ter verduidelijking worden ze dikwijls vervolgens algebraïsch uitgewerkt. Veel van de statistische achtergrond valt wat mij betreft onder 'te groot voor servet, te klein voor tafellaken.' Dat wil zeggen, methodologen en statistici vinden de uitleg dikwijls nogal kort en onvolledig en inhoudelijk georiënteerde onderzoekers (lees: minder vertrouwd met statistiek en matrix algebra) slaan de statistische uitleg bij voorkeur over.

Het boek van Stevens is wat het laatste betreft duidelijk methodologisch georiënteerd.

De statistische achtergronden worden gegeven in matrix notatie zonder nadere uitleg; het opnemen van een hoofdstuk matrix algebra is zonder meer nodig. De opbouw van het boek is minder strak en de volgorde van de hoofdstukken enigszins curieus. Zo is de behandeling van variantie analyse (MANOVA, MANCOVA, herhaalde metingen) verspreid over zeven hoofdstukken, waartussen soms hoofdstukken staan over andere technieken. Hiermee is ook de belangrijkste kern van het boek gegeven: een grondige bespreking van variantie analyse, inclusief de rol van de statistische aannamen, poweranalyse, en het gebruik van geplande en post hoc vergelijkingen. De bespreking van de computerprogramma's is minder nadrukkelijk dan bij Tabachnick en Fidell; er worden voorbeeldsetups gegeven, en de opmaak van veel tabellen volgt de vorm van de computeruitvoer, maar er zijn geen bladzijden met letterlijk overgenomen uitvoer.

Het boek van Hair et al. richt zich duidelijk op gebruik in het onderwijs. Formules worden zoveel mogelijk vermeden en matrix algebra ontbreekt volledig. Er wordt wel degelijk geprobeerd de lezer statistisch begrip bij te brengen, maar de aanpak is verbaal en intuïtief, ondersteund met grafische voorbeelden. Er wordt één dataset gebruikt voor alle voorbeeldanalyses. De betreffende dataset is opgenomen in een bijlage en een andere bijlage geeft geannoteerde setups voor SPSS, SAS, BMDP, PC-MDS en LISREL 7. Net als Tabachnick en Fidell worden de verschillende technieken volgens een vast stramien gepresenteerd: definitie van het onderzoeksprobleem, opstellen analyseplan, aannamen, berekenen van schattingen en passingsmaten, interpretatie en validatie.

Wanneer we de besproken boeken vergelijken op breedte en bruikbaarheid, dan is eenvoudig vast te stellen dat de boeken van Tabachnick en Fidell en van Hair et al. de meeste technieken bespreken. Tabachnick en Fidell gaan wat ruimer in op verschillende vormen van variantie analyse en logistische regressie. Hair et al. hebben een achtergrond in het marktonderzoek, hetgeen blijkt uit hun grondige bespreking van multidimensioneel schalen en conjuncte analyse. Het boek van Tabachnick en Fidell is van wat recentere datum, wat met

name blijkt uit het gebruik van de SIMPLIS commando's bij de uitleg van LISREL; Hair et al. gebruiken daarvoor nog de meer formele LISREL matrix commando's. Stevens' boek is wat beperkter van opzet. De uitleg van de variantie analyse is echter bijzonder grondig; onderzoekers die experimentele gegevens willen analyseren zullen hier veel nuttige informatie aantreffen. Alle drie de boeken besteden ruime aandacht aan praktische analyseproblemen zoals het screenen van gegevens en het omgaan met ontbrekende gegevens. Tabachnick en Fidell volgen ook daarbij weer een strak stramien, waarbij zij snel overgaan op het uitvoeren van normaliserende transformaties en het verwijderen van uitbijters. Persoonlijk adviseer ik de lezers daar wat terughoudender mee om te gaan dan Tabachnick en Fidell.

Terugkerend naar mijn oorspronkelijke vraag kunnen we vaststellen dat de besproken boeken wat uiteenlopen wat betreft het aantal technieken dat besproken wordt. Wanneer we als operationalisering van voor de sociale wetenschappen 'belangrijke' technieken eenvoudig nemen: alles wat SPSS biedt, dan constateren we dat geen van de boeken alle technieken behandelt. Het boek van Tabachnick en Fidell bevat de meeste technieken en is goed geschikt als naslagwerk. Het is een geschikt, maar pittig, leerboek voor studenten en onderzoekers. Het boek van Hair et al. is geschikt als naslagboek en leerboek voor studenten (en onderzoekers!) die liever niet worden lastig gevallen met matrix- en andere formules. Het is wat minder grondig dan het boek van Tabachnick en Fidell, maar alle belangrijke zaken staan erin. Het boek van Stevens is met name geschikt voor onderzoekers die zich willen verdiepen in complexe toepassingen van variantie analyse.

J.J. Hox

Beelden van bekwaamheid. Een onderzoek naar de implementatie en effecten van het programma 'Responsieve Instructie' voor leerkrachten in groep 2 van de basisschool.

Academisch proefschrift Universiteit Utrecht. Utrecht, 1996, 299 pagina's, ISBN 90 393 1025 4

De overgang van groep 2 naar groep 3 in de basisschool is voor veel kinderen een moeilijke stap. Kinderen in groep 2 mogen nog naar hartelust spelen, maar in groep 3 wordt van hen verlangd dat zij allerlei opdrachten maken en zich voegen naar het lesprogramma. Sommige kinderen zijn nog niet klaar voor deze overgang; ze kunnen niet goed geconcentreerd werken en zijn niet gemotiveerd om aan de nieuwe eisen te voldoen.

In zijn proefschrift *Beelden van bekwaamheid* doet J. Castelijns verslag van een veld-experiment dat als doel had kinderen in groep 2 beter op de stap naar groep 3 voor te bereiden. De auteur vroeg leerkrachten om met relatief eenvoudige middelen ongeconcentreerde en slecht gemotiveerde kinderen te helpen. De bedoeling was om op deze wijze de interactie tussen leerkracht en leerling te stimuleren. Castelijns' onderzoek maakt deel uit van het Utrechtse Orthopedagogische Project Afstemming, dat beoogt de responsiviteit van leerkrachten ten opzichte van hun leerlingen te vergroten.

De opzet was als volgt. De onderzoeker instrueerde leerkrachten om met behulp van tien eenvoudige technieken de interactie met de slecht geconcentreerde kinderen tijdens de les te verbeteren. Zij moesten bij voorbeeld vaker bij het kind gaan zitten en met hem of haar over het werk praten, ze moesten het aankijken en afspraken maken over het werk. Op drie meetmomenten deed Castelijns observaties in de klas en registreerde hoe vaak de leerkrachten die middelen toepasten, dus hoe vaak zij bij het kind gingen zitten, met hem of haar over het werk praatten, enz. Tegelijk werden de kinderen geobserveerd en werd vastgesteld hoe zelfstandig en aandachtig zij werkten.

De eerste meting was een vóórmeting, de tweede meting vond plaats vier weken na de invoering van het programma, en de derde meting na nog eens vier weken. Op elk van deze drie tijdstippen werd de interactie in de

klas op video geregistreerd. Onmiddellijk na de vóórmeting vond de instructie aan de leerkrachten plaats, waarbij met hen werd besproken hoe zij de nagestreefde technieken konden toepassen om de interactie met de slecht geconcentreerde kinderen te verbeteren. Na vier weken werden opnieuw opnamen gemaakt. Aan de hand hiervan werd met de docent besproken in hoeverre verbetering was opgetreden. Uiteindelijk was er een derde meting. Deze opzet maakte het mogelijk het succes van het programma af te meten aan de mate waarin de leerkrachten meer met de kinderen interacteerden en de mate waarin kinderen zelfstandiger en geconcentreerder werkten.

Aan het onderzoek namen twintig leerkrachten deel, werkzaam op 10 scholen. In elke klas werden de twee kinderen geobserveerd die volgens de leerkracht de meeste concentratieproblemen hadden. Eén van die kinderen werd aangewezen als experimenteel kind: aan de leerkracht werd gevraagd om aan dit kind speciaal aandacht te schenken. Zonder dat de leerkracht hierover was ingelicht werd ook de interactie met het andere ongeconcentreerde kind geobserveerd. De bedoeling hiervan was om vast te kunnen stellen of een toename in de concentratie van het experimentele kind wel toegeschreven kan worden aan de toepassing van de interactietechnieken. Bovendien geeft deze opzet de mogelijkheid om te kijken of het responsieve gedrag van de leerkracht beperkt blijft tot het experimentele kind of dat ook andere kinderen profiteren van de invoering van het programma.

Tenslotte werd onderscheid gemaakt in twee implementatiestrategieën. Sommige leerkrachten kregen tussendoor ondersteuning in de vorm van gesprekken met een schoolbegeleider en met andere leerkrachten die het programma uitvoerden. Aan anderen werd geen tussentijdse ondersteuning geboden.

Wanneer de scores van de leerkrachten tesa-men worden genomen blijkt het programma succesvol te zijn. Ten aanzien van de meeste criteria werden de doelstellingen bereikt, dat wil zeggen dat de responsieve interactie van de leerkrachten met de experimentele kinderen toenam. Het gedrag van de leerkrachten ten

opzichte van de controlekinderen bleek echter nauwelijks veranderd.

Wat betreft de aandacht van de kinderen waren de resultaten moeilijk te interpreteren. Voor de experimentele kinderen bleek er eerst een (significante) toename van hun concentratie te zijn (dus van meetmoment één naar meetmoment twee), terwijl er daarna (van meetmoment twee naar drie) een (significante) afname was. Het omgekeerde gold voor de controle kinderen: zij gingen in eerste instantie achteruit om vervolgens weer een grotere concentratie te tonen.

Een tweede set vragen is rond de implementatiestrategie. Zijn leerkrachten gebaat bij ondersteuning tussendoor? Het antwoord luidt bevestigend. Dit resultaat ligt voor de hand, aangezien, zoals de auteur zelf aangeeft, dit soort interventies alleen op de lange termijn succesvol zijn, als ze ingebed worden in een systeem van kwaliteitszorg. Toch is het resultaat niet triviaal, want het blijkt dat leerkrachten vooral steun ondervonden van gesprekken met andere leerkrachten die in het programma waren opgenomen en niet van telefoongesprekken tussendoor met een schoolbegeleider.

Het proefschrift maakt aan het einde een boeiende wending wanneer de auteur niet langer kijkt naar groepsscores maar naar de scores van individuele leerkrachten. Met behulp van een geaggregeerde schaal van de observatiematen berekende hij de responsiviteit op elk van de drie meetmomenten voor elke leerkracht afzonderlijk. Deze 'veranderingscurven' laten het leerproces van de individuele leerkrachten over de drie meetmomenten zien. Castelijns constateert dat er in lang niet alle gevallen sprake is van het patroon dat verwacht zou kunnen worden: een geleidelijke toename van responsiviteit en het bereiken van een optimum. Dit patroon wordt wel bij een aantal leerkrachten gevonden, maar er zijn ook heel andere patronen. Sommige leerkrachten laten tussen meetmoment twee en drie juist een afname van responsief gedrag zien, terwijl anderen pas na meetmoment twee vooruit gaan. De auteur concludeert dat "de ontwikkeling van de bekwaamheid van leerkrachten een buitengewoon persoonlijke aangelegenheid is met een grillig verloop" (p. 251). Het

enige dat hierover met zekerheid is te zeggen, zo stelt Castelijns in zijn slotbeschouwing, is de onvoorspelbaarheid van de veranderingsprocessen.

Op grond van deze analyses komt Castelijns tot een kritische heroverweging van de uitgangspunten van het programma en van de manieren waarop de effecten werden gemeten. Wie de effecten van het programma wil vaststellen met behulp van groepsscores ziet over het hoofd, dat leerkrachten de programmadoelen en de tien eenvoudige technieken niet mechanisch proberen te realiseren, maar dat zij rekening houden met het gedrag van het betreffende kind. Als een kind geconcentreerder werkt zal de leerkracht het niet nodig vinden te interveniëren. In de leercurve zal dit tot uitdrukking komen in een vermindering van responsief gedrag van de leraar.

De auteur realiseert zich dat de doelstellingen van het programma niet langer in absolute termen moeten worden gevat als alleen maar gericht op méér responsief gedrag van de leerkracht, maar dat het gaat om iets relatiefs: de mate waarin de interventie van de leerkracht in overeenstemming is met de behoeften van het kind. Niet responsiviteit op zichzelf is het doel, maar *adaptieve* responsiviteit, dat wil zeggen ondersteuning die aangepast is aan het gedrag van het betreffende kind op dat moment.

De zwakste leerlingen zouden de meeste ondersteuning moeten krijgen. Dat dit feitelijk ook gebeurt blijkt uit het negatieve verband tussen bij de kinderen gemeten aandacht bij spel of werk en de responsiviteit van de leerkracht. Bij aanvang van het onderzoek was verwacht dat grotere responsiviteit van de leerkracht juist tot meer aandacht van het kind bij het werk zou leiden. Nu blijkt het omgekeerde: als kinderen geconcentreerd werken, interacteert de leerkracht minder met hen. Voor de betreffende leerkracht wordt dan natuurlijk ook een leercurve gevonden waarin geen vooruitgang zit. Dit komt niet voort uit hardleersheid van de leraar, maar uit het feit dat hij of zij in dit geval het repertoire van responsieve technieken niet hoeft aan te spreken.

Omdat Castelijns zelf tot deze herinterpretatie van zijn onderzoek komt, is des te moeilijker te begrijpen dat hij vindt dat de leercur-

ves onvoorspelbaar zijn. Met het concept 'adaptief leerkrachtgedrag' is het grillig en persoonlijk verloop van de veranderingsprocessen juist wel te verklaren.

De kritische heroverweging in het slothoofdstuk zou mijns inziens behalve tot een andere interpretatie ook moeten leiden tot een andere kijk op de onderzoeksmethode. Het is jammer dat de auteur hieraan geen aandacht meer besteedt. Castelijns' operationalisering van responsiviteit is onvoldoende: adaptieve responsiviteit valt niet met eenvoudige kwantitatieve maten te meten. Hij heeft een formule, in ieder geval niet-inhoudelijke, benadering gekozen van de interactie tussen leerkracht en leerling. Hij stelt vast of en hoe vaak en hoe lang een leerkracht aandacht geeft aan een kind, maar hij kijkt niet naar de inhoud van de interactie. Dit lijkt mij niet goed te verdedigen. Het gaat er toch niet alleen om *dat* een leerkracht reageert op een kind, maar vooral *hoe* hij of zij reageert, niet alleen om de kwantiteit maar juist ook om de kwaliteit van die interactie. De leerkrachten moeten ook in inhoudelijk opzicht adaptief zijn: de feedback moet adequaat en stimulerend zijn en aangepast aan het begripsniveau van het kind. Een leerkracht kan wel met een kind over een taak praten, maar hij of zij kan het doel om de motivatie van het kind te vergroten voorbij schieten door een veel te moeilijke opdracht te geven, het kind te overdonderen met een woordenstroom of door het kind bij een simpele opgave zo uitbundig te prijzen dat het de indruk krijgt dat het niet veel kan. Feedback en aandacht kunnen ook averechts werken, en met die mogelijkheid is geen rekening gehouden. Een verschuiving van responsiviteit naar adaptieve responsiviteit vereist dat ook de inhoud van de gesprekjes tussen leerkracht en kind in het onderzoek betrokken wordt.

Deze interessante en vakbekwaam uitgevoerde studie laat zich lezen als het verslag van het leerproces van de auteur over responsieve interactie. Iedereen die op dit terrein onderzoek doet kan hiervan leren. De kritische heroverweging die de auteur aan het einde van zijn studie presenteert maakt duidelijk hoeveel klemmen en valkuilen een onderzoeker van interactie onderweg tegenkomt. Als het gaat

om het veranderingsproces van individuele leerkrachten laat Castelijns zien dat niet één bepaald verloop het beste is: er zijn allerlei configuraties die alleen geïnterpreteerd kunnen worden door de context mee te nemen. Verbetering van de interactie tussen leerkracht en leerling laat zich niet op eenvoudige wijze vaststellen: maten die geen rekening houden met de specifieke omstandigheden laten geen conclusies toe.

Die overtuigende conclusie bergt tevens de beperking van Castelijns' programma in zich. Het is te smal, het houdt geen rekening met de specifieke context en richt zich te veel op makkelijk te observeren kenmerken, terwijl de realiteit daarvoor te complex is. Leerkrachten kunnen zeker profiteren van de in dit programma beproefde interactietechnieken. Maar ze zullen ze in combinatie met andere didactische interventiemogelijkheden aanwenden als ze inschatten dat dat nodig is.

E. Elbers

S.J. van Zolingen

Gevraagd: sleutelkwalificaties.

Een studie naar sleutelkwalificaties voor het middelbaar beroepsonderwijs

Academisch proefschrift Katholieke Universiteit Nijmegen. Universiteitsdrukkerij, Nijmegen 1995, 436 pagina's, ISBN 90 9008388 X

Het proefschrift bevat een grondige verkenning van veranderingen in arbeidsorganisaties en de consequenties daarvan in termen van kwalificaties (ofwel: wat moet iemand in het kader van een beroeps- of functieuitoefening kunnen en kennen?). Van Zolingen gebruikt daarvoor zowel buitenlandse studies als onderzoek dat in Nederland voor de Nederlandse situatie is gedaan. Ze komt tot de conclusie dat kwalificaties zoals bijvoorbeeld klantgerichtheid en zelfstandigheid belangrijk zijn binnen arbeidsorganisaties. Bewust definieert ze deze voorlopig als 'algemene kwalificaties' om verandering te voorkomen met verwante 'beladen' begrippen als extra-functionele kwalificaties of breed toepasbare beroepskwalificaties.

Uit de bestudeerde literatuur blijkt dat als gevolg van met name commerciële veranderingen, maar ook door technologische en organisatorische ontwikkelingen werkgevers in

toenemende mate vragen om flexibel inzetbaar, breed opgeleide werknemers.

Kwalificaties ziet men veelal als hét middel om de aansluiting tussen onderwijs en arbeid te optimaliseren. Vanuit die gedachte beperkt Van Zolingen zich terecht niet alleen tot veranderingen in het arbeidssysteem en de gevolgen daarvan voor kwalificaties. Ze beschrijft ook veranderingen op de arbeidsmarkt en in het secundair beroepsonderwijs. Om het kwalificatiedenken zuiver te houden is het wat mij betreft echter (nog steeds) de vraag of en zo ja hoe ver onderwijs invloed mag hebben als het gaat om het opsporen, formuleren en vaststellen van kwalificaties.

Hoe dan ook, op basis van de literatuurstudie komt Van Zolingen tot de conclusie dat "algemene kwalificaties in brede kring, door werkgevers (organisaties), door de overheid en door werknemers (organisaties) van belang worden geacht voor het goed (kunnen blijven) functioneren van werknemers" (p. 86).

Aan de hand van literatuur verkent Van Zolingen vervolgens het begrip kwalificatie. Ook hier geldt dat steeds een link wordt gelegd naar de Nederlandse situatie. Uiteindelijk blijkt het begrip 'sleutelkwalificaties' het meest adequaat te zijn. Van Zolingen kiest er dan ook voor dit concept van Mertens (1974) te actualiseren en toe te spitsen op arbeidssituaties. Ze komt tot de volgende definitie:

"Sleutelkwalificaties zijn: kennis, inzicht, vaardigheden en houdingen die behoren tot de duurzame kern van een beroep of een groep verwante functies met de mogelijkheid tot transfer naar andere en nieuwe functies binnen dat beroep en tot innovaties binnen dat beroep, die bijdragen aan de bekwaamheidsontwikkeling van een beroepsbeoefenaar en die overgangen binnen de loopbaan vergemakkelijken" (p. 117).

Deze definitie is uitputtend en sluit prima aan bij de eerder geconstateerde behoefte aan breed en flexibel inzetbare medewerkers.

Van Zolingen onderscheidt vervolgens bij sleutelkwalificaties zes dimensies, die een synthese vormen vanuit de bestudeerde literatuur: een algemeen-instrumentele, een cognitieve, een persoonlijkheids-, een sociaal-communicatieve, een sociaal-normatieve en een strategische dimensie. De criteria waarop deze

indeling is gebaseerd, zijn echter niet vermeld. Daarmee wordt het lastig om als lezer de waarde van deze onderverdeling op voorhand vast te stellen.

In het laatste deel van de literatuurstudie gaat Van Zolingen in op het verwerven en het vertalen van sleutelkwalificaties. Die aspecten worden kort belicht. Het overzicht van mogelijke didactische werkvormen lijkt logisch, maar is tegelijkertijd willekeurig. Aardig daarbij is wel dat niet alleen niet-klassikale vormen als project- en probleemgestuurd onderwijs worden besproken, maar ook vormen van klassikaal onderwijs. Tegelijkertijd neemt Van Zolingen een voorschot op het empirische deel van de studie door sleutelkwalificaties aan werkvormen te koppelen: "De Leittextmethode lijkt bij uitstek geschikt voor het verwerven van sleutelkwalificaties zoals leren leren, omgaan met informatie, probleemoplossen en plannen" (p. 136). Elk van de zes genoemde dimensies krijgt zo verspreid over werkvormen een plek.

Het vertalen van sleutelkwalificaties in eindtermen en leerplannen blijkt een harde noot om te kraken volgens Van Zolingen. Zij typeert dit vooralsnog als een onopgelost probleem, waarvoor geen directe oplossingen voorhanden zijn. Jammer, want daarin zou ongetwijfeld een meerwaarde voor onderwijskundigen hebben gelegen. Wel geeft ze richtingen aan waarin oplossingen gezocht kunnen worden. Ik ben het volledig eens met Van Zolingen dat de combinatie van kernproblemen met sleutelkwalificaties perspectief biedt, zeker gezien vanuit flexibele en brede inzetbaarheid in beroepssituaties.

Het empirisch deel van de studie begint met de bespreking van enkele varianten van de Delphi-methode. Omdat het onderzoek niet curriculumontwikkeling beoogt, maar zich richt op het belang en de toepassing van sleutelkwalificaties kiest Van Zolingen voor de 'policy Delphi-aanpak'. Deze vorm richt zich op het genereren van beleidsalternatieven door middel van een gestructureerde dialoog (p. 155). Uiteindelijk zijn deskundigen vanuit onderwijs en arbeid betrokken bij het onderzoek. Gezien de doelstelling (beleidsalternatieven) is het jammer dat Van Zolingen uiteindelijk heeft besloten vertegenwoordigers vanuit de overheid niet te bevragen. Onderwijs

en arbeid te zamen in een panel onderbrengen, levert op voorhand echter al genoeg politieke discussie op, zeker als de vraag wordt voorgelegd wie waarvoor verantwoordelijk is!

In het eerste deel van het empirisch onderzoek inventariseert Van Zolingen de gevolgen van veranderingen in arbeidsorganisaties en op de arbeidsmarkt voor de gevraagde sleutelkwalificaties (zij valideert daarmee de resultaten van het literatuuronderzoek en de daaruit afgeleide sleutelkwalificaties en ze legt daarmee de basis van het empirische deel). Zij richt zich daarbij op drie sectoren: banken, verzekeringen en industrie. Non-profit organisaties zijn niet bij het onderzoek betrokken, waarmee conclusies over relevantie van sleutelkwalificaties over de breedte van de arbeidsmarkt niet mogelijk zijn. Daarnaast beperkt Van Zolingen zich tot sleutelkwalificaties in intredefuncties, wat vanuit het aansluitingsvraagstuk onderwijs-arbeid nuttig is.

Uiteindelijk worden 24 sleutelkwalificaties in het verdere verloop van het onderzoek meegenomen. In eerste instantie moeten dat er meer geweest zijn. Welke zijn afgefallen en waarom wordt echter niet duidelijk. De relevantie van de overgeblevene is echter evident. Willen sleutelkwalificaties betekenis krijgen en willen ze vertaald kunnen worden in eindtermen en leerplannen dan moeten ze contextspecifiek beschreven worden. Bijlage 5.6 laat zien dat dat mogelijk is. Zowel voor MEAO-C als voor MTS-E abiturienten zijn deze uitgewerkt en laten ze contextgebonden verschillen zien.

De overige drie onderzoeksvragen naar de rol van het MBO in de verwerving van sleutelkwalificaties, naar geschikte didactische methoden en naar vertaling in eindtermen leiden in het empirisch deel uiteindelijk niet tot verrassende uitkomsten. De resultaten geven richtingen aan, waarin verdere antwoorden gezocht kunnen worden, geen pasklare antwoorden. Maar daarin moet ook niet de grootste meerwaarde van het onderzoek gezocht worden. Dat zit in de conclusie dat het concept sleutelkwalificaties (opnieuw) interessant en relevant is.

Van Zolingen heeft met haar onderzoek ontegenzeggelijk het belang van sleutelkwalificaties aangetoond. En alhoewel keuzen en criteria binnen het onderzoek niet altijd geëxpliciteerd zijn, tonen het proefschrift en de onderzoeksresultaten dat ze zelf vele sleutelkwalificaties tot haar kwalificatiebagage mag rekenen. En dat het onderwerp inmiddels tot op beleidsniveau doorklinkt, blijkt wel uit de vraag die de minister van OCenW in een brief van 4 februari jongstleden voorlegt aan de SER: "Hoe kunnen sleutelkwalificaties vorm gegeven worden binnen beroepsonderwijs met een hoog gehalte aan realistische praktijksituaties?"

F.T.L. Blokhuis

Onderwijskundig adviseur CINOP

Mededelingen

Het laatste nummer van deze jaargang zal ongeveer het dubbele aantal pagina's omvatten en als boekeditie met de titel 'Opvoeding, onderwijs en sociale integratie' begin 1998 aan de abonnees worden toegezonden

Promoties

Universiteit van Amsterdam
Promovendus: Drs. M. van Hattum
Titel proefschrift: Pesten. Een onderzoek naar beleving, visie en handelen van leraren en leerlingen rondom pesten
Promotor: Prof.dr. P.G. Swanborn
Co-promotor: Dr. J.C. van der Wolf
Datum: 6 november 1997

Universiteit van Amsterdam
Promovendus: Drs. S.E. Severiens
Titel proefschrift: Gender and learning
Promotor: Prof.dr. B. van Hout-Wolters
Co-promotor: Dr. G. ten Dam
Datum: 13 november 1997

Universiteit van Amsterdam
Promovendus: Drs. P.J.C. de Bruijn
Titel proefschrift: Van een experiment naar een facet
Promotor: Prof.dr. G.W. Meijnen
Datum: 4 december 1997

Universiteit van Amsterdam
Promovendus: Drs. L. Sontag
Titel proefschrift: Vormgeving en effecten van onderwijs aan zesjarigen
Promotor: Prof.dr. G.W. Meijnen
Datum: 8 december 1997

Universiteit Twente
Promovendus: Drs. S.A. de Vries
Titel proefschrift: Aanpasbare courseware: Ontwerp en gebruik binnen het voortgezet onderwijs
Promotor: Prof.dr. J.C.M.M. Moonen
Co-promotoren: Dr. J.P.F. De Diana en Dr. P.A.M. Kommers
Datum: 3 december 1997

Universiteit Twente
Promovendus: Drs. C.A. Verwijs
Titel proefschrift: A mix of core and complementary media: New perspectives in media-decision making
Promotor: Prof.dr. J.C.M.M. Moonen
Co-promotor: Dr. B.A. Collis
Datum: 3 januari 1998

Rijksuniversiteit Groningen
Promovendus: Drs. J.F. Zijlstra
Titel proefschrift: Evaluatie van het programma 'Peuterrevalidatie vanuit opvoedingsperspectief'
Promotor: Prof.dr. H. Nakken
Datum: 16 oktober 1997

Katholieke Universiteit Nijmegen
Promovendus: Drs. C.M.T. Peters
Titel proefschrift: De professionaliteit van schoolleiders; een taakspecifieke benadering
Promotoren: Prof.dr. J.H.G.I. Giesbers en Prof.dr. T.C.M. Bergen
Co-promotor: Dr. P.J.C. Slegers
Datum: 17 september 1997

Katholieke Universiteit Nijmegen
Promovendus: Drs. A. Blonk
Titel proefschrift: Dyslexpert. The development and evaluation of an expert system for the diagnosis of reading and spelling problems
Promotor: Prof.dr. E.E.J. De Bruyn
Co-promotor: Dr. J.H.L. van den Bercken
Datum: 21 oktober 1997

Katholieke Universiteit Nijmegen
Promovendus: Drs. H. Wentink
Titel proefschrift: From graphemes to syllables: The development of phonological decoding skills in poor and normal readers
Promotoren: Prof.dr. R. Schreuder en Prof.dr. L. Verhoeven
Co-promotor: Dr. W.H.J. van Bon
Datum: 29 oktober 1997

Katholieke Universiteit Nijmegen
Promovendus: Drs. A.T.L.J. Monshouwer
Titel proefschrift: Ratio militans; perspectieven op Stephan Strasser
Promotor: Prof.dr. A.W. van Haften
Datum: 10 november 1997

Katholieke Universiteit Nijmegen
Promovendus: Drs. C. Koopmans
Titel proefschrift: Keynotes in music education. A philosophical analysis
Promotor: Prof.dr. A.W. van Haafden
Datum: 25 november 1997

Katholieke Universiteit Brabant
Promovendus: Drs. R.S. Irausquin
Titel proefschrift: Quality and use of phonological representations in poor and normal readers
Promotor: Prof.dr. B.L.M.F. de Gelder
Datum: 3 oktober 1997

Universiteit Utrecht
Promovendus: Drs. M.T. Glaudé
Titel proefschrift: Werkplek-opleiden als innovatie(middel). Leren en innoveren via opleiden op de werkplek
Promotor: Prof.dr. N.A.J. Lagerweij
Co-promotores: Dr. K.M. Stokking en Dr. A.M. Versloot
Datum: 17 oktober 1997

Recent gestart onderzoek

Universiteit van Amsterdam

Titel project: Bevorderende en belemmerende factoren voor het leven lang leren: een exploratief onderzoek op het microniveau
Contactpersonen: Dr. T.T.D. Peetsma en Dr. M.L.L. Volman, SCO-Kohnstamm Instituut

Titel project: Een onderzoek naar de predictieve validiteit van het Staatsexamen NT2
Contactpersonen: Drs. Y. Emmelot en Drs. E. van Schooten, SCO-Kohnstamm Instituut

Titel project: Intercultureel leren in de klas
Contactpersoon: Drs. G. Ledoux, SCO-Kohnstamm Instituut

Universiteit Twente

Titel project: Monitoring ICT in het onderwijs
Projectleiding: Prof.dr. Tj. Plomp en Dr. A.C.A. ten Brummelhuis

Oratie

Op 2 oktober 1997 heeft Prof.dr. J. Terwel zijn ambt van hoogleraar Onderwijspedagogiek aan de faculteit der Psychologie en Pedagogiek van de Vrije Universiteit Amsterdam aanvaard met het uitspreken van een rede, getiteld: *Strategisch leren in contexten*.

Benoeming

Prof.dr. C.M. de Glopper is per 1 september 1997 benoemd als directeur van het Onderwijsinstituut van de Faculteit der Pedagogische en Onderwijskundige Wetenschappen van de Universiteit van Amsterdam.

Inhoud andere tijdschriften

Comenius
17e jaargang, nr. 3, 1997

Themanummer: De toekomst van de historische pedagogiek
Inleiding op het themanummer, door P. Selten, N. Bakker en G. Biesta
De spanning van verborgen agenda's. Ontwikkelingen inzake het onderwijs in de historische pedagogiek, door M. Depaepe
Een markante omissie. Over de plaats van meisjes en vrouwen in Nederlandse historisch-pedagogische leerboeken, door N. Bakker en M. van Essen
Hulde aan het handwerk, kritiek op de theorie. Over de stand van zaken in de historische pedagogiek in Duitsland, door H.E. Tenorth

Pedagogisch Tijdschrift
22e jaargang, nr. 1/2, 1997

Themanummer: Sociaal leren
Inleiding op het themanummer, door D. Wildemeersch
Sociaal leren met het oog op maatschappelijke verantwoording. Het debat over landbouw, milieu en natuur, door J. Vandenabeele en D. Wildemeersch
Een riskante mythe. Volksopvoeding en volksgemeenschap, door B. van Gent
Sociaal leren in de context van sociale uitslui-

ting, door L. Veendrick en J. Zeelen
Sociaal leren in de risicomaatschappij: naar
een empirisch verantwoorde uitweg, door
B. Boog en M. van der Kamp
Civiele maatschappij, sociaal-cultureel werk
en sociaal leren, door R. van der Veen.

Pedagogisch Tijdschrift
22e jaargang, nr. 3, 1997

Voor het Nederlands als wetenschappelijke
taal, door A.W. van Haaften
Pestalozzi in Europa. Prolegomena van een
pedagogische werkingsgeschiedenis van zijn
erfgoed, door H. Van Crombrugge en
M. Depaepe

Tussen bewaking en verbetering van onder-
wijskwaliteit. Is total quality management
geschikt voor scholen? door J. Axe en R. Elte
Ethiek als expressie: een uitweg voor de peda-
gogiek? Besprekingsartikel, door P. Smeyers

Tijdschrift voor Orthopedagogiek
36e jaargang, nr. 7/8, 1997

Samenwerking tussen twee scholen voor spe-
ciaalonderwijs en de jeugdzorg, door
T.J.A. Schut, W.A.J.J. Willems en
S.W.M. Zijlmans

Heeft stellen nadelige gevolgen voor spellen?
door A.M.T. Bosman, K. Schep-Ottevanger en
W.H.J. van Bon

De interactie tussen spraaktaalgestoorde kin-
deren en hun ouders, door M.C.B. Drenthen
en J.M.A. Riksen-Walraven

Tijdschrift voor Orthopedagogiek
36e jaargang, nr. 9, 1997

Onderzoek naar de praktijktheorie over Prak-
tisch Pedagogische Gezinsbegeleiding bij
externaliserend probleemgedrag, door
G.M.P. Loots en P.A. de Ruyter
De Vineland Adaptief Gedragschaal bij kinde-
ren met cerebrale parese, door I.C.M. Ren-
tinck, A. Vermeer, D.A. van der Ploeg,
G.M.P. Loots en E.J. Karel

'De basis van de communicatie'. Kennisma-
king met een Russisch onderwijsprogramma
voor de sociale en emotionele opvoeding van
zes- tot zevenjarige kinderen, door
A.M.E. Strumphler

Ontvangen boeken

Blom, J.J.C. *Use-oriented courseware development
for agricultural education. An ecological appro-
ach* (dissertatie). Landbouwniversiteit Wage-
ningen, 1997.

Doef, P. van der, *Systematiek van orthopedagogisch
handelen*. Uitgeverij SWP, Utrecht, 1997, f 34,50.

Gerris, J.R.M. (red.), *Jongerenproblematiek: Hulp-
verlening en gezinsonderzoek*. Van Gorcum,
Assen, 1997, f 39,50.

Hattum, M.J.C. van, Pesten. *Een onderzoek naar
beleving, visie en handelen van leraren en leer-
lingen* (dissertatie). Universiteit van Amsterdam,
1997.

Houten, E. van, & Berg, G. van den, *Interacties in de
leefgroep. Interpersoonlijke processen tussen
groepsopvoeders en kinderen*. Uitgeverij SWP,
Utrecht, 1997, f 26,50.

Kef, S., Hox, J.J., & Habekothé, H.T., (On)Zichtbare
steun. *Onderzoek naar visueel gehandicapte
jongeren en hun netwerk*. Thesis Publishers,
Amsterdam, 1997, f 27,50.

Kemper, A., & Janssens, J., *Zin en onzin van Video-
Hometraining*. Uitgeverij SWP, Utrecht, 1997,
f 26,50.

Lange, M. de, & Klomp, M., *Agressieve jongeren.
Cognitie, emotie en gedrag*. Uitgeverij SWP,
Utrecht, 1997, f 26,50.

Levering, B., & Maanen, M. van, *Klein geheim. Inti-
miteit, privacy en ontwikkeling van identiteit*. De
Tijdstream, Utrecht, 1997, f 35,-.

Pameijer, N., & Beukering, T. van, *Handelingsgerich-
te diagnostiek*. Acco, Leusden, 1997, f 59,-.

Praamsma, J.M., *Nieuwe wereldburgers. Aantasting
van natuur en milieu als vraagstuk van algemene
vorming* (dissertatie). Universiteit Utrecht, 1997.

Rink, K. & Ott, W. (Eds.), *Youngsters between free-
dom and social limits*. Garant, Leuven/Apel-
doorn, 1997, f 45,50.

Severiens, S., *Gender and learning. Learning styles,
ways of knowing and patterns of reasoning* (dis-
sertatie). Universiteit van Amsterdam, 1997.

Stellingwerf, B.P., *Learning the rules of the game.
Strategies in teaching arithmetic word problem
solving* (dissertatie). Katholieke Universiteit Nij-
megen, 1997.

Visser, J., *Lezen op AVI-niveau* (dissertatie). KPC
Groep, 's-Hertogenbosch, 1997.