

Tijdschrift voor onderwijsresearch 1996

<https://hdl.handle.net/1874/208874>

TIJDSCHRIFT VOOR ONDERWIJS RESEARCH

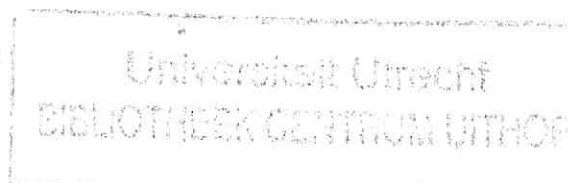
ISSN 0166-591 X

Jaargang 21
1996

Swets & Zeitlinger B.V.
Lisse

Redactie:

G. Wim Meijnen (voorzitter)
Klaas J. Westerhof (secretaris)
Anneke A. van der Hoeven-Van Doornum
Christiaan Hamaker
Bernadette van Hout-Wolters
Wim J. van der Linden
Joost Lowyck
Martin Mulder



SW

VERENIGING VOOR ONDERWIJSRESEARCH

RIJKSUNIVERSITEIT TE UTRECHT



2702 998 6

TIJDSCHRIFT VOOR ONDERWIJSRESEARCH

Een drie-maandelijks tijdschrift voor onderzoekers van onderwijs, gewijd aan theoretische, methodologische en professionele problemen in de onderwijsresearch.

Redactie:

G. Wim Meijnen (voorzitter)
Klaas J. Westerhof (secretaris)
Anneke A. van der Hoeven-Van Doornum
(redacteur Notities, Commentaren en
Boekbesprekingen)
Christiaan Hamaker
Bernadette van Hout-Wolters
Wim J. van der Linden
Joost Lowyck
Martin Mulder

Redactiesecretariaat:

Klaas J. Westerhof
Vakgroep Pedagogiek en Onderwijskunde
Rijksuniversiteit Groningen
Grote Rozenstraat 38
9712 TJ Groningen
tel.: 050-3636492/3636500
fax: 050-3636521
e-mail: K.J. Westerhof@ppsw.rug.nl

Manuscripten zende men in drievoud, getypt en persklaar aan de redactiesecretaris.
Manuscripten dienen te worden voorbereid overeenkomstig de Aanwijzingen voor auteurs.

Abonnementsprijs: Instituten f. 202,00
Particulieren f. 129,00
Studenten f. 102,50

De prijzen zijn inclusief porto, maar exclusief btw.

Men kan zich abonneren bij de uitgever: Swets & Zeitlinger, Heereweg 347, 2161 CA Lisse.
Fax 252-435111. E-mail: orders@Swets.nl

Voor leden van de Vereniging voor Onderwijsresearch (VOR) is het abonnementsgeld reeds inbegrepen in de contributie. Leden van de VOR hoeven zich dan ook *niet* afzonderlijk te abonneren bij de uitgever. Beëindiging van een abonnement geschiedt uitsluitend schriftelijk bij de uitgever: ten minste zes weken voor het einde van de abonnementsperiode. Bij niet-tijdige opzegging wordt het abonnement automatisch met één jaar verlengd.

Losse nummers zijn (à f. 50,00) verkrijgbaar bij de uitgever.

Advertenties dienen te worden gezonden aan het adres van de uitgever. Tarieven: 1/1 f. 450,00; 1/2 f. 260,00; 1/4 f. 190,00; exclusief zetkosten.

Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotocopie, microfilm of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

Het *Tijdschrift voor Onderwijsresearch* wordt uitgegeven door de Vereniging voor Onderwijsresearch.

Vereniging voor Onderwijsresearch
Secretariaat:
Dr. M. Mulder
Faculteit Toegepaste Onderwijskunde
Universiteit Twente
Postbus 217
7500 AE Enschede
tel.: 053-4893652
fax: 053-4329136
Mulder@edte.utwente.nl

Abonnementenadministratie:
Swets & Zeitlinger Publishers
Heereweg 347
2161 CA Lisse
tel. 0252-435111
fax: 0252-415888
giro 13984

Tijdschrift voor Onderwijsresearch is opgenomen in onderstaande online-data bestanden:
Bibliografie Nederlandse Sociale Wetenschappen.

Themanummer: het jonge kind in onderzoek

Inleiding: Onderwijs aan jonge kinderen nog nauwelijks onderzocht

Egbert Harskamp en Gerry Reezigt

*GION, Gronings Instituut voor onderzoek van onderwijs, opvoeding en ontwikkeling.
Rijksuniversiteit Groningen¹*

Het onderwijs aan jonge kinderen staat in Nederland al jarenlang in de belangstelling. In vaak hevige discussies rond de voorbereiding en de invoering van de Wet op het Basisonderwijs in 1985 werd veel aandacht geschonken aan de positie van wat toen nog het kleuteronderwijs werd genoemd. Die aandacht ontstond vanuit de vrees dat het kleuteronderwijs zijn eigenheid zou verliezen in de nieuwe basisschool, een vrees die enerzijds betrekking had op de veranderende rechtspositie van kleuterleidsters maar anderzijds ook op de inhoud van het onderwijs aan jonge kinderen. Een veel gehoorde klacht was dat de eigenheid van het kleuteronderwijs op de lange duur zou verdwijnen onder druk van de hogere leerjaren in de basisschool.

De grote aandacht voor het kleuteronderwijs is in de afgelopen jaren niet vertaald in een even grote aandacht voor onderzoek naar het onderwijs aan jonge kinderen zoals dat in Nederland op dit moment gestalte krijgt. Voorzover onderzoek is gedaan, betreft dat relatief vaak de initiële opleiding van leerkrachten, met de vraag of de eigenheid van het onderwijs aan jonge kinderen voldoende tot zijn recht komt in het onderwijsaanbod op PABO's. Daarnaast is onderzocht in hoeverre leerkrachten in de basisschool volledig inzetbaar zijn en ook als zodanig worden ingezet (Van de Ven, Vrieze & Claessen, 1988; Van den Bijtel, Blok & Deckers, 1992; Vermeulen & Bosker, 1992). Onderzoek naar wat er in de praktijk in groepen 1 en 2 gebeurt en naar wat de activiteiten opleveren in termen van opbrengsten bij leerlingen is de laatste jaren wel op gang gekomen. Zo is onder meer een longitudinaal onderzoek naar de invloed van gezinsfactoren en onderwijsfactoren op de ontwikkeling van kinderen gestart (Meijnen, 1990), waarin herhaalde metingen in de groepen 1 en 2 hebben plaatsgevonden. Maar vergeleken bij het onderzoek dat het afgelopen decennium is verricht naar allerlei aspecten van het onderwijs in de leerjaren 3 tot en met 8 steekt het onderzoek naar onderwijs aan jonge kinderen zeer mager af. Ook de verschillende aanbevelingen voor een geïntegreerd onderzoeksprogramma op dit terrein (ter voorkoming van diverse losse projecten) van de ARBO uit 1990 zijn niet overtuigend opgevolgd (ARBO, 1990). Het is dan ook geen wonder dat de Commissie Evaluatie Basisonderwijs moest concluderen dat er weinig gegevens bekend zijn over het feitelijke onderwijsaanbod en de opbrengsten van het onderwijs in groepen 1 en 2 (CEB, 1994). Ter verklaring voor deze situatie voert de CEB twee argumenten aan. Het eerste argument betreft de onduidelijkheid over het gewenste aanbod, het gewenste proces en de gewenste opbrengsten van het onderwijs aan jonge kinderen. Het tweede argument betreft de praktische problemen die het meten van opbrengsten bij jonge kinderen met zich meebrengt.

Op beide argumenten van de CEB valt het nodige af te dingen. Het is een feit dat vaardigheden van jonge kinderen moeilijker te toetsen zijn dan bijvoorbeeld prestaties van leerlingen in de bovenbouw van de basisschool, maar onmogelijk is het zeker niet (zie bijvoorbeeld Pellegrini & Galda, 1991). Daarbij komt dat de laatste jaren verschillende instrumenten zijn ontwikkeld (onder andere in het kader van leerlingvolgsystemen) die het toetsen van jonge kinderen vergemakkelijken (zie o.a. Van Kuyk, 1994). Wat het eerste argument betreft is het de vraag of er sprake is van onduidelijkheid over het gewenste aanbod, dan wel of er verschillende min of meer concurrerende meningen bestaan over het gewenste aanbod en in het verlengde daarvan over gewenste processen en gewenste opbrengsten. Voor dit laatste zijn enkele bewijzen aan te

¹ Adres: Postbus 1286, 9701 BG Groningen.

voeren. De laatste jaren zijn verschillende visies ontwikkeld op het onderwijs aan jonge kinderen (Van Oers & Janssen-Vos, 1992), die deels ook al in de onderwijspraktijk ingang vinden. Het uit België afkomstige initiatief van het ervaringsgeoriënteerde onderwijs (Laevers, 1994) en het project Basisontwikkeling van het Algemeen Pedagogisch Studiecentrum (Janssen-Vos, 1992) zijn op dit moment het meest populair in de onderwijspraktijk, maar andere, meer programmagerichte initiatieven zijn in ontwikkeling (Leseman & Cordus, 1994). Het is mogelijk dat, als deze ontwikkeling zich doorzet, binnen enkele jaren verschillende typen onderwijs in groep 1 en 2 naast elkaar bestaan. Vooralsnog is onduidelijk of de doelen die de initiatieven zelf nastreven worden bereikt en ook hoe deze doelen zich verhouden tot de algemene waarderingskaders die recent door de Inspectie zijn ontwikkeld en die ook door de CEB zijn gehanteerd (Inspectie, 1992; CEB, 1994). Voor het bevorderen van de kwaliteit van het onderwijs is het te hopen dat in de nabije toekomst onderzoek in uitvoering zal worden genomen dat het niet alleen mogelijk maakt uitspraken te doen over de relatie tussen onderwijsaanbod en opbrengsten in scholen die volgens de bovenbeschreven visies werken, maar ook over scholen die niet volgens een bepaalde uitgesproken visie te werk gaan. Dat deze laatste groep scholen zeer groot is, is door de CEB aangetoond (1994) en er is weinig aanleiding om te verwachten dat in die situatie snel verandering zal komen.

Door het ontbreken van empirisch onderzoek over de feitelijke situatie in de vroegere kleuterschool is het niet mogelijk vast te stellen hoe vroeger onderwijs werd gegeven en welke resultaten dat opleverde voor kinderen en in hoeverre dat onderwijs overeenkwam met de 'verworvenheden' van het kleuteronderwijs die nu worden gepropageerd. Een integraal onderzoek naar de wijze waarop nu onderwijs wordt gegeven in groep 1 en 2 en met welke resultaten voor kinderen ontbreekt eveneens. Duidelijk is wel uit recent onderzoek naar vaardigheden van 4- tot 6-jarigen dat er grote verschillen tussen kleuters zijn in kennis van begrippen die van belang zijn voor het leren lezen, schrijven en rekenen. Een grote groep kinderen beheerst de basisbegrippen niet als ze in groep 3 met schoolse vaardigheden worden geconfronteerd (Van Kuyk, 1994). Het is niet mogelijk na te gaan of er de afgelopen jaren veel is veranderd in het onderwijs aan en de resultaten van groep 1 en 2.

Wellicht is mede door het ontbreken van onderzoek naar het onderwijs en de opbrengsten in groep 1 en 2 nog steeds sprake van een hardnekkige mythevorming rond het onderwijs aan jonge kinderen. Discussies worden vaak gevoerd vanuit de veronderstellingen dat:

- het onderwijs vroeger beter was en meer recht deed aan de eigenheid van jonge kinderen,
- de Wet op het Basisonderwijs heeft geleid tot verschoolsing van het onderwijs in groep 1 en 2, waarbij wordt aangenomen dat verschoolsing automatisch een verslechtering betekent, en
- jonge kinderen nu worden gedwongen in een keurslijf van taken die niet op hun leeftijd zijn afgestemd terwijl ze vroeger onbekommerd konden spelen (zie voor een terugblik op de afgelopen jaren Janssen-Vos, 1991).

Impliciet wordt er vaak van uitgegaan dat vrij bezig zijn of spelen (activiteiten die overigens maar zelden duidelijk gedefinieerd worden) zonder meer is wat jonge kinderen de hele dag willen doen en dat leren iets heel anders is dan spelen. Wat dit betreft kan de titel van een van de programma's van het project Onderwijs en Sociaal Milieu uit 1979, "Kleuters Willen Ook Wel Eens Denken", nog steeds als een provocerende oproep worden beschouwd. Voorgestelde veranderingen, zoals het structureren van onderwijsactiviteiten in een gedifferentieerd onderwijsaanbod of het systematisch evalueren van vaardigheden van kinderen om daardoor een goed beeld van hun ontwikkeling te krijgen, worden vaak als verslechteringen afgedaan, waarbij dan vaak de suggestie wordt gewekt dat de huidige praktijk goed is of in elk geval niet beter zou kunnen worden door de voorgestelde veranderingen. Een voorbeeld is te vinden in het volgende citaat van Gammage (1994): "Our young children deserve to find themselves, to discover creativity, to learn through play and varied social encounters (...). The wide variability in children's development calls for a firm notion of care and individuality, not in uniform approaches of an externally constructed early childhood curriculum, but based on sound knowledge of what these children really are. We need to defend our children from uniformity and utilitarian approaches to education."

Dergelijke citaten floreren in de literatuur over het onderwijs aan jonge kinderen en maskeren, door hun gerichtheid tegen een meer gestructureerd onderwijsaanbod en duidelijke eisen, de noodzaak de huidige praktijk kritisch te beschouwen, juist ook om vast te stellen of het huidige aanbod voldoende inspeelt op de onderwijsbehoeften van leerlingen met achterstanden ten opzichte van hun leeftijdsgenoten.

De laatste tijd is echter een kentering op gang gekomen in kringen van degenen die zich met het kleuteronderwijs bezighouden. Steeds vaker verschijnen in vakbladen publicaties over de noodzaak om mondelinge en schriftelijk taalvaardigheid van jonge kinderen al vroeg te stimuleren door middel van een gericht onderwijsaanbod (zie o.a. het juninummer 1994 van het blad *De wereld van het jonge kind*, met als thema "Geletterd word je niet vanzelf!").

HET ONDERZOEK IN DIT THEMANUMMER

Het doel van dit themanummer is om door middel van het presenteren van onderzoeksgegevens bij te dragen aan de kennis over de feitelijke gang van zaken in het onderwijs en de ontwikkeling in cognitieve vaardigheden van jonge kinderen in groep 1 en 2. De discussie over het onderwijs zou aan kracht en importantie kunnen winnen als deze gevoerd wordt vanuit een empirisch onderbouwd beeld van de huidige onderwijspraktijk en van mogelijke veranderingen in die praktijk. Pas dan kan weloverwogen beslist worden of veranderingen nodig zijn en kans van slagen zullen hebben, ten einde zo veel mogelijk kinderen met succes de basisschool te laten doorlopen en vroegtijdige uitval te voorkomen.

Het eerste artikel in dit themanummer, *Onderzoek naar vormgeving en effecten van het onderwijs aan jonge kinderen in Nederland* door Gerry Reczigt, geeft een overzicht van de resultaten uit 13 onderzoeken die sinds 1985, het jaar dat de Wet op het Basisonderwijs in werking trad, zijn uitgevoerd. Bij het beschrijven van de verschillende onderzoeken wordt gebruik gemaakt van de centrale begrippen uit de literatuur over onderwijs-effectiviteit (Cremers, 1994). Het overzicht wijst uit dat van verschooling in groep 1 en 2 geen sprake is: leerkrachten hebben niet of nauwelijks kenmerken van het onderwijs in groep 3 tot en met 8 overgenomen. Daarmee is echter niet gezegd dat veranderingen in de huidige praktijk van groep 1 en 2 niet wenselijk of nodig zouden zijn; vooral een meer systematische keuze van onderwijsmaatregelen ten behoeve van het bereiken van welomschreven doelen bij verschillende groepen leerlingen is gewenst.

In het tweede artikel, *Stimulering van geletterdheid bij kleuters van de hand van Jef van Kuyk en Ludo Verhoeven*, wordt verslag gedaan van de effecten van verschillende onderwijsbenaderingen op de ontwikkeling van conceptuele taalvaardigheden, woordbewustzijn, fonembewustzijn en oriëntatie op boek en schrift. Een ontwikkelingsgerichte benadering, waarbij het accent ligt op het zelf kiezen van opdrachten door kinderen, wordt vergeleken met een gecombineerde benadering waarin ontwikkelingsgerichte elementen worden aangevuld met een programmagerichte benadering: naast het zelf kiezen van opdrachten krijgen kinderen één of twee keer per week vooraf vastgestelde taken aangeboden. De gecombineerde benadering leidt tot additionele effecten op enkele van de onderzochte effectmaten, met name de maten die betrekking hebben op het bewustzijn van kinderen dat woorden zijn opgebouwd uit klanken. Allochtone kinderen profiteren relatief sterk van de gecombineerde benadering.

Het derde artikel, *Ontwikkeling van tweetaligheid bij allochtone kleuters* door Ludo Verhoeven en Goretti Narain, gaat nader in op voorwaarden die aan het onderwijs aan allochtone kleuters gesteld zouden kunnen worden. De taalontwikkeling in de eerste en de tweede taal bij Turkse, Antilliaanse en Marokkaanse kinderen is onderzocht vanaf de leeftijd van 4 tot 7 jaar op verschillende aspecten (fonologisch, lexicaal, syntactisch en tekstueel). Deze ontwikkeling verloopt niet identiek in de onderzochte groepen: bij Turkse en Marokkaanse leerlingen blijft de eigen taal het Nederlands domineren, hoewel hun vooruitgang in het Nederlands sterker is dan in de eigen taal. Bij de Antilliaanse kinderen gaat het Nederlands de eigen taal overheersen. Voor het onderwijs in groep 1 en 2 impliceren de gegevens dat cognitieve stimulering in het

Nederlands zinvol is vanaf het moment dat allochtone kinderen een gebalanceerd niveau van tweetaligheid hebben bereikt. Bij Antilliaanse kinderen is dit doorgaans eerder het geval dan bij Turkse of Marokkaanse kinderen. Antilliaanse kinderen hebben echter ook nog grote achterstanden vergeleken bij Nederlandse kinderen (Kook & Vedder, 1994).

Het vierde artikel, *Actuele opvattingen over onderwijs aan jonge kinderen* door Luc Stevens en Egbert Harskamp, beschrijft een onderzoek dat zich richt op opvattingen en doelen van leerkrachten en het onderwijsaanbod in groep 2. Relaties tussen opvattingen en doelen van leerkrachten in het onderwijsaanbod en de beslissingen in de onderwijspraktijk zijn geanalyseerd. De meeste leerkrachten hebben een kindvolgende opvatting, waarbinnen ze taalontwikkeling, zelfstandigheid, werkhouding en motorische ontwikkeling de belangrijkste doelen vinden. Kindvolgende leerkrachten verschillen niet van meer kindsturende leerkrachten waar het gaat om concrete beslissingen in pedagogische situaties. Kindvolgende leerkrachten verschillen wel van kindsturende leerkrachten waar het gaat om het aanbod van specifieke leervoorwaarden voor rekenen en taal. Kindsturende leerkrachten bieden meer activiteiten op het gebied van taal- en rekenvoorwaarden dan kindvolgende leerkrachten. Het valt op dat vrijwel geen van de leerkrachten groot belang hecht aan het aanbieden van deze voorwaarden, terwijl alle kindsturende en ongeveer de helft van de kindvolgende leerkrachten de voorwaarden in hun aanbod hebben. Er zijn dus discrepanties tussen de opvattingen van leerkrachten ten aanzien van voorbereidend taal en leren lezen en hetgeen zij in de praktijk doen. Voor het relateren van opbrengsten van onderwijs aan gedrag van de leerkracht moet dan ook in ieder geval naar het feitelijk handelen worden gekeken.

Het vijfde en laatste artikel, *Onderwijs aan jonge kinderen: historische ontwikkelingen van de hand van Wim Meijnen*, schetst ten slotte een beeld van de geschiedenis van het onderwijs aan jonge kinderen. Het artikel maakt duidelijk dat scholen al heel lang een rol spelen in het onderwijs en oorspronkelijk ook vooral de (religieuze) opvoeding van het jonge kind, in ieder geval bepaalde groepen jonge kinderen uit verschillende maatschappelijke groeperingen. De rol die het onderwijs speelt bij het vergroten of verkleinen van maatschappelijke verschillen tussen kinderen is inmiddels goed gedocumenteerd en onderzocht. Duidelijk is geworden dat ook de eerste leerjaren een eigen bijdrage leveren, een inzicht dat lange tijd niet expliciet van invloed is geweest op de vormgeving van het onderwijs. Het onderwijsbeleid in de nabije toekomst zal verder moeten inspelen op dit gegeven, evenals op de behoefte aan mogelijkheden voor en kwaliteitsverbetering van opvang voor kinderen voorafgaand aan de basisschool.

Hoewel de artikelen verschillende aspecten van het onderwijs in groep 1 en groep 2 belichten, maken ze alle duidelijk dat het onderwijs in groep 1 en 2 beter kan worden afgestemd op de verschillende onderwijsbehoeften van leerlingen. Het onderzoek naar de vormgeving en effecten van het onderwijs en het onderzoek naar actuele opvattingen laten zien dat het huidige onderwijs in groep 1 en 2 niet overeenstemt met het ideaalbeeld dat zo vaak wordt geschetst. Leerkrachten hebben in het algemeen slechts globale doelen in hun hoofd, die geen richting geven aan hun dagelijks omgaan met kinderen en hooguit het aanbod van leeractiviteiten enigszins beïnvloeden. Leerkrachten slagen er niet vanzelfsprekend in ieder kind te geven wat hij op een bepaald moment nodig heeft, onder andere omdat ze niet systematisch nagaan hoe kinderen zich ontwikkelen.

Denken over een andere, meer gestructureerde aanpak van het onderwijs in het licht van welomschreven doelen is voor het planmatig leren handelen van leerkrachten van belang. Vooral voor het extra stimuleren van de ontwikkeling van allochtone kinderen, de groep die op dit moment de grootste onderwijsachterstanden heeft, is een planmatige inzet van leerkrachten gewenst. De onderzoeken naar de stimulering van geletterdheid en naar de ontwikkeling van tweetaligheid laten zien dat er mogelijkheden zijn om de achterstanden van allochtone kleuters met probate middelen aan te pakken.

De artikelen maken duidelijk hoeveel er nog niet is onderzocht. Het is in het belang van de kwaliteit van het onderwijs dat in de nabije toekomst meer onderzoek kan worden gedaan naar de

effecten van concrete onderwijsbenaderingen in groep 1 en 2, maar ook naar de effecten van de huidige praktijk in reguliere groepen 1 en 2 op de ontwikkeling van leerlingen. Daarbij kan ontwikkeling breed worden gedefinieerd, zonder een eenzijdige nadruk op aspecten van de cognitieve ontwikkeling. De affectieve ontwikkeling van jonge leerlingen is een zo mogelijk nog minder ontgonnen onderzoeksterrein. Voor een volledig beeld van de processen en de opbrengsten in de eerste jaren van de basisschool zou in deze situatie verandering moeten worden gebracht.

Een onderzoeksterrein dat in Nederland eveneens grotendeels onontgonnen is betreft de wijze waarop in groepen 1 en 2 met spel wordt omgegaan en welke effecten dit heeft op de ontwikkeling van kinderen. Spel wordt nu vaak beschouwd als iets dat in zichzelf goed is en waar weinig stimulering vanuit de leerkrachten bij nodig is (Harskamp, Pijl & Snippe, 1992). Uit onderzoek is echter bekend dat bepaalde vormen van spel gerelateerd zijn aan de mate van geletterdheid van kinderen, in ieder geval aan de beginstadia van het schrijven (Pellegrini & Galda, 1991). Bepaalde aspecten van spel, zoals het overleg van kinderen over het spel, hangen samen met begrip van teksten. Er is bovendien een sterke samenhang tussen de wijze van spelen en de intelligentie van kinderen, en er zijn aanwijzingen dat bepaalde typen spel op sommige kinderen andere effecten hebben dan op andere kinderen, bijvoorbeeld afhankelijk van de leeftijd. Een actieve rol van de leerkracht is vereist om een geschikte omgeving te bieden of om kinderen op een betekenisvolle wijze aan het spelen te krijgen en te houden (Christie, 1991). Afgezien van één tamelijk kleinschalig onderzoek naar spel in groep 1 en 2 (Harskamp, Pijl & Snippe, 1992) is nauwelijks iets bekend over spel in de eerste jaren van de basisschool. Bij het programmeren van nieuw onderzoek zou ook dat onderzoeksterrein in beschouwing moeten worden genomen.

LITERATUUR

- Adviesraad voor het Basisonderwijs (1990). *Spelen en jongleren*. Zeist: ARBO.
- Bijtel, J. van der, Blok, H., & Deckers, P. (1992). *Kleuters in het PABO-curriculum*. Den Haag: SVO.
- Commissie Evaluatie Basisonderwijs (1994). *Onderwijs aan jonge kinderen*. Den Haag: SDU.
- Christie, J.F. (1991). Play and early literacy development: summary and discussion. In J.F. Christie (Ed.). *Play and early literacy development*. Albany: State University of New York.
- Creemers, B.P.M. (1994). *The effective classroom*. London: Cassell.
- Gammage, P. (1994). In defence of children. In F. Laevers (Ed.). *Defining and assessing quality in early childhood education*. Leuven: University Press.
- 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*. Den Haag: SDU.
- Janssen-Vos, F. (1991). Van kleuteronderwijs naar basisonderwijs, wat zich in tien jaar tijd afspeelde. *De Wereld van het Jonge Kind*, 140-149.
- Janssen-Vos, F. (1992). *Basisontwikkeling*. Assen: Van Gorcum.
- Kuyk, J. van (1994). Does ordering lay the foundation of acquiring mathematics in young children? In F. Laevers (Ed.). *Defining and assessing quality in early childhood education*. Leuven: University Press.
- Kook, H., & Vedder, P. (1994). The development of reading skills in Antillean 4-6 year olds. In F. Laevers (Ed.). *Defining and assessing quality in early childhood education*. Leuven: University Press.
- Laevers, F. (1994). The innovative project Experiential Education and the definition of quality in education. In F. Laevers (Ed.). *Defining and assessing quality in early childhood education*. Leuven: University Press.
- Leseman, P.P.M., & Cordus, J. (1994). *(Allochtone) kleuters meer aandacht*. Rijswijk: WVC.
- Meijnen, G.W. (1990). *De schoolcarrière van leerlingen gezien in relatie tot de invloeden die maatschappelijke en binnenschoolse determinanten hierop hebben*. Amsterdam: SCO.
- Oers, B. van, & Janssen-Vos, F. (1992). *Visies op onderwijs aan jonge kinderen*. Assen/Maastricht: Van Gorcum.

- Pellegrini, A.D., & Galda, L. (1991). Longitudinal relations among preschoolers' symbolic play, metalinguistic verbs, and emergent literacy. In J.F. Christie (Ed.). *Play and early literacy development*. Albany: State University of New York.
- Ven, J. van de, Vrieze, G., & Claessen, J. (1988). *Grensverkenners in de basisschool*. Nijmegen: ITS.
- Vermeulen, C.J., & Bosker, R.J. (1992). *De omvang en gevolgen van deeltijdarbeid en volledige inzetbaarheid van leerkrachten in het basisonderwijs*. Enschede: UT/OCTO.

Manuscript ontvangen 6-9-1994

Definitieve versie ontvangen 5-4-1995

Onderzoek naar vormgeving en effecten van het onderwijs aan jonge kinderen in Nederland

Gerry Reezigt

*GION, Gronings Instituut voor Onderzoek van Onderwijs, opvoeding en ontwikkeling,
Rijksuniversiteit Groningen¹*

ABSTRACT

In this article, a review of 13 research studies on the education of young children (4-6 year old) is presented. To compare the findings of the studies, the main concepts of a general educational effectiveness model are used. With regard to instruction, teachers in grades 1 and 2 of Dutch primary schools tend to manage their classrooms in a rather orderly way. They pursue very general goals, plan their daily work on the basis of thematic concepts, do not act very directive, and do not systematically evaluate the progress of their students. Teachers provide their students with varied subjectmatter and task assignments, but they do not really implement adaptive instruction. Curricula for grades 1 and 2 are not systematically used and little attention is paid to activities directly related to reading, writing, and mathematics. Half of the time in grades 1 and 2 can be considered educational time (as opposed to time spent on playing, eating, drinking etc.) and in general 75 % of the students show involvement in classroom activities.

Further (longitudinal) research is necessary to investigate the influence of teacher and classroom characteristics on the educational progress of the students.

1. INLEIDING

Sinds de voorbereiding en de invoering van de Wet op het Basisonderwijs (WBO) is er een groeiende aandacht voor het onderwijs aan jonge kinderen, in dit artikel opgevat als het onderwijs in groep 1 en 2. Dat heeft echter lange tijd niet of nauwelijks geleid tot onderzoek naar de periode aan het begin van de basisschool, terwijl de hogere leerjaren wel uitgebreid zijn bestudeerd. De discussies over de vormgeving en mogelijke effecten van het onderwijs aan jonge kinderen hadden overwegend betrekking op de 'verworvenheden' van het kleuteronderwijs (Van Erp & De Visch Eybergen, 1983), die, zoals sommigen vreesden, verloren zouden gaan bij de integratie van kleuter- en lagere school. Argumenten ontleend aan empirisch onderzoek werden en worden in die discussies maar zelden gebruikt. Enerzijds zullen niet alle betrokkenen evenveel behoefte hebben aan empirisch gefundeerde argumenten, anderzijds zijn er beperkingen aan het onderzoek dat zowel nationaal als internationaal op het onderhavige terrein wordt verricht.

Er is weliswaar een overvloed aan onderzoek naar de wijze waarop kinderen zich ontwikkelen en de wijze waarop ze leren, maar deze literatuur is praktisch altijd afkomstig uit de ontwikkelingspsychologische hoek en zeer zijdelings gerelateerd aan het onderwijs (Beishuizen, 1983; Koops, 1983; Dijkstra, Dudink & Takens, 1983; Weinert, 1987; Goffin, 1989). Hoewel sommige ontwikkelingspsychologische theorieën duidelijke implicaties hebben voor het onderwijs, zoals de theorieën van Piaget (Lunzer, 1986; Weinert, Schrader & Helmke, 1987) en Vygotsky (Van Parreren, 1987, 1990), betreft het meeste onderzoek experimenten die tamelijk ver verwijderd zijn van de praktijk in schoolklassen. Onderzoek vanuit een onderwijskundige invalshoek heeft doorgaans juist wel betrekking op de feitelijke onderwijspraktijk en effecten daarvan, maar beperkt zich meestal tot leerlingen van zes jaar of ouder, zowel in Nederland als in andere

1. Adres: Postbus 1286, 9701 BG Groningen.

landen. Oorzaken daarvoor liggen in de start van de leerplicht, onduidelijkheid over de gewenste effecten en problemen bij het meten van effecten.

- In Nederland begint de leerplicht op vijfjarige leeftijd. In een aantal andere landen ligt dat anders: in Zweden bijvoorbeeld zijn kinderen pas op zevenjarige leeftijd leerplichtig (Pramling, 1986). In de Verenigde Staten zijn verschillen van staat tot staat in de start van de leerplicht, die variëren van vijf (één staat) tot acht jaar (vier staten; Glazer, 1985). In sommige landen is dus ook nog geen sprake van effecten van regulier onderwijs op vier- en vijf-jarige kinderen (Bridgman, 1989).
- In het onderwijs aan jonge kinderen bestaat onduidelijkheid over de gewenste vormgeving en effecten (Karweit, 1989). Ook in Nederland bestaat geen set van welomschreven doelstellingen voor groep 1 en 2. Daardoor is discussie mogelijk over wat het onderwijs precies in zou moeten houden en welke effecten bereikt zouden moeten worden. Dergelijke discussies zijn bijvoorbeeld nog gaande wat het voorbereidend lezen en rekenen betreft, waarbij het aanleren van 'klassieke' lees- en rekenvoorwaarden gecontrasteerd wordt met nieuwere ideeën over 'emergent literacy' en 'emergent numeracy' (Boland & Mommers, 1991; Harskamp & Willemsen, 1991; Van Kuyk, 1991a; Verhoeven, 1992). In hogere leeftijdsgroepen van het basisonderwijs is ook niet heel concreet omschreven wat leerlingen precies zouden moeten kennen en kunnen, maar er bestaat toch meer impliciete overeenstemming, onder andere afgeleid uit het onderwijsaanbod in de methoden die in de verschillende leerjaren worden gebruikt. Bovendien is het in dat geval voor onderzoeksdoeleinden eenvoudiger om uit gebruikte methoden af te leiden wat eigenlijk wordt nagestreefd, en wat dus in praktijk effecten zouden kunnen zijn (Gillijns & Verhoeven, 1992; Harskamp, Edelenbos & Suhre, 1993).
- Een laatste knelpunt in het onderzoek naar onderwijs aan jonge kinderen ligt in de praktische problemen, zowel op de korte termijn als (qua interpretatie) op de lange termijn. Klassikaal afnemen van toetsen is bij jonge leerlingen zo goed als onmogelijk, zodat alleen al om budgettaire redenen vaak besloten wordt onderzoek in hogere leerjaren te laten plaatsvinden. Daar komt bij dat, ook als inhoudelijke keuzes zijn gemaakt ten aanzien van effectmaten, in longitudinaal onderzoek problemen ontstaan met de vergelijkbaarheid van effectmaten in latere leerjaren (Davis, 1990). Voorzover gegevens bekend zijn over schoolse prestaties van jonge kinderen, gaat het vaak om onderzoek dat gekoppeld is aan de ontwikkeling van specifiek materiaal of specifieke programma's, waarbij kinderen ook uitsluitend worden getoetst op de vaardigheden die het materiaal geacht wordt aan te leren (o.a. Van Kuyk, 1991a). Dergelijk onderzoek levert geen gegevens op over het algemene prestatieniveau van kinderen in groep 1 en 2.

Ondanks deze verklaringen voor het ontbreken van onderzoek naar effecten, blijft het opvallend dat in de jaren rond de invoering van de WBO, toen talloze onderzoeken zijn uitgezet naar de vormgeving en de effecten van het onderwijs in de hogere leerjaren van de basisschool, ook geen enkel grootschalig onderzoek naar de vormgeving van het onderwijs in de groepen 1 en 2 heeft plaatsgevonden. Er is in de periode na de WBO wel onderzoek gedaan naar de consequenties van de WBO voor het onderwijsaanbod op PABO's en naar de inzetbaarheid van leerkrachten in de onderwijspraktijk (Van de Ven, Vrieze & Claessen, 1988; Van den Bijtel, Blok & Deckers, 1992; Vermeulen & Bosker, 1992). Daarnaast zijn gegevens over de situatie in groep 1 en 2 beschikbaar die ontleend zijn aan onderzoek dat zich in eerste instantie richtte op de hogere leerjaren, maar de groepen 1 en 2 'meenam'. De afgelopen jaren hebben enkele kleinschalige observatiestudies plaatsgevonden, die specifiek gericht waren op het onderwijs in groep 1 en 2 (Van Gennip & Ruijs, 1992; Harskamp, Pijl & Snippe, 1992). Ook is een longitudinaal onderzoek gestart naar de effecten van het onderwijs aan het begin van de basisschool, in het bijzonder voor verschillende groepen allochtone leerlingen (Meijnen, 1990; Van der Leij, Meijnen & Leseman, 1994). Het eerste grootschalige onderzoek naar de feitelijke situatie in groep 1 en 2 is pas verricht in het kader van de werkzaamheden van de Commissie Evaluatie Basisonderwijs in het najaar van 1992 (Inspectie, 1994b), ruim 7 jaar na de invoering van de WBO.

2. OPZET VAN HET ARTIKEL

Vraagstelling en doel

De bedoeling van dit artikel is om op basis van literatuurstudie een overzicht te geven van het empirisch onderzoek dat in Nederland in het laatste decennium is gedaan naar het onderwijs in groep 1 en 2, aan de hand van de vraag:

op welke wijze wordt het onderwijs in groep 1 en 2 in de praktijk vormgegeven en welke effecten op leerlingen worden bereikt door verschillen in vormgeving van het onderwijs in groep 1 en 2?

Idealiter zou uit onderzoek kunnen blijken welke vormgeving van het onderwijs leidt tot bepaalde effecten op leerlingen. Afhankelijk van de specifieke gewenste effecten zou het dan ook mogelijk zijn de empirisch meest effectieve vormgeving te beschrijven, waarna de huidige praktijk in groep 1 en 2 daar tegen afgezet zou kunnen worden. Voor een dergelijke werkwijze is echter onvoldoende onderzoeksmateriaal beschikbaar. Het is op dit moment ook niet mogelijk de huidige praktijk in groep 1 en 2 en de opbrengsten in termen van het functioneren van leerlingen volgens algemeen aanvaarde kwaliteitscriteria te beoordelen. Voor een dergelijke beoordeling zou gebruik gemaakt moeten kunnen worden van criteria en standaarden, een werkwijze die gevolgd is door de Inspectie (1992, 1994a) en de Commissie Evaluatie Basisonderwijs (1994a, 1994b). Omdat momenteel echter nog niet is vastgelegd welke leerinhouden leerkrachten geacht worden over te dragen in groep 1 en 2 en welke effecten op cognitief en affectief gebied bereikt zouden moeten worden, kunnen de gehanteerde standaarden van de Inspectie en de CEB voorlopig niet als richtinggevend gezien worden.

Het is daarentegen wel mogelijk onderzoeksbevindingen over het onderwijs in groep 1 en 2 in een historisch perspectief te plaatsen door ze af te zetten tegen de vaak geciteerde 'verworvenheden' van het vroegere kleuteronderwijs (Van Erp & De Visch Eijbergen, 1983; CEB, 1994a) en in een actueel perspectief door ze af te zetten tegen algemene theorieën en onderzoeksresultaten over onderwijseffectiviteit, die niet specifiek aan leerjaren of leeftijden zijn gekoppeld. Voor deze invalshoek is in dit artikel gekozen (zie verder 7).

Beschrijvingskader

Bij de beschrijving van onderzoeksresultaten wordt gebruik gemaakt van het onderwijseffectiviteitsmodel van Creemers (1991, 1994), waarbij op het niveau van de leerkracht de componenten instructie, differentiatie en materiaal centraal staan. Wat instructie betreft worden verschillende indicatoren onderscheiden, zoals management van de klas, duidelijkheid van doelstelling, structureren van de inhoud van het onderwijs, en registratie en evaluatie van de ontwikkeling van leerlingen (zie verder 3.1). Deze componenten beïnvloeden de gelegenheid om te leren en de daadwerkelijke taakgerichte leertijd ofwel de betrokkenheid van leerlingen en daardoor uiteindelijk hun cognitief functioneren. Naast het niveau van de leerkracht is in de context van de WBO op schoolniveau vooral de overeenstemming in werkwijze tussen leerkrachten van belang (o.a. Bowman, 1993). Figuur 1 laat een schematische weergave zien van de elementen van het model en de veronderstelde relaties.

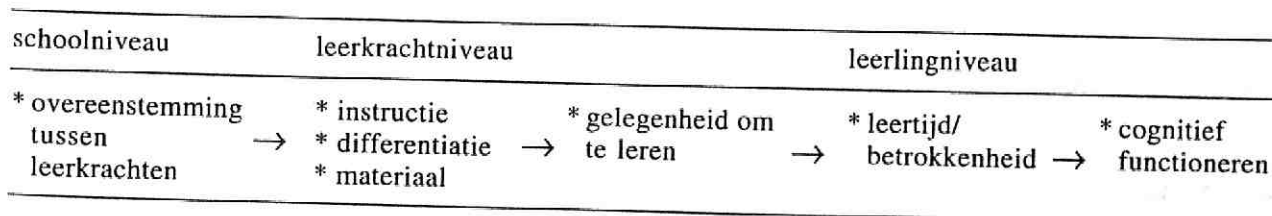


Fig. 1. Beschrijvingskader onderzoeksresultaten (ontleend aan Creemers, 1991, 1994).

Bij de keuze van een onderwijseffectiviteitsmodel als beschrijvingskader voor het onderwijs in groep 1 en 2 kan een tweetal kanttekeningen worden geplaatst, die betrekking hebben op de geldigheid van de in het effectiviteitsmodel onderscheiden begrippen voor het onderwijs in groep 1 en 2 en op de nadruk die het model legt op het cognitief functioneren.

Ten aanzien van de vraag of de begrippen uit het model voor het onderwijs in groep 1 en 2 voldoende geldigheid bezitten, kan worden opgemerkt dat de begrippen van algemene aard zijn en niet aan bepaalde leeftijdscategorieën van leerlingen gekoppeld. Dat neemt niet weg dat de meeste onderzoeksliteratuur, waar de empirische evidentie van de begrippen op gebaseerd is, betrekking heeft op leerlingen die ouder zijn dan de vier- tot zesjarigen die hier centraal staan. Onderzoek dat in de Verenigde Staten en ook in Nederland is gedaan naar effecten van programma's voor jonge kinderen (in achterstandssituaties) wijst echter uit dat, ondanks de enorme diversiteit in programma's, grotendeels dezelfde kenmerken voor deze leeftijdsgroep effectief lijken te zijn als voor hogere leeftijdsgroepen (o.a. Karweit, 1989). Ondanks een grote verscheidenheid in programma's, verschillen in nagestreefde doelen en verschillen in gehanteerde effectmaten (Stallings & Stipek, 1986) komen het stellen van duidelijke doelen, het structureren van activiteiten, het nauwkeurig volgen van de ontwikkeling van kinderen, het optimaliseren van de gelegenheid tot leren door adequaat klasmanagement en het volgen van eenzelfde aanpak over leerjaren heen consequent als effectieve kenmerken naar voren (Meijnen & Riemersma, 1992; Reezigt, Guldemon & Ros, 1993; Leseman & Cordus, 1994). Bovendien zijn belangrijke begrippen uit visies op onderwijs aan jonge kinderen, die momenteel furor maken (Laevers, 1991; Janssen-Vos, 1992), voor een deel ook terug te voeren op de begrippen uit het effectiviteitsmodel, ook al bestaat niet altijd consensus over na te streven effecten en ook al is discussie mogelijk over de concrete invulling van de begrippen in het geval van groep 1 en 2. Zo benadrukt Laevers (1991) bijvoorbeeld het belang van aantrekkelijk materiaal en het voortdurend observeren van leerlingen, met als doel eventuele problemen zo snel mogelijk te signaleren en aan te pakken. Daarnaast stelt hij de betrokkenheid van kinderen bij wat ze doen centraal. Ook Simonis (1993) benadrukt de betrokkenheid als cruciale intermedierende variabele tussen de aanpak van de leerkracht en de effecten op leerlingen. Janssen-Vos (1992, 1993a, 1993b) benadrukt o.a. het belang van een zorgvuldige selectie van activiteiten, het belang van differentiatie en het belang van het gericht observeren van leerlingen en het registreren van het onderwijsaanbod, effecten van het aanbod en consequenties voor vervolgactiviteiten.

In het effectiviteitsmodel worden geen effecten in het affectieve domein gespecificeerd. Voorzover in dit artikel onderzoeksresultaten worden beschreven die betrekking hebben op effecten, gaat het ook alleen om effecten op ofwel de betrokkenheid van leerlingen, ofwel het cognitief functioneren. De keuze hiervoor is van pragmatische aard. Wat de onderzoeksliteratuur op het terrein van onderwijs-effectiviteit betreft geldt dat effecten van onderwijs in het affectieve domein minder frequent zijn onderzocht en dat minder goed modelmatig te specificeren valt welke effecten van welke factoren op het niveau van de leerkracht of de school verwacht kunnen worden. De grote variatie aan effectmaten in het affectieve domein (motivatie, attitudes, attributies, zelfbeeld, schoolwelbevinden e.d.) vergroot deze onduidelijkheid nog. Wat de ontwikkeling van jonge kinderen in het affectieve domein betreft wordt maar zelden naar de rol van het onderwijs gekeken: welke concrete gedragingen leerkrachten zouden kunnen of moeten vertonen om bepaalde gewenste effecten op de affectieve ontwikkeling van kinderen te bewerkstellingen is niet duidelijk (o.a. McLeod, 1988; Leune, 1992; Bowman, 1993). Voorzover literatuur betrekking heeft op het doelgericht beïnvloeden van de affectieve ontwikkeling door het onderwijs gaat het doorgaans om specifieke groepen leerlingen: hoogbegaafde kinderen en kinderen met ernstige leer- en gedragsproblemen (o.a. Halsted, 1990; Van Lieshout, 1992; Monks & Peters, 1992; Winnubst, 1992), vaak in de hogere leerjaren.

De keuze voor het cognitief functioneren brengt noodzakelijkerwijs beperkingen met zich mee. Er is echter enige reden om aan te nemen dat factoren die effectief zijn op het cognitieve terrein dat ook zijn op het affectieve terrein (Zaal, 1980; Knuver, 1993),

Selectie van onderzoeken

Voor het overzicht van onderzoek in dit artikel is gebruik gemaakt van empirisch onderzoek vanaf 1985 naar de vormgeving van het onderwijs en de effecten daarvan op het cognitief functioneren van leerlingen. Het belangrijkste selectie criterium is dat het onderzoek betrekking heeft op de doorsnee klaspraktijk in groep 1 en 2. Literatuur die uitsluitend betrekking heeft op specifieke groepen leerlingen (zoals leerlingen uit lager sociaal milieu, allochtone leerlingen) is alleen in beschouwing genomen als er onderzoeksgegevens gepresenteerd zijn over de reguliere werkwijze in klassen. Onderzoek naar effecten van onderwijsexperimenten op specifieke groepen leerlingen (b.v. Wagenaar, 1992) is om die reden niet opgenomen. Onderzoek waarin opbrengsten zijn gemeten zonder dat koppelingen gemaakt kunnen worden met het onderwijs dat kinderen hebben genoten (Van Kuyk, 1991a; Verhoeven & Van Kuyk, 1991) blijft eveneens buiten beschouwing.

In totaal zijn 13 onderzoeken geselecteerd. Tabel 1 geeft per onderzoek een korte beschrijving van de volgende gegevens:

- de rapportages die ten behoeve van dit artikel zijn gebruikt,
- de periode waarin de data zijn verzameld,
- het aantal leerkrachten en leerlingen in de steekproef,
- de wijze waarop leerkrachten en leerlingen zijn geselecteerd,
- de wijze waarop data zijn verzameld en, voorzover van observatie sprake is, of het gaat om data die weinig (low inference) dan wel veel (high inference) interpretatie van de observator vergen,
- de componenten op leerkracht- en leerlingniveau uit het onderwijseffectiviteitsmodel waarover data beschikbaar zijn. Veel onderzoeken doen ook uitspraken over het schoolniveau, meestal gebaseerd op gegevens van leerkrachten uit verschillende leerjaren. Omdat zelden data op schoolniveau worden verzameld, is het schoolniveau niet als apart niveau in tabel 1 opgenomen.

Tabel 1 laat zien dat de meeste onderzoeken (4 tot en met 13) van zeer recente datum zijn (gegevens verzameld in 1990 of later). Voor een deel zijn ze nog niet afgerond (10 en 11, die dezelfde groep leerkrachten betreffen). De drie onderzoeken die gegevens uit 1985 rapporteren (1, 2 en 3) hebben kort voor de officiële invoering van de WBO plaatsgevonden. Daarna is er een periode niets gebeurd, waarna een grotere stroom onderzoeken op gang is gekomen.

Het meeste onderzoek is tamelijk kleinschalig (uitgezonderd 2, 3 en 13). Oorzaken daarvoor liggen in de meeste gevallen in de gehanteerde onderzoeksmethode (observaties), het feit dat sprake was van een onderzoek in een beperkte regio (1) of van een deel- of voorstudie (5, 9). Verschillende onderzoeken betreffen in hun totaliteit feitelijk grotere aantallen leerkrachten uit andere leerjaren van de basisschool (2, 3, 4, 7), omdat ze niet speciaal op de groepen 1 en 2 gericht waren. In de 8 onderzoeken waar leerkrachten geobserveerd worden gaat het meestal om een eenmalige observatie (uitgezonderd 9, 10 en 13). Omdat leerkrachten niet altijd hetzelfde gedrag vertonen maar vrij grote verschillen laten zien (Reezigt, Guldmond & Ros, 1993) kunnen de eenmalige observatiegegevens vertekend zijn. In 3 van de 8 observatiestudies (6, 12, 13) is sprake van high inference observaties, waarbij niet zozeer waarneembaar gedrag van de leerkracht is gescoord, maar waarbij door de observatoren beoordeeld is hoe de leerkracht zich gedraagt. In 2 studies (12 en 13) zijn de observaties niet verricht door onderzoekers, maar door inspecteurs van het onderwijs. De vraag is in dat geval in hoeverre de werklessen die geobserveerd zijn als 'normale' werklessen te beschouwen zijn. Voor de vragenlijstgegevens geldt dat sociale wenselijkheid een rol kan spelen, evenals voor de gegevens die door middel van logboeken of interviews verzameld zijn.

Uit Tabel 1 blijkt duidelijk dat de meeste onderzoeken zich uitsluitend richten op het niveau van de leerkracht. Gegevens op het niveau van de leerlingen zijn zeer schaars: data over de betrokkenheid van leerlingen bij hetgeen in de klas gebeurt en bij hun opdrachten en taken zijn slechts in drie onderzoeken verzameld (8, 9, 10) en data over het cognitief functioneren slechts

Tabel 1. Kenmerken van onderzoeken naar onderwijs aan jonge kinderen.

auteurs	data uit	aantal leerkr/leerl.	steekproef	wijze van data-verzameling	componenten effectiviteitsmodel waar informatie over wordt gegeven:					
					leerkracht		leerling			
					instr	diff	mat	gel	betr	cogn
1. Van der Meer, Appelhof, ten Doeschate & Kramer (1986)	1985	28 lkr. groep 1 en 2	regio Utrecht	observatie lkr (low inf), 1 ochtend per lkr	*	*	—	*	—	—
2. Reezigt, van Dijk & Bosveld (1986)	1985	146 lkr. groep 1 en 2	aselect	vragenlijst leerkrachten	*	*	—	—	—	—
3. Wolbert, Schaap-Hummel & Span (1986)	1985	462 lkr. groep 1 en 2	stratificatie op innovatiestatus	vragenlijst leerkrachten	*	*	*	*	—	—
4. Knuver & Reezigt (1991)	1990	56 lkr. groep 1 en 2	stratificatie op % zittendblijvers	vragenlijst leerkrachten	*	*	*	—	—	—
5. Harskamp & Willemsen (1991)	1990	33 lkr. groep 2/	regio Noord-Nederland	vragenlijst leerkrachten	—	—	*	—	—	*
		66 lln. groep 2 (2 lln per groep)	niveau (gem.)	rekentoets						
6. Van Gennip & Ruijs (1992)	1990/1991	20 lkr. groep 1 en 2	regio Nijmegen	observatie lkr (high inf), 1 ochtend per lkr	*	*	*	*	—	—
7. Vermeulen & Bosker (1992)	1991	45 lkr. groep 2	stratificatie op deeltijd/inzet lkr.	vragenlijst leerkrachten	*	*	—	*	—	*
		768 lln. groep 2	hele klas	5 taal- en rekentoetsen						

Tabel 1 (vervolg).

Tabel 1 (vervolg).

auteurs	data uit	aantal leerkr./leerl.	steekproef	wijze van data-verzameling	componenten effectiviteitsmodel waar informatie over wordt gegeven:					
					leerkracht leerling					
					instr	diff	mat	gel	betr	cogn
8. Harskamp, Pijl & Snippe (1992)	1991	65 lkr. groep 2	regio Noord-Nederland vragenlijsten, logboeken	observatie lkr (low inf) en klas, 1 schooldag;	—	—	—	*	*	—
9. Reezigt, Guldemond & Ros (1993)	1991	5 lkr. groep 1 en 2	regio Groningen	observatie lkr (low inf) en klas, 6 ochtenden; vragenlijsten, logboeken, interviews	*	*	—	*	*	—
10. Schonewille & van der Leij (1993)	1992	45 lkr. groep 1 en 2	stratificatie op etnische herkomst lln	observatie lkr (low inf) en klas, 4 ochtenden	—	—	—	*	*	—
11. Sontag & Meijnen (1993)	1992	zie boven	zie boven	zie boven, + logboeken en interviews	*	—	—	—	—	—
12. Inspectie (1994)	1992	80 lkr. groep 1 en 2	aselect	observatie lkr (high inf), 1 werkles per lkr	*	*	*	—	—	—
13. Van der Grift (1994)	1992	383 lkr. groep 1 en 2	aselect	observatie lkr (high inf), 2 werklessen per lkr	*	*	*	*	—	—
aantal onderzoeken dat informatie levert per component:					10	9	6	8	3	2

in twee (5, 7). Hierbij moet worden aangetekend dat in het onderzoek van Schonewille en Van der Leij (1993) en Sontag en Meijnen (1993) ook data op dit terrein worden verzameld, waarover nog nader gepubliceerd zal worden.

In de volgende paragrafen wordt een overzicht gegeven van resultaten uit de geselecteerde onderzoeken over het gedrag van de leerkracht, de gelegenheid tot leren en de betrokkenheid van leerlingen en het cognitief functioneren van leerlingen. Daarna wordt aandacht besteed aan factoren op het niveau van de school, i.c. de mate waarin het onderwijs van leerkrachten in de onderbouw van de basisschool overeenstemt. Ten slotte worden de resultaten enerzijds afgezet tegen de 'verworvenheden' die het onderwijs in groep 1 en 2 volgens velen bezit en anderzijds

tegen gegevens uit literatuur op het terrein van onderwijseffectiviteit. Daarna volgen conclusies en voorstellen voor verder onderzoek.

3. DE LEERKRACHT

Op het niveau van de leerkracht worden, uitgaande van het beschrijvingskader (zie figuur 1) de componenten instructie, differentiatie en materiaal onderscheiden. Deze drie componenten hebben, afzonderlijk en in samenhang, invloed op de gelegenheid tot leren en de leertijd van leerlingen.

3.1 Instructie

Binnen de component instructie worden door Creemers (1994) 11 verschillende effectiviteitsindicatoren onderscheiden: management en een ordelijke atmosfeer, huiswerk ofwel verlenging van de leertijd door taken mee naar huis te geven, hoge verwachtingen, duidelijkheid van doelstellingen, structureren van de inhoud van het onderwijs, heldere presentatie, vragen stellen, leerstof onmiddellijk inoefenen, registratie en evaluatie, feedback naar leerlingen en correctief/remedieel onderwijs aan leerlingen die dat nodig hebben.

De geselecteerde onderzoeken bevatten geen van alle gegevens over het meegeven van taken of werkjes naar huis, met als doel de leertijd te verlengen. Het ligt niet voor de hand in een groep 1 en 2 huiswerk, zoals dat in de bovenbouw van de basisschool voorkomt, te verwachten. Het is echter wel mogelijk dat leerkrachten alle leerlingen of alleen zwakke leerlingen werkjes meegeven om extra te oefenen, in overleg of onder begeleiding van de ouders. Of dit daadwerkelijk gebeurt en om wat voor soort werkjes het dan gaat is niet bekend. Er zijn evenmin onderzoeksgegevens over de verwachtingen die leerkrachten van groep 1 en 2 van hun leerlingen hebben wat hun cognitief functioneren betreft, over de wijze van vragen stellen en over de wijze waarop leerstof of begrippen worden ingeoefend.

Over de andere indicatoren zijn wel gegevens beschikbaar. Per indicator volgt een korte beschrijving van de onderzoeksresultaten.

Management/ordelijke atmosfeer

Of een leerkracht in staat is een klas goed te managen blijkt in eerste instantie uit de tijd die nodig is voor procedurele zaken en ordehandhaving; gegevens hieromtrent komen in paragraaf 4 aan de orde. Een ordelijk klimaat in de klas is nodig om leerlingen in staat te stellen geconcentreerd te werken. Sontag en Meijnen (1993) onderzochten bij 41 leerkrachten door observaties hoe het klimaat in hun klas te kenschetsen was. Criteria voor een ordelijk klimaat waren een zekere structuur in de activiteiten van de klas; daarnaast werd beoordeeld of de leerkracht overzicht houdt over alle activiteiten, volgt wat alle leerlingen doen, aan de leerlingen duidelijk maakt wat er moet gebeuren en aangeeft wanneer de ene activiteit ophoudt en een volgende begint. Een deel van de leerkrachten (15 %) wordt als chaotisch gekenschetst: ze hebben ordeproblemen, leerlingen werken niet geconcentreerd en weten niet altijd wat er van hen verwacht wordt, de activiteiten zijn niet goed voorbereid en er zijn relatief veel kinderen die niets doen, schreeuwen of vechten. Ongeveer een derde van de leerkrachten (29 %) bereikt een zeer ordelijk klimaat: leerlingen werken rustig en geconcentreerd, zelfs als de leerkracht even uit de klas is, en de activiteiten zijn goed voorbereid. De meeste leerkrachten (56 %) zitten tussen de beide andere groepen in. Er is geen samenhang tussen de mate van ordeijkheid en de mate van gezelligheid in de klas (zoals ingeschat door de observatoren).

Duidelijkheid van doelstellingen

Wat doelstellingen betreft is niet alleen van belang of leerkrachten duidelijk voor ogen hebben wat ze met hun leerlingen proberen te bereiken, maar ook dat ze niet teveel verschillende doelstellingen nastreven en nadruk leggen op enkele basisvaardigheden. Als aan leerkrachten van groep 1 en 2 gevraagd wordt welke doelen ze nastreven of welke onderwijsidealen ze

proberen te bereiken, worden meestal heel globale doelen genoemd in de pedagogische sfeer: kinderen maximale ontplooiingskansen bieden door een breed vormingsaanbod (Reezigt, Van Dijk & Bosveld, 1986), een goede basis leggen voor de verdere schoolloopbaan (Reezigt, Guldemon & Ros, 1993), zorgen dat kinderen zich veilig en prettig voelen op school, zorgen dat kinderen met elkaar leren omgaan en dat ze zelfstandig leren werken (Van Gennip & Ruijs, 1992). De meeste leerkrachten van groep 1 en 2 vinden de ontwikkeling van het cognitief functioneren ook wel van belang, maar niet op de eerste plaats (Van Gennip & Ruijs, 1992). De sociaal-emotionele en de creatief-motorische ontwikkeling worden belangrijker gevonden (Sontag & Meijnen, 1993; zie ook de bijdrage van Stevens en Harskamp aan dit themanummer). Ongeveer hetzelfde blijkt ook uit scores van leerkrachten van groepen 1 en 2 op schalen voor gerichtheid op het cognitief functioneren, die nooit in de hogere regionen van de schaal liggen (gemiddelde score 2 op een schaal die loopt tot 4, Knuver & Reezigt, 1991; gemiddelde score 15 op een schaal die loopt tot 36, Vermeulen & Bosker, 1992). Waarschijnlijk mede door het globale karakter van de doelen, dat het moeilijk maakt om richtlijnen voor de dagelijkse gang van zaken af te leiden, stimuleert slechts ongeveer de helft van de leerkrachten de ontwikkeling van leerlingen doelgericht (Inspectie, 1994a).

Op het niveau van de school bestaan evenmin specifieke doelen voor groep 1 en 2: in 60 % van de schoolwerkplannen zijn geen leer- en ontwikkelingsdoelen beschreven voor groep 1 en 2 en ook een beschrijving van het leer- en vormingsaanbod blijft vaak achterwege (CEB, 1994a). Het is dus niet zo dat leerkrachten van groep 1 en 2 globale doelen hanteren terwijl op een hoger niveau wel specifieke doelen geformuleerd zijn.

Structureren van de inhoud van het onderwijs

Structureren van het onderwijs kan, over afzonderlijke lessen heen, tot uiting komen in een planning van leerstof voor de langere termijn en daarnaast, in iedere afzonderlijke les, in het aansluiten bij voorkennis van leerlingen en in het terug- en vooruitblikken op leerstof of onderwerpen. Structurering voorkomt dat onderwerpen of leerstof uit de lucht komen vallen.

De meeste leerkrachten van groep 1 en 2 werken aan de hand van een jaarplanning (Van Gennip & Ruijs, 1992). In de jaarplanning wordt doorgaans aangegeven welke thema's op welk moment aan de orde moeten komen, welke projecten er in de loop van een jaar plaatsvinden en welke materialen aangeboden zullen worden. Het gaat meestal om niet meer dan zeer open, globale richtlijnen die geen duidelijke sturing geven aan de dagelijkse praktijk. Volgens Van Gennip & Ruijs (1992) is er soms in de praktijk niet veel verschil tussen het wel of niet volgens een jaarplan werken. Leerkrachten werken soms ook met planningen voor een week of een dag, maar dan gaat het er vooral om welke opdrachten leerlingen moeten maken en welke materialen daarbij klaargezet moeten worden.

Op het niveau van afzonderlijke lessen is het niet gebruikelijk dat leerstof of thema's gekoppeld worden aan leerstof die al aan de orde is geweest of aan leerstof die nog aan de orde zal komen (Van Gennip & Ruijs, 1992).

Heldere presentatie

Zonder een heldere presentatie van leerstof of begrippen en zonder een duidelijke uitleg bij taken en opdrachten zullen leerlingen al snel in de knoop komen met taken die ze moeten maken. In hoeverre leerkrachten van groep 1 en 2 er in slagen informatie goed voor het voetlicht te krijgen is niet bekend. Wel is onderzocht of leerkrachten gebruik maken van passend materiaal, een handelwijze die doorgaans geschikt voor deze leeftijdsgroep wordt geacht. Volgens Van der Meer, Appelhof, Ten Doesschate en Kramer (1986) maakt ongeveer de helft van de leerkrachten gebruik van aanschouwelijk materiaal bij het klassikaal uitleggen van nieuwe begrippen. Volgens inspecteurs die werklessen en muzisch-creatieve lessen observeerden hebben leerkrachten hun instructie meestal voorbereid en maken ze regelmatig gebruik van passend materiaal (Inspectie, 1994b). In hoeverre leerkrachten tijdens het geven van instructie checken of de leerlingen alles nog kunnen volgen en op basis daarvan instructie aanpassen of herhalen is niet bekend.

Harskamp, Pijl en Snippe (1992) gingen na hoeveel tijd de leerkracht tijdens een speelwerkles besteedt aan het aandacht geven aan leerlingen, en om wat voor aandacht het dan gaat: algemeen stimulerende aandacht of directe sturing van werkzaamheden waar leerlingen mee bezig zijn. Directe sturing geeft aan dat een leerkracht een leerling instrueert over taken of werkjes. Meer dan de helft van de tijd van de speelwerkles (54 %) is de leerkracht bezig met andere zaken dan leerlingen aandacht geven (management 30 %, afwezig/geen specifieke aandacht voor de klas 14 %, rondlopen 10 %). De tijd die leerkrachten wel besteden aan aandacht geven aan leerlingen betreft vooral algemene stimulering van leerlingen (40 %) en in veel mindere mate directe sturing (6 %). Ook Schonewille en Van der Leij (1993) onderzochten (op individueel niveau) het soort aandacht dat leerlingen krijgen. Een individuele leerling krijgt 2 tot 3 % van de tijd aandacht van de leerkracht; praktisch altijd gaat het daarbij om algemeen stimulerende aandacht. Leerkrachten treden dus in het algemeen maar zelden direct sturend op als leerlingen eenmaal aan het werk zijn.

Registratie en evaluatie

Om goed te kunnen volgen in hoeverre de leerlingen vorderen en om eventueel aanpassingen in het onderwijs tot stand te brengen is het regelmatig registreren en evalueren van leerlingvorderingen noodzakelijk, naast het doelmatig gebruiken van evaluatiegegevens.

De meeste leerkrachten registreren met een zekere regelmaat wat de leerlingen doen; slechts 24 % geeft aan dat niet systematisch te doen (Reezigt, Van Dijk & Bosveld, 1986). Het gaat echter in de meeste gevallen alleen om een registratie van het onderwijsaanbod, zoals bijvoorbeeld het in een klassenschrift bijhouden van wie in welke hoek heeft gespeeld, wie met welk materiaal heeft gewerkt e.d. (Van Gennip & Ruijs, 1992; Reezigt, Guldmond & Ros, 1993; Inspectie, 1994b). Knuver en Reezigt (1991) melden de volgende percentages:

- 90 % van de leerkrachten schrijft regelmatig op wat er gedaan is,
- 59 % van de leerkrachten schrijft regelmatig op of een activiteit of een werkje goed of slecht is verlopen,
- 60 % van de leerkrachten schrijft regelmatig op of er problemen optraden bij activiteiten of werkjes.

De beide laatste percentages stemmen overeen met recente gegevens van de Inspectie (1994a): 60 % van de leerkrachten registreert in voldoende mate hoe het met activiteiten en vorderingen is gesteld. Daarnaast registreert 61 % ook gegevens over het gedrag van leerlingen. Gegevens worden vastgelegd in logboeken, leerlingklappers, observatieformulieren of mappen waarin voor ieder kind bepaalde werkjes worden bewaard (Van Gennip & Ruijs, 1992). Niet alle activiteiten worden overigens even vaak geregistreerd. Registratie van activiteiten tijdens werklessen komt vaker voor dan registratie van activiteiten tijdens muzisch-creatieve lessen (Inspectie, 1994b).

Het regelmatig evalueren van leerlingen door gestandaardiseerde toetsen of systematische observatie is niet gebruikelijk. Ongeveer de helft van de leerkrachten geeft aan dergelijke activiteiten uit te voeren, bijvoorbeeld aan het begin van het nieuwe schooljaar (Reezigt, Van Dijk & Bosveld, 1986) of in de loop van het jaar (Wolbert, Schaap-Hummel & Span, 1986). Hetzelfde percentage wordt vermeld door Van Gennip en Ruijs (1992). Daarbij zijn dan wel alle soorten toetsen meegerekend, ook werkbladen en zelfgemaakte toetsbladen. Standaardtoetsen van het CITO of soortgelijk materiaal worden bijna nooit gebruikt, toetsen uit op school aanwezige methodes zelden. Observeren, een activiteit die veel leerkrachten wel eens uitvoeren, vindt niet met grote regelmaat plaats: 74 % van de leerkrachten doet dat minder dan eens per maand (Knuver & Reezigt, 1991). De vraag is dan ook nog of het wel altijd gaat om systematisch observeren: gegevens van de Inspectie (1994a) wijzen uit dat leerkrachten vooral gegevens verzamelen door informele observaties (74 % van de leerkrachten) en minder met behulp van voorgestructureerde screeningslijsten (61 %).

Het is niet duidelijk in hoeverre registratie van activiteiten in de klas en observaties gegevens over individuele leerlingen opleveren; het is ook denkbaar dat leerkrachten zich daarbij meer op de klas als geheel richten en bijvoorbeeld vooral bekijken in hoeverre een bepaalde activiteit organisatorisch gezien een succes is. De laatste jaren zijn in het kader van de ontwikkeling van leerlingvolgsystemen diverse toetsen voor de leerjaren 1 en 2 ontwikkeld, die zijn afgestemd op toetsen die in hogere leerjaren worden gebruikt. Gegevensverzameling op individueel niveau in een dergelijk systematisch kader heeft een meerwaarde boven verzamelen van gegevens op het niveau van de klas en het incidenteel gebruik van los van elkaar staande toetsen of observatielijsten. Volgens Sontag en Meijnen (1993) is er op de meeste scholen al wel een leerlingvolgsysteem aanwezig vanaf groep 3, maar gebruiken de leerkrachten van groep 1 en 2 in meerderheid hun eigen materiaal. Ook op scholen waar een leerlingvolgsysteem vanaf groep 1 aanwezig is worden de materialen voor groep 1 en 2 nog niet vanzelfsprekend door leerkrachten gebruikt. Er is ook een kleine groep leerkrachten die helemaal niets meer registreert, in afwachting van de afronding van het volgsysteem op hun school.

Wat het gebruik van registratie- en evaluatiegegevens betreft, meldt de Inspectie (1994a) dat bij 47 % van de leerkrachten van een 'doelmatig' gebruik sprake is, d.w.z. gebruik voor diagnose, evaluatie, planning en reflectie op het eigen handelen. Een grotere groep leerkrachten (76 resp. 70 %) gebruikt de gegevens vooral om ouders over hun kinderen in te lichten en om de leerkracht van de volgende groep op de hoogte te stellen. Gericht inzet van gegevens voor het volgen van het ontwikkelingsverloop van kinderen of het signaleren van uitval op bepaalde terreinen komt minder frequent voor (44 resp. 56 % van de leerkrachten). Vergelijkbare resultaten worden vermeld door Van Gennip & Ruijs (1992).

Feedback naar leerlingen

Door feedback te geven over activiteiten en werkjes kunnen leerkrachten aan hun leerlingen duidelijk maken hoe ze het hebben gedaan en wat er anders of beter zou kunnen. Onderzoeksresultaten geven aan dat dit in groepen 1 en 2 geen gebruikelijke praktijk is. Volgens de Inspectie (1994b) laten leerkrachten wel met enige regelmaat merken of activiteiten wel of niet goed lopen; het is echter niet duidelijk of het om individuele feedback gaat of om opmerkingen tegen de hele klas. Leerkrachten zelf noemen het geven van feedback tijdens werklessen nauwelijks als een activiteit die ze regelmatig uitvoeren (Van Gennip & Ruijs, 1992). Uit het onderzoek van Knuver en Reezigt (1991) blijkt dat een kwart van de leerkrachten aangeeft geen feedback aan leerlingen te geven; de resterende leerkrachten zeggen in meerderheid (61 %) dat ze leerlingen wel eens prijzen als ze een werkje af hebben of ze een stempeltje of sticker geven, maar dat ze niet speciaal vertellen hoe ze het hebben gedaan, wat anders of beter had gekund e.d. Volgens Reezigt et al. (1993) hebben interacties tussen leerkrachten en individuele leerlingen maar zeer zelden betrekking op feedback over een bepaalde prestatie.

Correctief/remedieel onderwijs

Vooral in de eerste leerjaren van de basisschool is het, om groeiende problemen en achterstanden te voorkomen, van belang dat leerkrachten of anderen (zoals remedial teachers) op tijd correctief/remedieel onderwijs geven. Idealiter zou een school ook in staat moeten zijn om leerlingen die juist ver voor zijn op anderen een passend onderwijsaanbod te bieden.

Volgens de Inspectie (1994b) is de situatie op de meeste scholen nog niet optimaal te noemen. De meeste scholen kunnen wel in staat worden geacht problemen bij jonge leerlingen op te sporen (78 %), maar adequaat hulp bieden is minder vanzelfsprekend (66 % van de scholen). Er zijn wel voorzieningen voor extra ondersteuning voor kinderen met leermoeilijkheden (76 % van de scholen), maar niet voor hoogbegaafde jonge leerlingen (22 %). Ook het opsporen van en hulp bieden aan hoogbegaafde jonge kinderen is een probleem: slechts 28 resp. 23 % van de scholen wordt daartoe in staat geacht.

Samenvatting

Samenvattend blijkt uit het overzicht van de gegevens bij de afzonderlijke indicatoren dat er nog maar weinig details bekend zijn over de manier waarop leerkrachten in groep 1 en 2 leerstof aanbieden en opdrachten geven en over de mate waarin ze tijdens activiteiten in de gaten houden of leerlingen wel begrijpen wat er van hen wordt verwacht. Duidelijk is wel dat leerkrachten eerder als algemeen stimulerend dan als direct sturend te kenschetsen zijn. Dit impliceert dat er voor leerlingen ruimte overblijft om een eigen invulling aan activiteiten te geven. Onbekend is echter in hoeverre leerlingen altijd in staat zijn zelf voor een zinvolle invulling te zorgen.

De meeste leerkrachten hebben een redelijk ordelijk klimaat in hun klas, maar ze verschillen vrij sterk van elkaar op dit aspect. Opvallend is dat een zeer ordelijk klimaat niet automatisch samengaat met een onaangename sfeer, een associatie die vaak wordt gelegd. De dagelijkse praktijk in groep 1 en 2 wordt niet gestuurd door concrete, richtinggevendende doelen. Leerkrachten streven zeer globale doelen na op basis van globale jaarplannen, die eerder een ordening van thema's en onderwerpen behelzen dan een planning van leerstofinhouden of begrippen. Hoewel praktisch alle leerkrachten wel op enigerlei wijze registreren wat er in de klas is gebeurd en aan de orde is geweest, is niet duidelijk wat voor gegevens op het niveau van de individuele leerling worden verzameld en hoe ze worden gebruikt. Feedback over activiteiten op individueel niveau is eerder uitzondering dan regel. De meeste scholen hebben wel enige voorzieningen voor de opvang van zwakke leerlingen, maar niet voor goede of zeer goede leerlingen.

3.2 Differentiatie

Als kinderen op vier-jarige leeftijd aan de basisschool beginnen verschillen ze onderling al sterk. Verhoeven en Van Kuyk (1991) onderzochten de conceptuele en metalinguïstische kennis van 298 kinderen bij hun intrede in het basisonderwijs. Op sommige toetsen haalt de ene vierjarige een nul-score, terwijl een ander met gemak het maximum bereikt. Soortgelijke resultaten zijn ook gevonden voor rekentaken (Van Kuyk, 1991a). Sommige vierjarigen hebben praktisch alles wat gemiddeld genomen in groep 1 en 2 wordt aangeboden al onder de knie als ze op school komen, terwijl sommige zesjarigen aan het eind van groep 2 nog niet halverwege dat niveau komen. Er is dus, uitgaande van deze situatie, alle reden om in groep 1 en 2 te differentiëren. De noodzaak tot differentiatie stijgt nog als rekening gehouden moet worden met verschillen in etnische herkomst en daarmee samenhangende verschillen in bijvoorbeeld de aanspreekbaarheid van kinderen en hun woordenschat (Damhuis & De Gloppe, 1993).

Leerkrachten van groep 1 en 2 worden vaak ten voorbeeld gesteld aan leerkrachten van hogere leerjaren als het gaat om het realiseren van differentiatie in de klas. Men gaat er praktisch altijd van uit dat leerkrachten van groep 1 en 2 goed in staat zijn te differentiëren en te zorgen voor een onderwijsaanbod dat aansluit bij de individuele leerling (zowel wat diens niveau als interesse betreft), omdat ze gewend zijn flexibel te werken en minder strak vasthouden aan frontaal-klassikaal lesgeven. Verschillende onderzoeken wijzen uit dat leerkrachten van groep 1 en 2 inderdaad positiever denken over differentiatie (ze zien het minder als een problematische vernieuwing, meer als iets dat er vanzelfsprekend bij hoort) en ook vrij consequent hoger scoren op schalen die de mate van differentiatie meten dan de leerkrachten uit hogere leerjaren. Leerkrachten van groep 1 en 2 noemen voornamelijk klassegrootte en tijdgebrek als belemmerende factoren voor differentiatie, terwijl leerkrachten uit hogere leerjaren ook relatief vaak klagen over ongeschikt materiaal, gebrek aan ruimte en het feit dat ze aan een combinatieklas moeten lesgeven (dat juist in 1 en 2 vaak als bevorderende factor wordt gezien; Reezigt et al., 1986; Van Gennip & Ruijs, 1992). Vermeulen en Bosker (1992) melden een gemiddelde score van 23 op een differentiatieschaal met een maximum van 35 voor leerkrachten van groep 1 en 2 en een score van 13 voor leerkrachten uit groep 3; Reezigt et al. (1986) melden een gemiddelde score van 70 op een schaal met een maximum van 110 voor leerkrachten van groep 1 en 2 en een score van 53 voor leerkrachten uit hogere leerjaren. Deze schalen geven voornamelijk aan in hoeverre leerkrachten toestaan dat leerlingen uit hun klas op hetzelfde moment aan verschillende taken werken; dit is in groep 1 en 2 dagelijks het geval als de leerlingen tijdens de werkles in verschillende hoeken bezig zijn met verschillende materialen en

activiteiten. De uitspraak van de Inspectie (1994a), dat bij 61 % van de leerkrachten sprake is van een veelzijdig en gedifferentieerd aanbod, heeft op deze situatie betrekking. Klassikaal werken, de gebruikelijke gang van zaken in groep 3 en hoger, is in groep 1 en 2 niet dominant. In de meeste klassen variëren de groeperingsvormen sterk, maar overheersen de perioden waarin kinderen individueel werken (Van der Meer et al., 1986). Differentiatie modellen die leerkrachten in hogere leerjaren gebruiken (basisstof-herhalingsstof-verrijkingsstofmodel, werken in niveaugroepen, individueel werken) komen in groep 1 en 2 niet op dezelfde wijze voor (Bosveld, Van Dijk, Reezigt & Weide, 1988).

De positieve houding van leerkrachten ten opzichte van differentiatie, de relatief korte perioden van klassikaal werken en de toegestane variatie in materialen en activiteiten zegt op zich echter nog niet veel over het differentiërend handelen van de leerkracht. Om een goed beeld te krijgen van de differentiatiepraktijk in groep 1 en 2 is informatie nodig over de manier waarop leerkrachten doelbewust het onderwijs afstemmen op het niveau of de interesse van de individuele leerling. Van belang in dit verband zijn de vragen:

- in hoeverre streven leerkrachten naar het bereiken van bepaalde doelen door middel van differentiatie op basis van het niveau of de interesse van leerlingen;
- verzamelen leerkrachten gegevens over het niveau of de interesse van leerlingen, met de bedoeling op basis daarvan een aangepast onderwijsaanbod te verzorgen;
- in hoeverre bieden leerkrachten daadwerkelijk een naar niveau of interesse aangepast onderwijsaanbod aan?

In het voorgaande is al beschreven dat leerkrachten van groep 1 en 2 in het algemeen globale doelen formuleren. Dat geldt ook als gevraagd wordt wat ze met differentiatie proberen te bereiken. Volgens Wolbert et al. (1986) en Van Gennip en Ruijs (1992) zeggen leerkrachten dat ze differentiatie voornamelijk toepassen om de leerling in zijn totaliteit tot zijn recht te laten komen; differentiatie wordt niet gezien als een mogelijkheid om de effectiviteit van het onderwijs te vergroten (door bijvoorbeeld met kinderen die op een bepaald punt uitvallen extra te oefenen).

Het is, zoals al is gebleken, niet gebruikelijk dat leerkrachten leerlingen observeren, gerichte vragen stellen of toetsen, en het is evenmin gebruikelijk dat ze dat doen met als doel het aanbod voor die leerling aan te passen. Als kinderen op vierjarige leeftijd de school binnenkomen gebeurt dat evenmin (Knuver & Reezigt, 1991). Praktisch alle leerkrachten vinden het weliswaar belangrijk om te weten waar de interesse van leerlingen ligt, maar voorzover ze daar actief achter proberen te komen gaat het doorgaans om interesses van de hele klas waar aandacht aan besteed wordt tijdens een kringgesprek (Reezigt et al., 1986).

In het onderwijs inspelen op de interesses van kinderen kan onder andere door kinderen tijdens de kring te laten vertellen wat hen bezighoudt, kinderen in die hoeken te laten werken die ze zelf kiezen (met een zekere beperking zodat niet altijd hetzelfde wordt gekozen) en kinderen te laten kiezen uit opdrachten (Reezigt et al., 1986). Volgens de Inspectie (1994a) hebben leerkrachten niet altijd voldoende oog voor de ervaringen van kinderen (44 % van de leerkrachten wordt als voldoende op dit aspect beoordeeld), maar is er in de meeste klassen wel voldoende gelegenheid voor kinderen om zelf keuzen te maken en initiatieven te nemen (63 % van de klassen). Leerkrachten houden bij het aanbieden van materialen ook rekening met de belangstelling van kinderen (Inspectie, 1994b). Rekening houden met de interesse van kinderen leidt er echter zelden of nooit toe dat er in het onderwijsaanbod iets verandert: de leerkrachten wijken praktisch nooit van hun geplande activiteiten af (Van Gennip & Ruijs, 1992; Reezigt et al., 1993).

Hoewel leerkrachten van groep 1 en 2 inspelen op niveauverschillen van belang vinden, geven ze in meerderheid aan dat alle kinderen in hun klas in principe hetzelfde programma volgen (Reezigt et al., 1986). Meestal werkt niet de hele klas op hetzelfde moment aan dezelfde opdracht, maar doorgaans moeten wel alle kinderen volgens een rouleersysteem dezelfde opdracht maken (Wolbert et al., 1986). Omdat kinderen binnen een klas volgens leerkrachten sterk verschillen in bijvoorbeeld algemeen niveau, taalgebruik en zelfstandigheid (Reezigt et al.,

1993), worden niet aan alle leerlingen dezelfde eisen gesteld. Als leerlingen opdrachten tijdens werklessen maken (het gaat dan bijvoorbeeld om vouw-, plak- en knipwerkjes, knutselwerkjes of werkbladen), worden aan sommige kinderen hogere eisen gesteld dan aan andere op het gebied van zelfstandig werken, de moeilijkheidsgraad van de opdracht, het tempo waarin de opdracht moet worden afgemaakt of het aantal opdrachten (Knuver & Reezigt, 1991). Leerlingen die zwakker worden ingeschat krijgen soms extra beurten of makkelijker beurten in de kring (Reezigt et al., 1993). Op grond waarvan leerkrachten besluiten van het ene kind meer te eisen dan van het andere is niet duidelijk, hoewel de leeftijd een rol speelt (Van Gennip & Ruijs, 1992): oudere leerlingen krijgen vaker opdrachten dan jongere, jongere leerlingen mogen vaker zelf kiezen wat ze willen doen. De hoeken in de klas worden niet als differentiatiemaatregel ingezet; leerkrachten geven maar zelden aan dat taalzwakke kinderen bijvoorbeeld extra vaak in de leeshoek of de luisterhoek werken (Reezigt et al., 1993).

Differentiatie in de betekenis van rekening houden met niveauverschillen tussen leerlingen betekent in groep 1 en 2 dan ook vooral het accepteren van niveauverschillen en niet zozeer het actief en doelgericht inspelen op die verschillen, eventueel met het oogmerk veranderingen tot stand te brengen.

3.3 Materiaal

Wat materiaal in groep 1 en 2 betreft moet onderscheid gemaakt worden tussen de materialen waar leerlingen min of meer zelfstandig mee kunnen werken (speelleermateriaal) en materiaal dat de leerkracht gebruikt bij het verzorgen van onderwijs (programma's en methodes).

Het materiaal waar leerlingen zelf mee werken zonder begeleiding van de leerkracht is meestal ondergebracht in de hoeken binnen de klas. Gemiddeld heeft een klas 5 hoeken (Van Gennip & Ruijs, 1992); de meest voorkomende zijn de poppenhoek, de bouwhoek, de lees- en boekenhoek, de expressiehoek (tekenen, verven) en de zand- en watertafel (Van Gennip & Ruijs, 1992; Harskamp et al., 1992). Taal/luisterhoeken en hoeken met rekenmaterialen of rekenspelletjes komen veel minder frequent voor. Ontwikkelings-materialen zoals lotto's en puzzels kunnen meestal niet door kinderen zelf gepakt worden. Naar de mening van de Inspectie (1994a) heeft ongeveer tweederde van alle klassen hoeken van voldoende kwaliteit. Er is veelzijdig leer- en ontwikkelingsmateriaal, dat toegankelijk en ordelijk is opgesteld.

Leerkrachten hebben vaak wel een programma voor algemene cognitieve ontwikkeling in bezit (73 %) en methodisch materiaal voor schrijven (50 %). Methoden of programma's op andere, meer specifieke gebieden (taal, rekenen, muzisch-creatieve ontwikkeling, sociaal-emotionele ontwikkeling) zijn veel minder algemeen verspreid (35-44 %). De laagste verspreidingsgraad hebben methodes en programma's voor lezen en motorische ontwikkeling (29 %; Inspectie, 1994b). Als gevraagd wordt wat er nu precies aanwezig is, worden vooral Taalactivering Malmberg, Doen Denken Durven en Praatspel genoemd, naast door de leerkracht zelf ontwikkeld materiaal (Knuver & Reezigt, 1991; Harskamp et al., 1993). De aanwezigheid van materialen of methoden zegt nog niets over het feitelijk gebruik door leerkrachten. Volgens Van Gennip en Ruijs (1992) wordt aanwezig materiaal overwegend als ideeënbron benut. Materiaal zoals Doen Denken Durven wordt door 25 % van de leerkrachten vrij strak volgens de aanwijzingen van de auteurs gebruikt, 50 % gebruikt gedeelten en de resterende 25 % ontleent er alleen ideeën aan (Wolbert et al., 1986).

Harskamp en Willemsen (1991) onderzochten het gebruik van programma's voor voorbereidend rekenen en speelleermateriaal op het gebied van rekenen. Een meerderheid van de leerkrachten zegt materiaal op dit terrein te bezitten (79 %); daarbij is vooral Kleuters Willen Ook Wel Eens Denken favoriet. Voorlopers van methoden voor groep 3 tot en met 8 (zoals Operator Rekenen en De Wereld in Getallen) worden veel minder vaak genoemd. Een deel van de leerkrachten (12 %) die wel een programma in huis hebben gebruikt het in de onderzochte periode (een kwartaal) niet. De overige leerkrachten werken er gemiddeld eens per twee weken mee en richten zich dan vooral op vergelijken van hoeveelheden en kennis van cijfers. Activiteiten rond meten, seriëren en werken met hoeveelheden zijn minder populair. Oefeningen op het terrein van erbij doen of eraf halen van hoeveelheden komen weinig voor. Speelleermateriaal op

het gebied van rekenen wordt 2 à 3 keer per week gebruikt om opdrachten aan groepjes kinderen te geven. Materiaal dat bedoeld is om aantallen te leren vergelijken, materiaal om te ordenen en materiaal gericht op cijferkennis worden het vaakst gebruikt. Leerkrachten maken vooral gebruik van de traditionele materialen van het kleuteronderwijs waar leerlingen individueel mee kunnen werken. Wat taalactiviteiten betreft zijn leerkrachten vooral gericht op het uitbreiden van de woordenschat, het laten luisteren naar klanken in woorden en het laten vertellen van verhalen bij prentenboeken of platen. Er wordt weinig gedaan aan het herkennen van letters, woorden of zinnen. Leerkrachten die programma's voor taalontwikkeling in hun bezit hebben bieden gemiddeld twee keer per week activiteiten daar uit aan (Harskamp et al., 1993).

Op de meeste scholen is materiaal voor verschillende leer- en vormingsgebieden aanwezig. Aanwezigheid van materiaal zegt nog niets over de mate en wijze van gebruik. Hoe vaak de leerkrachten materiaal gebruiken, hoe ze onderdelen selecteren, wanneer ze zelf materiaal maken en hoe ze dat precies doen is voor de meeste gebieden niet gedetailleerd bekend.

4. GELEGENHEID OM TE LEREN EN BETROKKENHEID VAN LEERLINGEN

Om na te gaan welke effecten onderwijs op leerlingen heeft, is het nodig te kijken naar de gelegenheid die leerlingen krijgen om bepaalde kennis of vaardigheden op te doen. Het gaat daarbij, in oplopende volgorde van specificiteit, om de tijd die volgens het lesrooster aan bepaalde gebieden moet worden besteed, om de tijd die leerkrachten er aan zeggen te besteden en om de tijd die er daadwerkelijk aan besteed wordt. In het laatste geval wordt de netto-onderwijstijd bedoeld die overblijft na correctie voor lesuitval, pauzes, tijd voor procedurele aspecten e.d. (zie verder 4.1).

De netto-onderwijstijd geeft aan welke gelegenheid leerlingen in de praktijk krijgen om iets te leren. De veronderstelling is dat leerlingen ook inderdaad op school aanwezig zijn en dat ze betrokken zijn bij hetgeen ze geacht worden te doen. Geen van beide veronderstellingen zijn vanzelfsprekend. Over de absentie van leerlingen in groep 1 en 2 zijn geen gegevens bekend. Absentie is, voorzover het om vierjarige leerlingen gaat die nog niet leerplichtig zijn, toegestaan; daarbij kan het voorkomen dat leerlingen alleen 's ochtends op school komen of op bepaalde dagen thuis blijven, bijvoorbeeld aan het eind van de week. Daarnaast zijn er aanwijzingen dat absentie bij vijfjarige leerlingen aanzienlijk kan zijn (Reezigt et al., 1993): in sommige groepen 1 en 2 zijn dagelijks enkele leerlingen afwezig (vaak dezelfde) en er zijn leerlingen die heel vaak te laat komen, snel thuis worden gehouden bij ziekte of om onbekende redenen niet verschijnen. De verschillen tussen leerlingen in de tijd die ze in groep 1 en 2 daadwerkelijk op school doorbrengen kunnen dus aanzienlijk zijn, nog afgezien van de verschillen in betrokkenheid tijdens de perioden dat ze wel aanwezig zijn (zie verder 4.2).

4.1. Gelegenheid om te leren

Verschillende onderzoeken geven informatie over de tijdsbesteding van leerkrachten in de groepen 1 en 2. De vergelijkbaarheid van de gegevens wordt bemoeilijkt omdat sommige onderzoeken alleen de roostertijd meten (Inspectie, 1994b) of alleen de tijd die leerkrachten zeggen aan bepaalde onderdelen te besteden (Wolbert et al., 1986; Vermeulen & Bosker, 1992), terwijl andere onderzoeken betrekking hebben op de feitelijk geobserveerde tijdsbesteding op hele of halve schooldagen of tijdens bepaalde lessen (Van der Meer et al., 1986; Van Gennip & Ruijs, 1992; Harskamp et al., 1992; Reezigt et al., 1993; Schonewille & Van der Leij, 1993). In het laatste geval zijn de geobserveerde activiteiten ingedeeld in door de onderzoekers bepaalde categorieën, in het eerste geval gaat het om benamingen die leerkrachten zelf voor hun werkzaamheden hanteren. Leerkrachten gaan daarbij meer uit van een indeling in werkvormen (kringgesprek, werkles, gymles, arbeid naar keuze, werken met ontwikkelingsmateriaal) dan van een indeling in leer- en vormingsgebieden. Zo wordt de tijdsindeling in het onderzoek van Vermeulen en Bosker dan ook gedomineerd door de categorie 'overige activiteiten', waar 90 % van alle

werkzaamheden in is ondergebracht; daarnaast worden nog geringe percentages genoemd voor rekenen en taal/lezen. Hierdoor wordt de interpretatie van gegevens die op leer- en vormingsgebieden betrekking hebben enigszins onbetrouwbaar. Bovendien kunnen leer- en vormingsgebieden gemakkelijk door elkaar lopen: tijdens een gymles kan bijvoorbeeld alleen lichamelijke oefening aan de orde komen, maar ook reken- en taalaspecten (spelletjes met teloefeningen, luisteroefeningen, liedjes met opdrachten e.d.).

Wat de roostertijd betreft is o.a. op basis van de gegevens van de Inspectie (1994b) door de CEB (1994a) aangegeven hoe de 22 beschikbare klokuren per week gemiddeld worden besteed; daarbij is geen onderscheid naar leerjaar 1 en leerjaar 2 aangebracht. Bijna de helft van de tijd gaat op aan zintuiglijke en lichamelijke oefening (40 %, 9 uur). De resterende tijd wordt in beslag genomen door expressie (25 %, 5 uur), taal (17 %, 4 uur) en kennisgebieden (14 %, 3 uur). Rekenactiviteiten nemen slechts 2 % (1 uur) van de tijd in beslag.

In de indeling van Wolbert et al. (1986) lopen leer- en vormingsgebieden en werkvormen door elkaar. Uitgaande van 22 uur wordt 7 uur besteed aan lichamelijke ontwikkeling, 4 uur aan opdrachten tijdens werklessen, 4 uur aan arbeid naar keuze, 2 uur aan taalactiviteiten, 2 uur aan muziek, 2 uur aan kringgesprekken en 1 uur aan onderwijsleergesprekken. Evenals bij de roostertijd het geval was, overheerst dus de tijd voor zintuiglijke/lichamelijke oefening. Ook bij Harskamp et al. (1992) en Van Gennip en Ruijs (1992) is dit het geval, terwijl in deze beide onderzoeken de tijd die specifiek voor voorbereidende taal- en rekenactiviteiten wordt gereserveerd eveneens zeer gering is (hooguit 2 uur per week voor de beide leer- en vormingsgebieden samen). Het onderzoek van Harskamp et al. maakt het mogelijk gegevens die leerkrachten zelf rapporteren (in logboeken) te vergelijken met gegevens van observatoren. Grote verschillen treden in de meeste onderscheiden categorieën niet op, maar juist wel in de categorie voorbereidende taal- en rekenactiviteiten: volgens leerkrachten zelf besteden ze er ongeveer 2 uur per week aan, maar de observatoren komen slechts tot krap 1 uur. Ook de categorie 'organisatie' leidt tot verschillen (volgens de leerkrachten 2 uur per week, volgens de observatoren 5 uur).

De onderzoeken die observatie-gegevens rapporteren maken het mogelijk om vast te stellen hoeveel tijd in groep 1 en 2 als onderwijstijd kan worden beschouwd, als tijd voor organisatie van activiteiten en als tijd voor ontspanning en verzorging (zie tabel 2).

Bij onderwijstijd is een ruime definitie gehanteerd. Hierbij gaat het om tijd besteed aan kringen en onderwijsleergesprekken, perioden van individueel werken, geven en uitleggen van opdrachten en gymlessen waarbij de leerkracht een actief instruerende rol speelt. Bij tijd voor organisatie gaat het om de tijd die gespendeerd wordt aan de overgang van de ene activiteit naar de andere, administratieve routines en wachttijd waarin in feite niets gebeurt. Tijd voor ontspanning en verzorging omvat inlooptijd aan het begin van de ochtend, buiten spelen, eten en drinken en vrije gymlessen waarbij de leerlingen mogen doen wat ze willen. Resttijd is de tijd die verloren gaat aan storingen van buiten af (telefoontjes in de klas, leerkrachten of leerlingen uit andere klassen die binnenkomen) en de tijd die de leerkracht niet in de klas aanwezig is.

De verschillende onderzoeken lopen wat onderwijstijd betreft uiteen van minimaal 44 % tot 60 %; daardoor verschillen ook de overige percentages. Naar alle waarschijnlijkheid zijn deze

Tabel 2. Daadwerkelijk bestede tijd in groep 1 en 2, in percentages.

	Van der Meer et al. (1986)	Harskamp et al. (1992)	Reezigt et al. (1993)
onderwijs	60	51	44
organisatie	13	22	14
ontspanning/ verzorging	26	27	40
resttijd	1	—	2

verschillen veroorzaakt door de betrekkelijk geringe steekproefgroottes. Omdat leerkrachten sterk van elkaar verschillen in hun tijdsbestedingspatroon kunnen de gemiddelden al snel uiteen gaan lopen.

De verschillende onderzoeken maken duidelijk dat de vakgebieden die vanaf groep 3 het onderwijs domineren niet of nauwelijks herkenbaar te traceren zijn in groep 1 en 2. Activiteiten die op deze vakgebieden betrekking hebben zijn ondergebracht in diverse werkvormen (werklessen, arbeid naar keuze, werken in hoeken). De meeste onderwijstijd gaat naar de "klassieke" kleuter-activiteiten zintuiglijke/lichamelijke oefening en expressie. Een groot deel van de tijd is niet onderwijsgericht. Ten slotte wordt een vrij aanzienlijk percentage tijd besteed aan organisatie van activiteiten.

4.2. Betrokkenheid van leerlingen

Naast de gelegenheid om te leren, die afhankelijk is van het gedrag en de tijdsbesteding van de leerkracht, is het nodig dat leerlingen betrokken zijn bij hetgeen ze geacht worden te doen. Het gaat daarbij niet alleen om opdrachten en taken die de leerkracht hen opgeeft; het gaat er ook om dat ze bijvoorbeeld in de tijd die ze in de bouwhoek mogen spelen ook daadwerkelijk met het materiaal in de bouwhoek aan de slag gaan en blijven.

Naar het oordeel van de Inspectie (1994a) zijn kinderen in 63 % van de groepen 1 en 2 voldoende betrokken op hun werk. Enkele observatiestudies leveren nadere informatie op dit punt. Volgens Reezigt et al. (1993) is gemiddeld over leerkrachten en activiteiten 75 % van de kinderen betrokken bij wat ze doen; 13 % van de leerlingen is niet betrokken, 11 % wacht op nadere instructie of hulp van de leerkracht en 1 % is uit de klas (bijvoorbeeld om iets uit een ander lokaal te halen). De verschillen van klas tot klas zijn aanzienlijk: de percentages betrokken leerlingen lopen uiteen van 69 tot 82. De betrokkenheid is ook afhankelijk van het soort activiteit dat aan de orde is (zie Tabel 3).

Bij klassikale activiteiten of bij individueel werk met begeleiding van de leerkracht zijn relatief de meeste leerlingen betrokken. De percentages betrokken leerlingen zijn het laagst als de leerkracht opdrachten geeft of als de ene activiteit in de andere overgaat. Dat laatste is begrijpelijk, omdat leerlingen op dergelijke momenten al snel afgeleid kunnen worden of ergens op moeten wachten. De geringe betrokkenheid tijdens opdrachten geven door de leerkracht is problematischer: als leerlingen juist dan niet opletten is de kans groot dat ze later, tijdens perioden van zelfstandig werk, niet weten wat er van hen wordt verwacht.

Harskamp et al. (1992) onderzochten de mate van betrokkenheid van leerlingen tijdens spel- en werksituaties; 73 % van de leerlingen is tamelijk of zeer intens bij het spel betrokken, terwijl 27 % niet of slechts oppervlakkig betrokken is. De percentages in werksituaties verschillen nauwelijks van de percentages in spelsituaties. Het gedrag van de leerkracht is van invloed op de mate van betrokkenheid: algemeen stimulerend en direct sturend gedrag correleren beide positief (.20 resp. .24), terwijl het percentage tijd dat leerkrachten besteden aan organisatie en management negatief correleert (-.34).

Tabel 3. Percentages betrokken leerlingen bij verschillende activiteiten.

groepsgesprek	88
vertellen in de kring	84
individueel werk (leerkracht begeleidt)	81
eten en drinken	74
individueel werk (leerkracht begeleidt niet)	74
leerkracht geeft opdrachten	65
overgangen tijdens lessen	57

Schonewille en Van der Leij (1993) vergeleken de mate van betrokkenheid van leerlingen tijdens groepswork of klassikale situaties en tijdens individueel werk. Leerlingen zijn in hogere mate betrokken bij groeps- en klassikaal werk (76 % tamelijk of zeer intens betrokken) dan bij individueel werk (66 % tamelijk of zeer intens betrokken). Ook in dit onderzoek zijn positieve correlaties gevonden van algemeen stimulerend en direct sturend gedrag van de leerkracht en de mate van betrokkenheid van leerlingen (.18 resp. .19).

Hoewel de betrokkenheid van leerlingen niet in alle onderzoeken op dezelfde wijze is gemeten, wijzen de resultaten steeds dezelfde richting uit. In alle onderzoeken is ongeveer 75 % van de leerlingen in redelijke mate betrokken bij werk of spel. Naarmate de leerkracht actiever optreedt stijgt de betrokkenheid van de leerlingen. De betrokkenheid is relatief laag als leerlingen individueel moeten werken zonder direct contact met de leerkracht.

5. EFFECTEN VAN HET ONDERWIJS OP HET COGNITIEF FUNCTIONEREN VAN LEERLINGEN

Zoals tabel 1 al liet zien, zijn er nauwelijks onderzoeksgegevens beschikbaar die uitspraken over effecten van onderwijs op het cognitief functioneren van jonge kinderen mogelijk maken. Slechts twee onderzoeken leveren enige informatie op dit punt. In het onderzoek van Vermeulen en Bosker (1992) zijn verschillende taal- en rekentoetsen afgenomen bij 768 leerlingen van groep 2. Hoewel wel enkele gegevens over instructiegedrag waren verzameld, zijn in de analyses op leerkrachtniveau alleen de variabelen deeltijd/voltdarheid en volledige inzetbaarheid gerelateerd aan de toetsscores van leerlingen, aangezien de centrale onderzoeksvragen zich daarop richtten. Effecten van deze variabelen konden overigens niet worden geconstateerd. In het onderzoek van Harskamp en Willemsen (1990) zijn rekenvaardigheden getoetst bij 66 leerlingen uit groep 2 (twee leerlingen van gemiddeld prestatieniveau per klas). Er zijn geen significante verbanden gevonden tussen de prestaties op de rekentoets en de mate waarin leerkrachten rekenactiviteiten uit programma's of speelleermateriaal voor rekenen aanbieden.

De huidige stand van zaken in het onderzoek van het onderwijs aan jonge kinderen laat nauwelijks uitspraken toe over effecten op het cognitief functioneren van leerlingen. In onderzoek naar prestaties van kinderen in groep 1 en 2 op het terrein van lezen/taal en rekenen/wiskunde (Verhoeven & Van Kuyk, 1991) zijn tot nu toe zelden variabelen op het niveau van de leerkracht opgenomen, terwijl in onderzoek naar het gedrag van leerkrachten in groep 1 en 2 tot nu toe zelden variabelen op het niveau van de leerling zijn opgenomen. Voorzover dat wel is gebeurd, zijn nog geen duidelijke relaties te leggen tussen bepaalde kenmerken van het onderwijs en het cognitief functioneren van leerlingen (zie verder de bijdragen van Van Kuyk & Verhoeven en van Verhoeven & Narain aan dit themanummer).

6. FACTOREN OP SCHOOLNIVEAU

Zoals in de inleiding al werd beschreven was het reduceren van de verschillen in onderwijsaanpak tussen de groepen 1 en 2 en de hogere leerjaren een van de hoofddoelen van de WBO. Volgens de CEB (1994a) is dat doel nog niet bereikt. Ook volgens 51 % van de leerkrachten is er nog steeds een kloof tussen groep 2 en groep 3 (Logman, 1994). Voorzover de in dit artikel beschreven onderzoeken vergelijkingen tussen de werkwijze in groep 1 en 2 en hogere leerjaren mogelijk maken, blijken op alle fronten nog verschillen te bestaan. Deze verschillen zijn grotendeels terug te voeren op het vroeger gebruikelijke onderscheid van een nadruk op spelen in groep 1 en 2 en een nadruk op leren en presteren in groep 3 en hoger (Ministerie van Onderwijs en Wetenschappen, 1990; Harskamp et al., 1992). Wat dit betreft onderscheiden de 'vroege' onderzoeken (ten tijde van de invoering van de WBO) zich niet van de meer recente.

Ondanks het feit dat leerkrachten uit groep 1 en 2 deelnemen aan de teamvergaderingen, doorgaans wel regelmatig overleggen met de leerkracht van groep 3 en ook afspraken hebben

gemaakt over begrippen die leerlingen moeten kennen en over het uitwisselen van informatie over leerlingen, hebben veel leerlingen toch nog steeds problemen als ze in groep 3 komen (Knuver & Reezigt, 1991). Het gaat dan om problemen met het wennen aan de klassikale manier van werken, het vele schriftelijke werk, het lange stilzitten en het gebrek aan mogelijkheden om zelf activiteiten te kiezen. Het lijkt er op dat de leerkrachten uit de verschillende leerjaren in de onderbouw elkaar wel beter hebben leren kennen, maar dat dit nog niet heeft geleid tot meer overeenstemming in de dagelijkse werkwijze in de klas. Voorzover al veranderingen zijn opgetreden gaat het meer om een wat lossere sfeer in groep 3 dan om een schoolsere sfeer in groep 1 en 2 (Van Gennip & Ruijs, 1992).

Volgens de Inspectie (1994a) is op slechts 22 % van de scholen sprake van een vloeiende overgang naar het leren van de basisvaardigheden lezen, schrijven, taal en rekenen. Op papier, in het schoolwerkplan, wordt op 40 % van de scholen weliswaar aandacht besteed aan een ononderbroken ontwikkeling, maar slechts 16 % van de scholen zet dit principe om in een goed schoolbeleid voor het jonge kind. De problemen zitten ook bij de leerkrachten zelf: minder dan de helft (38 %) heeft voldoende deskundigheid op het gebied van onderwijs aan het jonge kind. De Inspectie (1994a) schetst, uitgaande van de drie standaarden schoolwerkplan, schoolbeleid en deskundigheid een somber beeld: slechts 4 % van de scholen voldoet aan alle drie standaarden, terwijl 34 % op dit moment aan geen van de drie standaarden voldoet.

Overigens wijzen Van der Meer et al. (1986) er op dat er niet alleen breuklijnen zijn tussen groep 1 en 2 enerzijds en de hogere leerjaren anderzijds, maar ook binnen de hogere leerjaren onderling. Door de nadruk op de breuk tussen groep 2 en 3 is dat aspect meestal onderbelicht gebleven. Hun uitspraak "de school is een losse verzameling klassen" beperkt zich dus niet tot de scheiding tussen het vroegere kleuter- en lager onderwijs, maar strekt zich uit van groep 1 tot en met groep 8.

7. SAMENVATTING EN CONCLUSIES

In deze laatste paragraaf worden de resultaten uit de 13 onderzoeken samengevat en afgezet tegen enerzijds de verworvenheden van het kleuteronderwijs en anderzijds kenmerken van effectief onderwijs. Daarna volgen suggesties voor verder onderzoek naar onderwijs aan jonge kinderen.

Verworvenheden van het kleuteronderwijs

In de discussies rond het onderwijs aan jonge kinderen spelen de verworvenheden van het kleuteronderwijs (Van Erp & De Visch Eijbergen, 1983; Van Gennip & Ruijs, 1992; Janssen-Vos, 1993a; CEB, 1994a) al jarenlang een rol. De verworvenheden worden daarbij doorgaans in de strijd geworpen om een dreigende verschoolsing van het onderwijs aan jonge kinderen tegen te gaan. Van Kuyk (1991b) wijst er op dat het kleuteronderwijs door de invoering van de WBO voor het eerst van buiten af beïnvloed werd en toen, vanuit een min of meer bedreigde positie, de eigen waardevolle kenmerken (de verworvenheden) moest gaan verwoorden. De vrees voor verschoolsing van het onderwijs aan jonge kinderen is overigens geen typisch Nederlandse aangelegenheid: ook in de Verenigde Staten is een vergelijkbare discussie gaande over de eisen die al of niet aan jonge kinderen gesteld kunnen worden (Elkind, 1986; Karweit, 1989; Willis, 1993).

In de loop der jaren kregen de verworvenheden een geur van heiligheid, zonder dat duidelijk was in hoeverre de feitelijke praktijk op de voormalige kleuterscholen en in de huidige groepen 1 en 2 de verworvenheden weerspiegelden en zonder dat duidelijk was wat de concrete effecten van de verworvenheden op leerlingen waren. De relevantie van de verworvenheden wordt overigens recent steeds vaker ter discussie gesteld (Janssen-Vos, 1993a). Onderzoek op dit terrein betreft voornamelijk onderzoek naar de mening van deskundigen over de gewenste vormgeving van het onderwijs en niet onderzoek van de klaspraktijk voor of na de invoering van de WBO.

In verschillende publikaties worden lijsten verworvenheden beschreven. Sommige verworvenheden betreffen de materiële condities: er moeten verschillende hoeken in het lokaal zijn, er moet constructie- en compositiemateriaal aanwezig zijn en ouders moeten in het lokaal aanwezig kunnen zijn. Enkele andere verworvenheden zijn globaal geformuleerde uitgangspunten, die niet zo duidelijk richtinggevend zijn voor de praktijk (bijvoorbeeld: het kind moet als een totaliteit worden gezien, het lokaal is een leefruimte). De meeste verworvenheden zijn onder twee noemers te rangschikken: differentiatie en ruimte voor eigen initiatief van kinderen.

Onder *differentiatie* vallen verworvenheden als:

- er is een pluriforme onderwijssituatie, d.w.z. individuele kinderen en groepjes kunnen aan verschillende taken bezig zijn,
- de leerkracht accepteert en honoreert verschillen tussen kinderen,
- de leerkracht gaat uit van ervaringen en interesses van kinderen.

Onder *eigen initiatief* van kinderen vallen verworvenheden als:

- er zijn mogelijkheden voor zelfontdekkend en toevallig leren,
- er is gelegenheid voor gesprekken tussen kinderen,
- kinderen bepalen zelf wat ze gaan doen,
- kinderen kunnen spelen om het spelen zelf.

Op basis van de onderzoeken die in dit artikel beschreven zijn kan worden nagegaan in hoeverre de praktijk in groep 1 en 2 overeenkomt met de verworvenheden. Een pluriforme onderwijssituatie (verschillende kinderen doen verschillende dingen) is in de meeste groepen 1 en 2 een tamelijk gangbaar verschijnsel. Accepteren en honoreren van verschillen komt eveneens voor, in die zin dat leerkrachten niet van alle leerlingen hetzelfde verwachten en onderscheid maken in de eisen die ze aan hen stellen. Leerkrachten letten in dat geval vooral op niveauverschillen tussen kinderen, vaak gekoppeld aan hun leeftijd. Omdat de meeste leerkrachten de kinderen niet vaak toetsen of gericht observeren, gaat het met name om hun inschatting van niveauverschillen, meer dan om objectief geconstateerde verschillen. Leerkrachten willen in principe wel rekening houden met interesses van kinderen, maar ze gooien er hun vooraf vastgestelde planning meestal niet voor om. Globaal gezien is de verworvenheid *differentiatie* dus terug te vinden in de dagelijkse onderwijspraktijk. In de voorgaande tekst is echter al aangetekend dat het daarbij meer gaat om variatie (van activiteiten, van eisen aan kinderen) dan om differentiatie in de betekenis van actief en doelgericht door middel van een bewust gekozen onderwijsaanbod inspelen op objectief geconstateerde verschillen tussen kinderen. Het *eigen initiatief* van kinderen wordt in de meeste groepen 1 en 2 gehonoreerd door kinderen de gelegenheid te geven in de verschillende hoeken van de klas te spelen. Daardoor zijn de vier genoemde verworvenheden in principe gerealiseerd. De vraag is echter of een groot beroep op het eigen initiatief van kinderen altijd wel zo wenselijk is. Gegevens over kinderen uit achterstandssituaties maken duidelijk dat zelfontdekkend leren niet altijd spontaan optreedt en dat sommige kinderen meer leren door een sterkere sturing van de kant van de leerkracht (Leseman & Cordus, 1994), juist ook in perioden dat ze vrij spelen. Voor alle kinderen geldt dat ze de meeste betrokkenheid bij activiteiten tonen waar ook de leerkracht actief bij betrokken is. Toch is het juist in groep 1 en 2 zo dat kinderen veel tijd individueel, zonder begeleiding van de leerkracht, doorbrengen. In vergelijking met groep 3 wordt van hen een relatief grote zelfstandigheid verwacht in het zinvol en actief invullen van de tijd, een zelfstandigheid die mogelijk niet bij alle kinderen voldoende aanwezig is (o.a. Van Kuyk, 1992; Bowman, 1993).

De verworvenheden van het kleuteronderwijs zijn dus weliswaar enigermate terug te vinden in de huidige praktijk van groep 1 en 2, maar dat laat de vraag onverlet in hoeverre de verworvenheden zelf betekenis hebben als het gaat om de meest gewenste vormgeving van het onderwijs aan jonge kinderen en de potentiële effecten daarvan.

Kenmerken van effectief onderwijs

Welk onderwijs in de praktijk effectief is voor jonge kinderen is moeilijk te zeggen, enerzijds omdat niet duidelijk is welke effecten precies worden beoogd en wanneer vanuit dat oogpunt

sprake is van effectief onderwijs (gedefinieerd als het onderwijs dat de gewenste effecten realiseert), anderzijds omdat nauwelijks onderzoek beschikbaar is waarin gegevens op leerling-niveau zijn verzameld, zodat het niet mogelijk is op basis van de empirie uitspraken over effectief onderwijs (gedefinieerd als het onderwijs dat de hoogste scores op maten voor cognitief functioneren bij leerlingen realiseert) te doen.

In dit artikel is gebruik gemaakt van een onderwijseffectiviteitsmodel als beschrijvingskader voor de resultaten uit verschillende onderzoeken. Daardoor is het in ieder geval wel mogelijk de huidige praktijk in groep 1 en 2 af te zetten tegen kenmerken waarvan op theoretische en empirische basis effectiviteit kan worden verwacht in het cognitieve domein. Niet over alle begrippen uit het model zijn gegevens beschikbaar: zo ontbreekt informatie over manieren om de leertijd te verlengen, de verwachtingen van leerkrachten over hun leerlingen, de wijze waarop leerkrachten nieuwe leerstof of begrippen aanbieden en de wijze waarop ze vragen stellen aan de klas. Evenmin is voldoende onderzoek beschikbaar om duidelijke uitspraken over effecten van onderwijskenmerken op het cognitief functioneren van leerlingen te doen. Wel zijn, uitgaande van het beschrijvingskader (zie figuur 1) op basis van de beschreven onderzoeken de volgende conclusies ten aanzien van *instructie* mogelijk:

- de helft van de leerkrachten heeft een redelijk ordelijk klimaat in de klas, de andere helft is chaotisch of juist zeer ordelijk,
- de meeste leerkrachten gaan uit van heel globale doelen en vinden cognitieve ontwikkelingsdoelen niet het belangrijkste,
- leerkrachten werken volgens een globale thematische planning, niet volgens een leerstofinhoudelijke,
- in kontakten met leerlingen treden leerkrachten zelden direct sturend op; ze hebben een voorkeur voor algemeen stimulerende opmerkingen,
- leerkrachten houden wel bij wat er in de klas gebeurt en wie wat heeft gedaan, maar minder duidelijk is welke gegevens ze systematisch op het niveau van de individuele leerling verzamelen en hoe ze evaluatiegegevens gebruiken,
- leerkrachten geven leerlingen zelden feedback over hun activiteiten,
- er zijn wel voorzieningen voor remedieel onderwijs, maar niet voor hoogbegaafde kinderen.

Ten aanzien van *differentiatie*:

- in groep 1 en 2 is wel sprake van variatie in leerstof en taken, maar niet van doelgerichte differentiatie.

Ten aanzien van *materiaal*:

- het "klassieke" kleutermateriaal is praktisch overal aanwezig; een groot deel van de leerkrachten heeft ook een of meerdere programma's in bezit, maar deze worden vooral als bronnenboeken gebruikt; er wordt weinig gedaan om kinderen gerichte voorbereidende lees-, taal-, schrijf- en rekenactiviteiten te bieden.

Ten aanzien van de *gelegenheid om te leren*:

- 40 tot 60 % van de tijd in groep 1 en 2 valt, ruim gedefinieerd, als onderwijstijd te kenschetsen.

Ten aanzien van de *leertijd/betrokkenheid* van leerlingen:

- gemiddeld is 75 % van de leerlingen betrokken bij de activiteiten in de klas; naarmate de leerkracht actiever optreedt stijgt de betrokkenheid van leerlingen.

Deze opsomming maakt duidelijk dat van de door zovelen gevreesde 'verschoolsing' van het onderwijs in groep 1 en 2 geen sprake is. Leerkrachten zijn bijvoorbeeld niet erg geneigd de leerlingen actief voor te bereiden op de situatie in groep 3, werken weinig frontaal-klassikaal, werken niet of nauwelijks volgens methoden, hanteren een globale planning van activiteiten en hebben geen gestructureerde leerstofindeling. Hoewel er nauwelijks gegevens over de periode voor de invoering van de WBO zijn, lijkt het erop dat het onderwijs in groep 1 en 2 sindsdien niet essentieel is veranderd en niet opzienbarend te lijden heeft gehad van druk vanuit hogere leerjaren.

Vanuit een effectiviteitsstandpunt kan echter de vraag worden gesteld of enige 'verschool-

sing' juist niet wenselijk zou zijn. Daarmee wordt niet gepleit voor het rigoureuus invoeren van 'drill and practice' of voor het klakkeloos navolgen van methodes (Bredenkamp & Shepard, 1989), maar voor het doelgericht en systematisch kiezen en inzetten van onderwijsmaatregelen. Het is van groot belang na te gaan op welke gronden de huidige situatie in groep 1 en 2 eigenlijk te verdedigen zou zijn. Zonder een pleidooi te willen houden voor de zo vaak gevreesde 'werkbladenterreur', is het nodig de praktijk en het onderwijsaanbod in groep 1 en 2 kritisch door te lichten. Wat is bijvoorbeeld het voordeel van weinig directe sturing door leerkrachten, als juist blijkt dat kinderen het meest betrokken zijn bij hun werk als ook de leerkracht een actieve rol speelt; waarom zou het in groep 1 en 2 niet nodig zijn de vorderingen van kinderen objectief te bepalen en te registreren; waarom zouden kinderen van 4 en 5 jaar geen behoefte hebben aan feedback over hun prestaties; waarom zou differentiatie beperkt moeten blijven tot variatie in einddoelen ofwel verschillende eisen aan kinderen stellen (met het risico dat kinderen gaan onderpresteren), wat is de waarde en de betekenis van spelend leren als leerkrachten onvoldoende nagaan in hoeverre en wat kinderen eigenlijk leren als ze spelen en waarom zouden kinderen, zeker de kinderen die de begrippen die in groep 1 en 2 worden aangeleerd allang kennen, geen baat hebben bij gerichte lees-, taal-, schrijf- en rekenactiviteiten? Hiermee is overigens niet gezegd dat de situatie ten opzichte van deze kenmerken in hogere leerjaren essentieel anders of beter is. Het eindrapport van de CEB (1994b) verschaft wat dit betreft voldoende duidelijkheid. In de weergave van resultaten van onderzoeken kon in het algemeen slechts per aspect worden gerapporteerd op welke wijze leerkrachten zich gedragen, omdat maar een of enkele aspecten waren onderzocht. Een van de onderzoeken van de Inspectie (1994a) maakt wel uitspraken mogelijk over meerdere kenmerken heen. Dat leidt, uitgaande van de kwaliteitseisen die geformuleerd zijn voor het onderwijs aan jonge kinderen en die merendeels ontleend zijn aan de literatuur over effectief onderwijs, tot de volgende conclusies:

- 58 % van de leerkrachten voldoet aan de eisen voor de onderwijsleerorganisatie (variatie in werkvormen, eigen initiatief van leerlingen),
- 51 % voldoet aan de eisen voor de leeromgeving (voldoende en goed materiaal),
- 44 % voldoet aan de eisen voor het pedagogisch klimaat (veiligheid bieden, oog hebben voor ervaringen van kinderen),
- 34 % voldoet aan de eisen voor registratie, evaluatie en planning,
- 16 % voldoet aan de eisen voor het onderwijsaanbod (gedifferentieerd, doelgericht, aansluiten bij kenmerken van leerlingen; Inspectie, 1994a).

Slechts 4 % van de onderzochte klassen voldoet aan alle geformuleerde standaarden van de Inspectie (waarbij aangetekend moet worden dat de situatie in groep 3 nog minder rooskleurig is). Op basis van de resultaten ten aanzien van de verschillende effectiviteitskenmerken en op basis van de eisen die aan onderwijs aan jonge kinderen gesteld kunnen worden, is de conclusie gerechtvaardigd dat in de meeste gevallen aanzienlijke verbeteringen wenselijk en mogelijk zijn. Daarbij moet wel worden aangetekend dat het, bij de huidige stand van zaken in het onderzoek naar jonge kinderen in Nederland, bijzonder moeilijk is om uitspraken te doen over de te verwachten concrete effecten op het cognitief functioneren. Voorzover positieve effecten worden verwacht, zijn de verwachtingen gebaseerd op algemene theorieën en onderzoeksgegevens op het terrein van onderwijs-effectiviteit, die niet speciaal gericht zijn op jonge kinderen. Omdat het daarbij, zoals ook al eerder in dit artikel is beschreven, gaat om algemene en essentiële kenmerken van het onderwijsleerproces, is er op voorhand geen reden om aan te nemen dat de effecten van deze kenmerken, aangetoond bij leerlingen uit hogere leerjaren, niet zouden gelden voor leerlingen in groep 1 en 2.

Suggesties voor verder onderzoek

De suggesties voor verder onderzoek op het terrein van het onderwijs aan jonge kinderen hebben betrekking op het type onderzoek dat gewenst is en op variabelen die mogelijk in onderzoek opgenomen zouden kunnen worden.

Op dit moment is een longitudinaal onderzoek in uitvoering dat zich met name richt op de

ontwikkeling van allochtone kleuters (Meijnen, 1990; Van der Leij, Meijnen & Leseman, 1994). In dit onderzoek zijn de meeste kenmerken van effectief onderwijs, zoals in dit artikel beschreven, als variabelen opgenomen. Afgezien van dit onderzoek is er nog geen longitudinaal onderzoek beschikbaar dat onderwijskenmerken kan koppelen aan effecten op de cognitieve en zo mogelijk ook de affectieve ontwikkeling van jonge kinderen. Alle longitudinale studies uit de laatste decennia (Meijnen, 1984; Boland & Mommers, 1991; Slavenburg & Peters, 1989; Van der Werf, Weide & Tesser, 1991) hebben betrekking op groep 3 tot en met 8. In het recent gestarte PRIMA-cohort is wel voorzien in een meting in groep 2, maar groep 1 blijft ook hier buiten beschouwing. Idealiter zou onderzoek moeten starten bij de intrede in de basisschool of zo snel mogelijk daarna.

Op het niveau van de leerling zou gekeken moeten worden naar achtergrondkenmerken, voorspellers van schoolsucces en effectmaten. Bij de achtergrondkenmerken gaat het, naast de geijkte kenmerken zoals intelligentie, milieu van herkomst, sekse en etnische herkomst, met name om de mate waarin kinderen in de voorschoolse periode voorbereid zijn op de basisschool. Nagegaan zou moeten worden in hoeverre vaardigheden als het vermogen instructies te begrijpen, het vermogen eigen vragen of wensen kenbaar te maken en vertrouwdheid met regels en afspraken later schoolsucces voorspellen. Wat effectmaten betreft zou nagegaan moeten worden hoe leerlingen zich ontwikkelen op gebieden die in het onderwijs aan de orde zijn geweest (zoals bijvoorbeeld kennis van begrippen, kleuren), gebieden die nog niet expliciet aan de orde zijn geweest (zoals bijvoorbeeld letterkennis, hoeveelheden erbij of eraf doen) en ten aanzien van metacognitieve vaardigheden die het directe onderwijsaanbod overstijgen.

Op het niveau van de leerkracht zouden gegevens verzameld moeten worden op de gebieden instructie, differentiatie en materiaal. Daarbij zou dan vooral gelet moeten worden op de mate waarin leerkrachten er in slagen instructie, differentiatie en materiaal in onderlinge samenhang te hanteren.

Op het niveau van de school zijn tot nu toe nog niet frequent gegevens verzameld; meestal zijn uitspraken over het schoolniveau, wat groep 1 en 2 betreft, gebaseerd op geaggregeerde leerkrachtgegevens. Relevante variabelen op dit niveau hebben betrekking op de mate waarop het schoolbeleid leidt tot een consequente uitvoering van het onderwijs in verschillende klassen (afspraken op het gebied van doelen, inhouden en materialen, het gebruik van de leertijd, het registreren en evalueren van vorderingen van leerlingen).

Wat de methode van onderzoek betreft verdient in eerste instantie gegevensverzameling door observatie de voorkeur (waarbij het belangrijk is dat feitelijke gegevens worden verzameld die niet sterk onderhevig zijn aan interpretaties van de observatoren), zo mogelijk gekoppeld aan andere methoden zoals vragenlijsten, interviews of logboeken.

Het onderwijs aan jonge kinderen is nu al jarenlang het onderwerp van verhitte discussies en vormde een van de peilers voor het innovatieproces basisschool. Gezien het belang van de eerste fase van het basisonderwijs voor de verdere schoolloopbanen van leerlingen is nader onderzoek dat uitspraken toelaat over de effecten van verschillen in onderwijskenmerken op het functioneren van leerlingen niet alleen gewenst maar vooral ook zeer noodzakelijk.

LITERATUUR

- Beishuizen, M. (1983). Handelingspsychologie en vormgeving van instructie. In S. Dijkstra, A.C.M. Duijndijk, & R.J. Takens (Red.), *Psychologie en onderwijs*, 139-157. Lisse: Swets & Zeitlinger.
- Bijtel, J. van der, Blok, H., & Deckers, P. (1992). *Kleuters in het PABO-curriculum*. Den Haag: SVO.
- Boland, Th., & Mommers, M.J.C. (1991). Het relatieve belang van het beheersen van leesvoorwaarden. *Pedagogische Studiën*, 68, (1), 390-403.
- Bosveld, J.J.F., Dijk, M.H. van, Reezigt, G.J., & Weide, M.G. (1988). Differentiatie modellen in het kleuter- en lager onderwijs. *Pedagogische Studiën*, 65, (6), 250-259.
- Bowman, B. (1993). Early childhood education. In L. Darling-Hammond (Ed.). *Review of Research in Education*, 19, 101-135.

- Bredenkamp, S., & Shepard, L. (1989). How best to protect children from inappropriate school expectations, practices and policies. *Young Children*, 44, (3), 14-25.
- Bridgman, A. (1989). The IEA preprimary project: how a dream became reality. *Childhood Education*, 65, (3), 157-159.
- Commissie Evaluatie Basisonderwijs (1994a). *Onderwijs aan jonge kinderen*. Den Haag: SDU.
- Commissie Evaluatie Basisonderwijs (1994b). *Zicht op kwaliteit*. Den Haag: SDU.
- Creemers, B.P.M. (1991). *Effectieve instructie, een empirische bijdrage aan de verbetering van het onderwijs in de klas*. Den Haag: SVO.
- Creemers, B.P.M. (1994). *The effective classroom*. London: Cassell.
- Damhuis, R., & Glopper, K. de (1993). Leerkrachtoordelen over de wenselijke woordenschat van 4- tot 6-jarigen betrouwbaarheid, validiteit en gebruiksmogelijkheden. *Pedagogische Studiën*, 70, (5), 322-336.
- Davis, A. (1990). *The use of a continuous scale to measure growth in writing*. Boston: Paper for the Annual Conference of the American Educational Research Association.
- Dijkstra, S., Dudink, A.C.M., & Takens, R.J. (Red.). (1983). *Psychologie en onderwijs*. Lisse: Swets & Zeitlinger.
- Elkind, J. (1986). Formal education and early childhood education: an essential difference. *Phi Delta Kappan*, 67, (9), 631-636.
- Erp, M van, & Visch Eybergen, A. de (1983). *Op basis van de kleuterschool*. Amsterdam: SCO.
- Gennip, H. van, & Ruijs, A. (1992). *Aan de basis: onderzoek naar de pedagogisch-didactische kwaliteit van het onderwijs aan jonge kinderen*. Nijmegen: ITS.
- Gillijns, P., & Verhoeven, L. (1992). Het CITO-leerlingvolgsysteem: met het oog op de praktijk. *Pedagogische Studiën*, 69, (4), 291-296.
- Glazer, J.S. (1985). Kindergarten and early education: issues and problems. *Childhood Education*, 62, (1), 13-18.
- Goffin, S.C. (1989). Developing a research agenda for early childhood education: what can be learned from the research on teaching? *Early Childhood Research Quarterly*, 4, 187-204.
- Halsted, J.W. (1990). *Guiding the gifted reader*. Washington: Office of Educational Research and Improvement.
- Harskamp, E.G., Pijl, Y.J., & Snippe, J. (1992). *Spel en spelstimulering in het primair onderwijs*. Groningen: RION.
- Harskamp, E.G., Edelenbos, P., & Suhre, C. (1993). *Aanbod voor Nederlandse taal, rekenen/wiskunde en aardrijkskunde in het basisonderwijs*. Groningen: RION.
- Harskamp, E.G., & Willemsen, T. (1991). Programma's en speelleermaterialen voor voorbereidend rekenen in de basisschool. *Pedagogische Studiën*, 68, (1), 404-414.
- Inspectie van het onderwijs (1992). *Goed onderwijs aan jonge kinderen, een evaluatieve beschrijving van aspecten van het onderwijs aan jonge kinderen in de basisschool*. Den Haag: SDU.
- Inspectie van het onderwijs (1994a). *De toestand van het onderwijs aan jonge kinderen: de onderwijsleersituatie, het schoolbeleid en de actualiteit en breedte van de deskundigheid van onderwijsgeven*. De Meern: Inspectie van het onderwijs.
- Inspectie van het onderwijs (1994b). *Technisch rapport van het onderzoek onder 386 basisscholen ten behoeve van de evaluatie van het basisonderwijs*. De Meern: Inspectie van het onderwijs.
- Janssen-Vos, F. (1992). *Basisontwikkeling*. Assen: Van Gorcum.
- Janssen-Vos, F. (1993a). Stand van zaken in het onderwijs aan jonge kinderen: inleiding schakeldagen Jonge Kinderen 1993. *De wereld van het jonge kind*, 21, (4), 102-107.
- Janssen-Vos, F. (1993b). Perspectief voor onderwijs aan kleuters: APS-project Onderbouw Basisschool. In C.G. Derksen & C. Vreugdenhil (Red.). *De kern van beter onderwijs*. Utrecht: APS.
- Karweit, N.L. (1989). Effective kindergarten programs and practices for students at risk. In R.E. Slavin, N.L. Karweit, & N.A. Madden (Eds.). *Effective programs for students at risk*, 103-143. Boston: Allyn and Bacon.
- Knuver, J.W.M. (1993). *De relatie tussen klas- en schoolkenmerken en het affectief functioneren van leerlingen*. Groningen: RION.
- Knuver, J.W.M., & Reeziht, G.J. (1991). *Zittenblijven in het basisonderwijs*. Groningen: RION.
- Koops, W. (1983). Is ontwikkelingspsychologie relevant voor de onderwijspraktijk? In S. Dijkstra, A.C.M. Dudink, & R.J. Takens (Red.). *Psychologie en onderwijs*, 194-203. Lisse: Swets & Zeitlinger.
- Kuyk, J.J. van (1991a). *Does ordering lay the foundation of acquiring mathematics in young children?* Paper for the first European Conference on the Quality of Early Childhood Education, Leuven.
- Kuyk, J.J. van (1991b). Een schets voor het onderwijs aan jonge kinderen (4-8 jaar). *Pedagogische Studiën*, 68, (1), 378-389.

- Kuyk, J.J. van (1992). Hoe kunnen we kinderen helpen in groep 1 en 2? *Jeugd in school en wereld*, 77, (2), 17-21.
- Laevers, F. (1991). *The innovative project Experimental Education (EXE), an introduction*. Paper for the first European Conference on the Quality of Early Childhood Education, Leuven.
- Leseman, P.P.M. & Cordus, J. (1994). *(Allochtone) kleuters meer aandacht*. Rijswijk: WVC.
- Leune, J. (1992). Attitude- en persoonlijkheidsvorming op school. *Meso*, 12, (63), 2-6.
- Leij, A. van der, Meijnen, W., & Leseman, P. (1994). School career from 4 to 7: outline of a longitudinal study. In F. Laevers (Ed.). *Defining and assessing quality in early childhood education*. Leuven: Leuven University Press.
- Lieshout, C.F.M. van (Red.) (1992). *Basisschoolleerlingen met sociaal-emotionele problemen: diagnostiek en begeleiding*. Nijmegen: KUN, Vakgroep ontwikkelingspsychologie.
- Logman, P. (1994). Onderwijskwaliteitsshow. *NGL-blad*, 26, (29), 15-16.
- Lunzer, E.A. (1986). Child development and education: theory and practice. In E.A. Lunzer (Ed.). *Child development at primary school*, 98-113. Lisse: Swets & Zeitlinger.
- McLeod, D.B. (1988). Affective issues in mathematical problem solving: some theoretical considerations. *Journal for Research in Mathematics Education*, 19, (2), 134-141.
- Meer, T.A.M. van der, Appelhof, P.N., Doesschate, J.J. ten, & Kramer, G.B. (1986). *Het onderwijs in de klas tussen 9 en 12*. Utrecht: SAC.
- Meijnen, G.W. (1984). *Van zes tot twaalf*. Harlingen: SVO.
- Meijnen, G.W. (1990). *De schoolcarrière van leerlingen gezien in relatie tot de invloeden die maatschappelijke en binnenschoolse determinanten hierop hebben*. Amsterdam: SCO.
- Meijnen, G.W. & Riemersma, F.S.J. (1992). *Schoolcarrières: een klassenkwestie?* Amsterdam: SCO.
- Ministerie van Onderwijs en Wetenschappen (1990). *Zo hard gelopen en toch nog zo ver te gaan*. Den Haag: SDU.
- Monks, F.J., & Peters, W.A.M. (Eds.) (1992). *Talent for the future: social and personality development of gifted children*. Assen: Van Gorcum.
- Parreren, C.F. van (1987). *Ontwikkend onderwijs*. Leuven/Amersfoort: ACCO.
- Parreren, C.F. van (1990). Hoe stimuleren en hoe observeren we "ontwikkeling"? *De wereld van het jonge kind*, 17, (3), 198-202.
- Pramling, I. (1986). A pedagogical bridge from preschool to primary. In E.A. Lunzer (Ed.). *Child development at primary school*, 28-33. Lisse: Swets & Zeitlinger.
- Reezigt, G.J., Dijk, M.H. van, & Bosveld, J.J.F. (1986). *Differentiatie op de basisschool*. Den Haag: SVO.
- Reezigt, G.J., Guldmond, H., & Ros, A.A. (1993). *Effectief onderwijs aan jonge kinderen*. Groningen: RION.
- Schonewille, B., & Leij, A. van der (1993). Ethnic background, classroom instruction and child behaviour in kindergarten: the role of the teacher to stimulate children in the classroom by giving individual attention. *European Early Childhood Education Research Journal*, 1, (2), 11-25.
- Simonis, L. (1993). Inspectierapport over het jonge kind: bruikbaar, maar niet volledig. *De wereld van het jonge kind*, 20, (8), 228-233.
- Slavenburg, J.H., & Peters, T.A. (Red.) (1989). *Het project Onderwijs en Sociaal Milieu: een eindbalans*. Rotterdam: SAD.
- Sontag, L., & Meijnen, G.W. (1993). Kindergarten education: a study of effective education for young children at risk. *European Early Childhood Education Research Journal*, 1, (2), 51-65.
- Stallings, J.A., & Stipek, D. (1986). Research on early childhood and elementary school teaching programs. In M.C. Wittrock (Ed.). *Handbook of research on teaching*, 727-754. New York: McMillan.
- Ven, J. van de, Vrieze, G., & Claessen, J. (1988). *Grensverkenners in de basisschool*. Nijmegen: ITS.
- Verhoeven, L. (1992). *Ontluikende geleterdheid: schets van een onderzoeksterrein*. Tilburg: Katholieke Universiteit Brabant.
- Verhoeven, L., & Kuyk, J.J. van (1991). Peiling van conceptuele en metalinguïstische kennis bij de aanvang van het basisonderwijs. *Pedagogische Studiën*, 68, (1), 415-425.
- Vermeulen, C.J., & Bosker, R.J. (1992). *De omvang en gevolgen van deeltijdarbeid en volledige inzetbaarheid van leerkrachten in het basisonderwijs*. Enschede: Universiteit Twente/OCTO.
- Wagenaar, E. (1992). *Tweetaligheid in het aanvangsonderwijs: een onderzoek naar de effecten van tweetalig kleuteronderwijs op de schoolloopbaan van Marokkaanse kinderen*. Amsterdam: Het Spinhuis.
- Weinert, F.E. (1987). Developmental processes and instruction. In E. de Corte, H. Lodewijks, R. Parmentier, & P. Span (Eds.). *Learning & instruction, European research in an international context, Vol. 1*, 4-8. Oxford: Pergamon.

- Weinert, F.E., Schrader, F.W., & Helmke, A. (1989). Quality of instruction and achievement outcomes. In F.E. Weinert (Ed.). The relation between education and development. *International Journal of Educational Research*, 13, 895-912.
- Werf, M.P.C. van der, Weide, M.G., & Tesser, P. (1991). *Het onderwijsvoorrangsbeleid in de school en in de klas*. Groningen/Nijmegen: RION/ITS.
- Willis, S. (1993). Teaching young children: educators seek "developmental appropriateness". *Curriculum Update*, 4, 1-8.
- Winnubst, T. (1992). Het begeleiden van kinderen met ernstige gedragsproblemen: invoeringsprogramma "sociaal-emotionele ontwikkeling". *Speciaal Onderwijs*, 65, (5), 161-164.
- Wolbert, R.G.M., Schaap-Hummel, J., & Span, P. (1986). *Individualisering en differentiatie in de basisschool*. Den Haag: SVO.
- Zaal, J.N. (1980). Relaties tussen sociaal emotioneel functioneren en schoolprestaties van leerlingen in kleuter- en basisonderwijs. *Pedagogische Studiën*, 57, (4), 257-267.

Manuscript ontvangen 6-9-1994

Definitieve versie ontvangen 5-4-1995

Stimulering van geletterdheid bij kleuters

Jef van Kuyk¹ en Ludo Verhoeven²

¹CITO, Arnhem en ²Letterenfaculteit, KUB

ABSTRACT

The present article deals with the teaching of early literacy skills to children at kindergarten level. The effects of a developmental and curriculum-based intervention were tested following a pretest-posttest design. The tests measured vocabulary skills, narrative skills, objectivation skills and phonological awareness skills. The informants were divided into three groups: high SES, low SES Dutch children and ethnic minority children.

The results show that the developmental intervention had a positive effect on narrative skills and objectivation skills. The curriculum-based approach yielded mainly a positive effect on phonological awareness skills. For vocabulary and narrative skills a clear group effect could also be evidenced.

INLEIDING

In de traditie van het basisonderwijs wordt met lezen en schrijven begonnen na de kleuterperiode, ofwel bij de aanvang van groep 3. Het onderwijs in schriftelijke vaardigheden is momenteel sterk in beweging. In algemene zin wordt gezocht naar mogelijkheden om de continuïteit in mondelinge en schriftelijke taalverwerving van kinderen beter te kunnen garanderen. Recent onderzoek laat zien dat reeds bij de aanvang van het basisonderwijs sprake is van een grote individuele variatie in essentiële vaardigheden binnen het domein van vroege geletterdheid (Verhoeven & Van Kuyk, 1991). Het gevaar dreigt dat kinderen met een gebrekkige voor-schoolse taalverwerving in het basisonderwijs met een ontwikkelingsachterstand aan de basisvorming in functionele geletterdheid beginnen. Een belangrijke vraag is hoe de (voor)schoolse taalverwerving van kinderen met ontwikkelingsachterstanden zodanig kan worden gecompenseerd dat latere lees- en schrijfproblemen zoveel mogelijk worden ingeperkt. Daarvoor is inzicht nodig in de aard van (voor)schoolse taalverwervingsprocessen en in de factoren die de variatie in tempo van die processen bepalen.

In dit artikel wordt verslag gedaan van een onderzoek naar de effecten van twee gerichte interventies in het kleuteronderwijs, bedoeld om de ontwikkeling van conceptuele en metalingvistische kennis van jonge kinderen te stimuleren en hun gerichtheid op de geschreven taal. In de ene interventie wordt een ontwikkelingsgerichte benadering gevolgd, waarbij de nadruk ligt op functionele activiteiten die ook binnen gezinnen vaak worden gebezigd, zoals voorlezen en taalspel. In de andere interventie, een zgn. gecombineerde benadering, wordt naast een ontwikkelingsgerichte- een programmatische benadering gevolgd waarbij het hoofdaccent ligt op het meer expliciet onderwijzen van kennis en vaardigheden. Alvorens de resultaten van dit onderzoek te presenteren zullen we kort ingaan op de achtergronden van een ontwikkelings- en programmagerichte benadering van het taalonderwijs aan kleuters.

STIMULERING VAN VROEGE GELETTERDHEID

Twee opvattingen

Taalontwikkeling en leren lezen zijn de hoofddoelstellingen van het basisonderwijs. In Nederland zijn kleuter- en lager onderwijs geïntegreerd en de vraag is welke rol het onderwijs aan kleuters daarbij het beste kan worden toebedeeld. De plaats van de taalontwikkeling is daarin onomstreden, maar de plaats van het leren lezen of de voorbereiding daarop is nog allerm minst vastomlijnd. Deze kwestie is niet nieuw. Meer dan twintig jaar geleden werd aan het eind van het kleuteronderwijs de vraag gesteld of kinderen rijp waren om met het leren lezen te beginnen (Sixma, 1972). Uit veel onderzoek is gebleken (Stevens, 1987, Ellwein, Walsh, Eads & Miller, 1991) dat het gebruik van school- en leesrijpheidstests en de beslissing om het kind nog een jaar langer in het kleuteronderwijs te houden onvoldoende specifiek is en weinig effectief is gebleken. Ook in de geïntegreerde basisschool blijken kinderen onvoldoende hulp te krijgen om de leesvaardigheid onder de knie te krijgen, zoals blijkt uit de sterke stijging van het speciaal onderwijs (Stevens, 1987). Kennelijk is het basisonderwijs onvoldoende in staat om kinderen met achterstanden of problemen binnen het basisonderwijs hulp te bieden. Er is in de afgelopen jaren gezocht naar mogelijkheden om problemen vroegtijdig in het kleuteronderwijs op te sporen en binnen het kleuteronderwijs gerichte hulp te bieden. We kunnen in dit verband het traditionele voorwaardendenken onderscheiden en de nieuwere opvatting van de ontluikende geletterdheid.

1. Voorwaarden denken

In het traditionele denken wordt uitgegaan van een terugredenering, dat kinderen moeten voldoen aan een aantal voorwaarden die uit het formele leesproces kunnen worden afgeleid. Dit denken heeft veel weerstand opgeroepen. Het sloot onvoldoende aan op het curriculum en de werkwijze van het kleuteronderwijs, een verschijnsel dat ook veel compensatie- en stimuleringsprogramma's hebben doen mislukken (Groenendaal, 1978, Van Kuyk, 1985; Verhoeven & Mommers, 1989). Vervroeging van het formele leren zou leiden tot overbelasting en overhaasting (Elkind, 1986). Tenslotte werd er onvoldoende rekening gehouden met het ontwikkelingsniveau van jonge kinderen en hun specifieke karakteristieken. Uit onderzoek van Boland en Mommers (1991) is gebleken dat leesvoorwaarden voor een korte duur van belang zijn en een beperkte betekenis hebben voor het begrijpend lezen. Dit model beschouwen wij als minder geschikt voor het onderwijs aan jonge kinderen.

2. Ontluikende geletterdheid

De laatste jaren dient zich een nieuwe aanpak aan die betere garanties lijkt te bieden voor het onderwijs aan jonge kinderen. Het is de aanpak van ontluikende geletterdheid (vgl. Verhoeven, 1992, 1994). Volgens dit idee wordt de ontwikkeling tot geletterdheid beschouwd als een spontaan verlopend ontwikkelingsproces zoals bijvoorbeeld de taalontwikkeling, als er zich in de directe omgeving van het kind voldoende geschreven taalprikkels zijn. Volgens Sulzby en Teale (1991) zijn vooral het vertellen, het spontaan schrijven door kinderen en de ontwikkeling van het metalinguïstisch bewustzijn de belangrijkste factoren die de ontluikende geletterdheid bepalen. Ontluikende geletterdheid is beter toepasbaar in het onderwijs aan jonge kinderen omdat het aansluit op spontane ontwikkelingsprocessen. Deze aanpak is niet zozeer specifiek gericht op bepaalde leesvoorwaarden, maar sluit veeleer aan bij de algehele taalontwikkeling. Als zodanig sluit de aanpak beter aan op het curriculum van het kleuteronderwijs. Het wil echter niet zeggen dat aansluiting bij spontane ontwikkelingsprocessen alleen leidt tot succes. Volgens Sulzby en Teale (1991) is enerzijds ouderlijke attentie van groot belang, maar ook instructie in de schoolsituatie.

Belangrijk is het gegeven dat kinderen al een basiservaring hebben met geletterdheid. Uitgaande van dit model zijn er meerdere benaderingen mogelijk, een ontwikkelingsgerichte benadering waarbij spontane ontwikkelingsprocessen leidraad zijn voor het onderwijs en een programmagerichte benadering waarbij de spontane ontwikkelingsprocessen weliswaar uitgangspunt zijn maar waarbij deze gestuurd worden in de richting van geletterdheid.

Ontwikkelings- en programmagerichte stimulering

Met betrekking tot de schoolsituatie kunnen we onderscheid maken tussen ontwikkelings- en programmagericht onderwijs (vgl. Van Parreren, 1985). Van Parreren bepleitte een ontwikkelingsgerichte benadering voor het kleuteronderwijs en een meer programmagerichte benadering voor het lager onderwijs. In de kleuterschool gaat men in de regel niet van een vastomlijnd programma uit. Daarentegen probeert de leerkracht, uitgaande van incidentele leersituaties, de ontwikkeling van kinderen te stimuleren door een rijke omgevingscontext te scheppen. Met het oog op de begeleiding van kinderen uit risicogroepen wordt er echter vaak voor bepleit om met behulp van programma's ontwikkelingsachterstanden weg te werken. Uit onderzoek naar programma's voor kinderen met achterstanden blijkt dat deze gestructureerd en doelgericht moeten zijn om tot succes te leiden; er moet een evaluatiesysteem beschikbaar zijn om problemen tijdig te kunnen opsporen en het programma moet aangepast zijn aan het ontwikkelingsniveau van de individuele leerling (Slavin, Madden, Karweit, Livermon en Dolan, 1990, Van Kuyk, 1991b). Voor deze kinderen is naast een ontwikkelingsgerichte een meer programmagerichte benadering wenselijk. In feite betekent dit een geïntegreerde aanpak van ontwikkelings- en programmagerichte elementen. (Van Kuyk, 1992, 1994). Alvorens deze geïntegreerde benadering te schetsen geven we in het kader van de ontluikende geletterdheid de kenmerken van de ontwikkelings- en programmagerichte benadering.

Ontwikkelingsgerichte benadering

Onderzoek vanaf de jaren vijftig tot nu heeft aangetoond dat voorlezen in de thuis- en schoolsituatie bijdraagt aan de ontwikkeling van geletterdheid. Analyse van voorleessessies heeft aangetoond dat voorlezen fundamenteel een daad van constructie is: volwassene en kind hebben er een actieve rol in. Door voorspelbare actieschema's leert het kind hoe te participeren in deze activiteit. Die actieschema's blijken in de loop van de tijd te veranderen. Aanvankelijk is het voorlezen zeer interactief. Na het derde levensjaar lezen de ouders langer achter elkaar en tenslotte wordt een verhaal bijna in zijn geheel achter elkaar gelezen. Dan worden pas vragen gesteld naar aanleiding van het verhaal waarbij het kind de tekstrepresentatie leert vast te houden (Teale & Sulzby, 1986). Door deze interactie leren kinderen niet alleen hun woordenschat ontwikkelen, maar kunnen zij ook het boek 'herlezen'. Aanvankelijk imiteert het kind de leeshandeling maar naarmate het ouder wordt neemt het begrip van de gesproken en geschreven taal toe. Een uitlokkende stijl van voorlezen en het reageren op de intenties van het kind blijken het meest effectief te zijn. Ouders zijn dus belangrijke mediërende factoren (Sulzby en Teale, 1991).

In een aantal opzichten is de verwervingscontext in een schoolse context vergelijkbaar met die in een gezin. In beide gevallen is noodzakelijkerwijs sprake van een aanbod aan geschreven taal en van een interactie tussen een oudere en het kind. Er zijn echter ook duidelijke verschillen. In het stimulerende gezin is sprake van interpersoonlijke interacties tussen ouder en kind waarbij kinderen op basis van een veilige relatie de mogelijkheid krijgen om op velerlei wijze in aanraking te komen met geletterdheid in functionele situaties. In een schoolcontext is doorgaans sprake van groepen van ongeveer 25-30 kinderen met uiteenlopende persoonlijke en achtergrondkenmerken. Een aanpak die aansluit op de individuele kennis en beleving is in feite uitgesloten.

Ook in het onderwijs kunnen echter interactiepatronen worden geïmplementeerd die de ontwikkeling van vroege geletterdheid positief beïnvloeden. Mason, McCormick en Bhavnagri (1986) pleiten in dit opzicht voor taalspel, voorlezen en vertellen in kleine groepssessies. Leerkrachten zouden zich er in hun instructie vooral op moeten richten dat kinderen zelfstandig vragen leren stellen naar aanleiding van het gelezene.

Ook het pre-conventionele schrijven kan bijdragen aan de ontwikkeling van geletterdheid. Schrijven moet dan niet gezien worden als uitsluitend gericht op letterformatie en de mechanische aspecten ervan, maar op de relatie tussen gesproken en geschreven taal. Kinderen maken eigen composities. Ze 'schrijven' woorden en verhaaltjes. We moeten dan niet letten op de vorm, maar op de onderliggende conceptualisaties van het schrijven. Het schrijven ontwikkelt

zich in fasen. Aanvankelijk heeft het geschrevene geen betekenis voor anderen. Geleidelijk aan ontdekken kinderen dat elke letter correspondeert met een klank en dat er correspondentie bestaat tussen gesproken en geschreven taal. Belangrijk bij het ontluikend schrijven zijn het voorbeeld geven en de sociale context (Sulzby en Teale, 1991).

Een derde aspect van de vroege ontwikkeling van geletterdheid is het metalinguïstisch bewustzijn. Het is de vaardigheid te reflecteren op de taal als object. Het metalinguïstisch bewustzijn omvat een aantal bekwaamheden die variëren van het kunnen waarden van een taalgrapje tot het separeren van woorden in fonemen. Er zijn vier niveaus te onderscheiden: fonologisch bewustzijn, woordbewustzijn, vormbewustzijn en pragmatisch bewustzijn. De eerste twee vaardigheden stellen het kind in staat te reflecteren op deeltjes van de taal zoals het woord en de klank. De twee laatste maken reflectie op de betekenis en appreciatie van grotere delen van de taal mogelijk. Via taalspel in de kring kan de leerkracht proberen op semantisch coherente wijze op kinderen te reageren zodat intuïties over taal kunnen worden opgebouwd. Voorbeelden van dergelijke interactievormen geven Mason, et al. (1986). Daarnaast kunnen kleuters in een lees-schrijfhoek in de kleuterklas leren om vanuit ervaringen met schrift hun metalinguïstische kennis uit te breiden. Volgens Sulzby en Teale (1991) is de precieze relatie tussen de ontwikkeling van het metalinguïstisch bewustzijn en ontluikende geletterdheid echter nog verre van duidelijk.

Programmagerichte benadering

Binnen programmagerichte stimulering wordt uitgegaan van een doelgerichte aanpak waarbij directe instructie van vaardigheden centraal staat.

Het systematisch voorlezen aan kinderen in een onderwijssituatie blijkt ook de woordenschatontwikkeling van kinderen te bevorderen. Uit een studie van Elley (1989) blijkt dat kinderen in de onderbouw van scholen in Nieuw Zeeland significant vooruitgaan wat woordenschat betreft als gevolg van voorleessessies zonder nadere uitleg. Bij het geven van nadere uitleg omtrent de achtergrond en het verloop van het verhaal blijkt de leerwinst nog groter. De leerwinst van de kinderen bleek in hoge mate consistent. Nader onderzoek wees uit dat het leren van afzonderlijke woorden afhankelijk was van het aantal keren dat een woord in de tekst voorkwam, het voorkomen van het desbetreffende concept in illustraties en de mate van redundantie van betekenisaspecten in de context. Onderzoek van Morrow en Smith (1990) laat verder zien dat het tekstbegrip van kinderen bij voorlezen op school mede afhankelijk is van de grootte van de groep waarin wordt voorgelezen. Kinderen in kleine groepjes of in individuele vertelsituaties bleken meer ideeën te genereren en beter te presteren qua tekstbegrip dan kinderen in klassikale leessituaties.

Voor de ontwikkeling van het fonologisch bewustzijn blijkt een doelgerichte instructie eveneens effectief te kunnen zijn. Uit onderzoek van Bradley en Bryant (1983) bleek dat kinderen die geïnstrueerd werden in fonologisch bewustzijn (herkenning van eerste, middelste en laatste klank in een woord) en daarnaast ook in klank-letterkoppeling significant hoger scoorden bij het lezen dan de controlegroep. Uit onderzoek van Cunningham (1988) bleek dat kinderen die leerden fonemen te segmenteren door directe instructie zonder klank-letterkoppeling succesvol waren met lezen. Kinderen die daarnaast nog bewust werden gemaakt waarvoor de oefeningen bedoeld waren, een vorm van metacognitie, hadden nog betere leesresultaten. Ball en Blachman (1991) toonden eveneens aan dat foneembewustzijn significant gerelateerd is aan succes in de beginfase van lezen en spellen. Kinderen die leerden fonemen te segmenteren waren succesvoller dan een controlegroep. Dat gold niet voor letterkennis. Wel bleken kinderen die naast het segmenteren instructie kregen in klank-letterkoppeling betere resultaten te hebben.

Een belangrijke studie naar de ontwikkeling van fonologisch bewustzijn in een onderwijskundige context is die van Olofsson en Lundberg (1985). Uitgaande van een longitudinaal perspectief gingen zij bij 235 Deense kleuters het effect na van een acht maanden durend trainingsprogramma bestaande uit metalinguïstische spelletjes en oefeningen. Het programma bleek geen effect te hebben op de ontwikkeling van de woordenschat of de luistervaardigheid

van kinderen. Wel kon een significante vooruitgang worden vastgesteld op een aantal metalinguïstische taken. Voor rijm en manipulatie van syllaben waren de effecten relatief klein. Voor manipulatie van fonemen waren de effecten extreem groot. Het onderzoek laat zien dat metalinguïstische kennis ontwikkeld kan worden buiten de context van lees- en schrijfinstructie en een positief effect heeft op het latere lezen (Lundberg, Frost en Petersen, 1988).

Tot een vergelijkbare conclusie komen Williams (1986) en Cunningham (1988). Cunningham pleit er echter voor dat kinderen een dergelijke training krijgen in de context van een leessituatie.

Uit verschillende onderzoeken blijkt dat programmagerichte stimulering in het kleuteronderwijs effectief kan zijn, zowel met betrekking tot betekenisaspecten als wat betreft metalinguïstisch bewustzijn.

Geïntegreerde aanpak

Uit de voorgaande beschrijving moet niet worden geconcludeerd dat beide benaderingen elkaar uitsluiten, integendeel. In de ARBO-nota Spelen en Jongleren (ARBO, 1990) wordt een synthese van deze benaderingen bepleit in het onderwijs aan vier tot achtjarige kinderen en Van Kuyk (1991a, 1992a) heeft herhaaldelijk een integratie van beide bepleit. Vanuit een complementair onderwijsmodel wordt een ontwikkelingsgerichte benadering voorgesteld daar waar mogelijk, namelijk daar waar de zelfsturing van het kind dit toelaat en een programmagerichte benadering waar dit nodig is, namelijk daar waar het aan zelfsturing ontbreekt en deze moet worden versterkt. Zoals eerder aangegeven hebben met name kinderen met achterstanden behoefte aan structuur en een doelgerichte aanpak. Deze benadering werkt het meest effectief als er sprake is van tutoring. (Dit is een vorm van individuele hulpverlening). (Slavin, Karweit & Wasik, 1994). Uitgangspunt is steeds een ontwikkelingsgerichte benadering waarin programmagerichte elementen toegevoegd worden waar dit voor de ontwikkeling van het jonge kind vereist is. We denken vooral aan kinderen die op taaltoetsen lage scores behalen. Dat is veelal het geval bij Nederlandse kinderen uit lagere sociale milieus en bij allochtonen kinderen (Van Kuyk, 1991a; Verhoeven, 1990; Verhoeven en Van Kuyk, in voorbereiding). Deze geïntegreerde aanpak vereist van de leerkracht een breed en flexibel gedragsrepertoire dat op een responsieve wijze moet worden ingezet: afgaande op de signalen van kinderen de juiste respons geven (Lamb en Easterbrooks, 1981). Voor de hoogscorende kinderen (niet-risicokinderen) zal in het algemeen een ontwikkelingsgerichte benadering effectief zijn, voor laag scorende kinderen (risico-kinderen) zal een ontwikkelingsgerichte benadering niet voldoende zijn. Er zullen programmagerichte elementen moeten worden toegevoegd. Daarnaast dienen programmagerichte elementen zoveel mogelijk gericht te zijn op individuele kinderen (tutoring) die structuur en instructie nodig hebben.

Vraagstelling

Het onderzoek gaat uit van de volgende vragen:

1. In hoeverre werkt een ontwikkelingsgerichte benadering van beginnende geletterdheid in het kleuteronderwijs beter dan een traditionele benadering?
2. In welke mate heeft een programmagerichte stimulering in aanvulling op de ontwikkelingsgerichte benadering een positief effect op de ontwikkeling van risicokinderen?

Het onderzoek richt zich op de ontwikkeling van beginnende geletterdheid bij jonge kinderen. Daarbij gaat het om conceptuele vaardigheden, metalinguïstisch bewustzijn en de oriëntatie op boek en schrift. Binnen conceptuele vaardigheden gaat het primair om woordenschat, definitievaardigheden en vertelvaardigheid. Ten aanzien van metalinguïstisch bewustzijn valt verder onderscheid te maken tussen woordbewustzijn en fonologisch bewustzijn. Bij woordbewustzijn gaat het om het inzicht in formele aspecten van woorden, zoals rijm en woordobjectivatie. Bij fonologisch bewustzijn staat het inzicht centraal dat woorden zijn opgebouwd uit klanken. Vaardigheden, zoals de segmentatie en synthese van lettergroepen en fonemen treden hierbij op de voorgrond. Wat betreft de oriëntatie op boek en schrift valt onderscheid te maken tussen

kennis omtrent de conventies van geschreven taal en kennis van de schriftcode. Belangrijke schriftconventies die kinderen moeten zien te verwerven vormen de leesrichting, en de opmaak van pagina's en boeken. Bij de schriftcode zelf gaat het om het inzicht dat grafemen in geschreven taal de representanten vormen van fonemen.

Met het oog op de eerste onderzoeksvraag gaan we na wat het effect is van het werken met een ontwikkelingsgerichte aanpak op de ontwikkeling van beginnende geletterdheid bij kleuters. In deze benadering staat het interactief voorlezen van verhalen binnen een aantal thema's die nauw aansluiten bij de leefwereld van kleuters centraal. In het onderzoek wordt uitgegaan van verschillende strata: kinderen uit gezinnen met een hoge of gemiddelde sociaal-economische status, kinderen uit gezinnen met een lage sociaal-economische status en allochtone kinderen. Op deze wijze is het mogelijk om vast te stellen of eenmaal gevonden effecten in vergelijkbare mate optreden voor kinderen uit verschillende milieus.

De tweede onderzoeksvraag wil nagaan wat binnen de ontwikkelingsgerichte aanpak het effect is van extra programmagerichte stimulering van risicokinderen. De programmagerichte stimulering heeft het kenmerk van een directe instructie van kennis en vaardigheden aan kinderen in kleine groepjes. Als risicokinderen worden aangemerkt: allochtone kinderen en autochtone kinderen uit lagere sociaal-economische milieus. Door dit onderscheid is het mogelijk om te bepalen in hoeverre eenmaal gevonden effecten van de extra programmagerichte stimulering voor allochtone en autochtone kinderen vergelijkbaar zijn.

OPZET VAN HET ONDERZOEK

Proefpersonen

In de longitudinale studie waren 189 kleuters betrokken, die van 4-6 jaar werden gevolgd. Bij de aanvang van het onderzoek, november 1990 hadden de kinderen een leeftijdsrange van 4.1 jaar - 5.1 jaar.

Ze waren afkomstig uit 37 klassen van 15 scholen uit Amsterdam, Tilburg, Arnhem en Nijmegen. De groep werd zodanig samengesteld dat er kinderen uit 3 sociaal-economische groepen in vertegenwoordigd waren: kinderen, waarvan ten minste een van de ouders een hoge beroepsopleiding heeft: een hoge en midden sociaal-economische status (SES), (stratum 1), kinderen waarvan beide ouders een lage beroepsopleiding hebben, een lage SES, (stratum 2), en kinderen die uit een ander land afkomstig zijn en waarvan beide ouders een lage beroepsopleiding hebben: allochtonen met een lage SES, (stratum 3). Er is daarbij gezocht naar een vergelijkbare representatie wat betreft het land van herkomst met betrekking tot de verschillende condities, met name wat betreft mediterrane en ex-koloniale groepen. Gedurende het onderzoek was er sprake van uitval. In de eerste plaats moest het onderzoek om financiële redenen worden ingekrompen. Het belangrijkste gevolg was dat stratum 1 kinderen niet aan de gecombineerde conditie deelnamen. Verder werd de uitval voor een deel veroorzaakt door verhuizing of verzuim in het zicht van de grote vakantie. Dit laatste veroorzaakt een grotere uitval bij allochtonen, die voor de laatste meting naar het moederland vertrokken.

De kinderen van de drie strata werden verdeeld over de verschillende condities. Aan de controleconditie namen 66 kinderen deel, aan de ontwikkelingsgerichte conditie 82 kinderen en aan de gecombineerde conditie 41 kinderen. Deze laatste conditie bevatte alleen stratum 2- en stratum 3 kinderen.

Instrumenten

De instrumenten die in het onderzoek gebruikt zijn, zijn voor een belangrijk deel ontleend aan toetsen die bij eerder onderzoek zijn ingezet. De betrouwbaarheid van deze toetsen is redelijk tot goed. Voor een uitgebreide beschrijving van de toetsen, zie Verhoeven en Van Kuyk (1991). De toetsen zijn verdeeld in clusters die eerder in dit artikel aan de orde zijn gekomen: conceptuele vaardigheden, metalinguïstisch bewustzijn, verdeeld in woordbewustzijn en foneembe-

wustzijn, en oriëntatie op boek en schrift. De conceptuele ontwikkeling is geoperationaliseerd in de volgende toetsen: passieve en actieve woordenschat, definitievaardigheid en vertelcapaciteit.

Het woordbewustzijn heeft betrekking op verschillende aspecten waarin betekenis- en vormelementen nog nauw verbonden zijn: rijm, objectivatie en woordanalyse in zinnen. Het foneembewustzijn verwijst naar het onderscheiden van spraakklanken, waarbij de relatie met betekenisaspecten afwezig is. Tenslotte de oriëntatie op boek en schrift verwijst naar de gerichtheid van het kind op de geschreven taal: boeken, leesbegrippen en letterkennis.

In het onderstaande geven we de toetsen aan.

Conceptuele vaardigheden

- 1 Woordenschat, passief: 98 woorden.
Aan de hand van 4 plaatjes wordt een woord genoemd en het juiste plaatje aangewezen.
- 2 Woordenschat, actief: 64 woorden.
Aan de hand van een plaatje moet het kind het juiste woord noemen.
Deze beide toetsen zijn ontleend aan de TAK (Verhoeven, Vermeer, 1986).
- 3 Definitietaak.
De toets bestaat uit 16 woorden. Per woord worden 3 vragen gesteld.
b.v. 1 Wat is een gieter (kun je planten/bloemen mee water geven)
2 Wat doe je met een gieter (sproeien, gieten)
3 Hoe ziet een gieter eruit (handvat, tuit, ding met sproeier).
Het kind ontvangt twee punten als het een goede definitie geeft op de algemene vraag. Eén punt voor ieder goed antwoord op de twee specifieke vragen.
- 4 Verteltaak.
Het kind krijgt twee keer een reeks plaatjes voorgelegd en moet aan de hand daarvan een verhaaltje vertellen. Het kind krijgt punten als het losse elementen van de plaatjes kan noemen en als het tussen de situaties verbanden kan leggen.

Woordbewustzijn

Bij woordbewustzijn zijn drie toetsen opgenomen waarin zowel op het betekenselement als op het vormelement gelet kan worden en waarin de vertrouwdheid met het betekenselement aanwezig is.

- 1 Rijmen.
De toets bestaat uit 10 rijmwoorden. Aan de hand van plaatjes moet het kind aangeven welk woord rijmt op een gegeven woord.
- 2 Woordobjectivatie.
De toets bestaat uit 10 sets korte en lange woorden. Het kind moet de korte of lange woorden noemen.
Hier moet het kind kunnen afzien van de betekenis om zich geheel te richten op het lengteaspect. Betekenselementen kunnen hier misleidend zijn (boom-lucifer).
- 3 Woordsegmentatie.
De toets bestaat uit 10 zinnen met 5 of meer woorden. Het kind moet uit voorgezegde zinnen alle afzonderlijke woorden isoleren.

Foneembewustzijn

Bij foneembewustzijn gaat het om drie toetsen, waarin het kind zich geheel op het vormaspect moet richten en de relatie tussen betekenis en vorm gering of afwezig is.

- 1 Lettergreepsynthese.
De toets bestaat uit 20 woorden. Het kind moet steeds 2 of meer lettergrepen tot een woord voegen, b.v. bo-ter, ra-di-o.
- 2 Foneemsynthese.
Deze toets bestaat uit 20 woorden. Het kind moet steeds een aantal geïsoleerd genoemde fonemen tot woorden samenvoegen: b-oo-t.

3 Foneemsegmentatie.

De toets bestaat uit 20 woorden. Het kind isoleert uit het genoemde woord steeds alle fonemen.

Oriëntatie op boek en schrift

Bij de omgang met boeken ligt de nadruk op de pragmatische aspecten van de geschreven taal, op functionele begrippen die bij het hanteren van de geschreven taal nodig zijn en op de letterkennis. Er zijn drie toetsen:

1 Schriftoriëntatie.

Aan de hand van een klein boekje geeft het kind aan wat voor- en achterkant zijn, waar begonnen wordt, wat een zin, woord, letter is, wat een punt is e.d. De toets bestaat uit 24 vragen.

2 Leesbegrippen.

In deze toets wordt onderzocht of het kind functionele leesbegrippen beheerst, zoals eerste, laatste woord/letter, vooraan, midden, achteraan e.d. Er zijn 20 begrippen. Het zijn begrippen die nodig zijn om een groepsgewijze leesinstructie te kunnen volgen.

3 Klank-letterkoppeling.

De toets bestaat uit een toets met 34 letters. Het kind moet bij iedere letter de juiste klank noemen.

Schema van toetsen

	maximum score
Conceptuele vaardigheden	
1 woordenschat receptief	98
2 woordenschat actief	60
3 definitietaak	64
4 verteltaken	20
Woordbewustzijn	
1 rijmen	10
2 woordobjectivatie	10
3 woordsegmentatie	10
Foneembewustzijn	
1 lettergreepsynthese	20
2 foneemsynthese	20
3 foneemsegmentatie	20
Oriëntatie op boek en schrift	
1 schriftoriëntatie	24
2 leesbegrippen	20
3 klank-letterkoppeling	34

De toetsen werden alle individueel afgenomen. Per afname werd niet meer dan 15 minuten getoetst. Gedurende een tijdsbestek van een maand werden de kinderen op 13 verschillende dagen buiten de klas individueel getoetst. Bij alle drie de metingen zijn steeds dezelfde toetsen gebruikt. Bij de eerste afname waren de resultaten op enkele toetsen bijzonder laag. Bij de laatste meting traden vooral bij stratum 1 kinderen op enkele toetsen plafondeffecten op.

Experimentele conditie

Controleconditie

De controleconditie bestond uit een normaal kleuterschoolprogramma zonder dat daarin bepaalde specifieke activiteiten waren opgenomen. De betreffende scholen mochten in ieder geval niet in het bezit zijn van het programma Schatkist (Mommers en Verhoeven, 1990) of van dat programma gebruik maken. Aan de controleconditie namen 24 leerkrachten met hun groep deel.

De ontwikkelingsgerichte conditie

In de ontwikkelingsgerichte conditie werd op een systematische manier gebruik gemaakt van de thema's van het programma Schatkist. Aan deze conditie werd deelgenomen door kinderen uit alle drie de strata. De leerkrachten (22) die aan deze conditie deelnamen maakten vooraf een plan waarin werd vastgelegd welke thema's per jaar werden uitgevoerd (6 tot 10 thema's per jaar). Deze conditie werd gedurende twee jaar uitgevoerd. Per thema, dat twee tot vier weken in beslag nam, vulden de leerkrachten een evaluatieformulier in waarin werd vastgelegd welke activiteiten waren uitgevoerd met de hele groep en welke activiteiten kinderen individueel gekozen hadden. Ook beoordeelden de leerkrachten in hoeverre deze activiteiten naar tevredenheid waren verlopen.

Voorafgaande aan het experiment en enkele keren per jaar werden de leerkrachten bij hun Schatkistactiviteiten begeleid en werden knelpunten besproken. In het programma Schatkist worden thema's uitgewerkt die beginnen met een uitgebreid verhaal. Op basis van dit verhaal worden allerlei taalspelletjes en opdrachten uitgewerkt zoals dramatiseren, taalspel, creativiteit en opdrachten op werkbladen.

Zowel aan conceptualisering, metalinguïstische aspecten als aan schriftoriëntatie wordt aandacht besteed, met een accent op de betekeniscomponent en minder op de metalinguïstische component. In deze conditie lag het accent meer op het zelf kiezen van de opdrachten dan op het verplichtende karakter daarvan.

Gecombineerde conditie

In de gecombineerde conditie werd de ontwikkelingsgerichte conditie gedurende twee jaar uitgevoerd zoals beschreven. In het tweede jaar werden aan de ontwikkelingsgerichte aanpak voor de dan oudste kleuters, die tot de risicokinderen behoorden, programmagerichte elementen toegevoegd. Gedurende plusminus twintig weken kregen kinderen één à twee keer per week een programmagerichte stimulering. Omdat de programmagerichte stimulering niet binnen de groep door de eigen leerkracht werd uitgevoerd, maar door een leerkracht die buiten de klas werkte, noemen we dit geen geïntegreerde aanpak maar een gecombineerde conditie.

Iedere sessie duurde twintig tot dertig minuten en werd gegeven aan groepjes van zes kinderen die aselekt gekozen waren uit de kinderen van stratum 2 en 3. In veel gevallen zaten in de groepjes Nederlandse kinderen en allochtone kinderen. De verschillen in taalontwikkeling en woordkennis bleken in veel gevallen groot te zijn. De programmagerichte stimulering was ontwikkeld door het CITO (Van Kuyk, in voorbereiding). De uitvoering van het programmagerichte deel vond plaats buiten de klas door drie leerkrachten met een onderwijzers- of kleuterleidsteropleiding en een psycholoog in opleiding. Deze medewerkers werden begeleid door het CITO. Bij deze uitvoering waren kinderen betrokken uit negen klassen, waarvan de leerkrachten ook aan de ontwikkelingsgerichte conditie deelnamen. Gedurende de uitvoering van het programmagerichte deel hielden de medewerkers een protocol bij en overlegden maandelijks over het verloop van het programma. In het programmagerichte deel worden, evenals bij de ontwikkelingsgerichte stimulering, conceptuele, metalinguïstische en schriftoriënterende activiteiten uitgewerkt, eveneens aan de hand van een bepaald thema. Ook hier wordt uitgegaan van een (kort) verhaal waaraan allerlei conceptuele taken worden verbonden zoals woordenschat en het stellen van vragen. Daarnaast zijn er metalinguïstische taken zoals rijm, woord- en foneemsynthese, en woord- en foneemsegmentatie en tenslotte is er oriëntatie op boek en schrift. Alle activiteiten werden uitgevoerd op initiatief van de leerkracht en de opdrachten werden door alle kinderen van de groepjes uitgevoerd. Terwijl het ontwikkelingsgerichte deel voor alle kinderen gewoon doorliep startte het programmagerichte deel in groep twee.

Procedure

Voorafgaande aan het onderzoek werden alle kinderen getoetst. Dit werd na één jaar interventie herhaald en nog eens na het tweede jaar.

Bij de data-analyse wordt allereerst een beschrijving gegeven van de gemiddelden en standaarddeviaties per cluster voor leerlingen uit de onderzoeksstrata 1 (autochtone leerlingen,

hoog en midden milieu), 2 (autochtone leerlingen, laag milieu) en 3 (allochtone leerlingen) in de ontwikkelingsgerichte interventiegroep en de controlegroep bij het begin van groep 1 en op het eind van groep 2.

Vervolgens zijn in groep 2 de resultaten van kinderen uit stratum 2 en 3 in de ontwikkelingsgerichte interventiegroep vergeleken met die van kinderen uit dezelfde strata in de gecombineerde interventiegroep. Hierbij wordt nagegaan in hoeverre een gecombineerde interventie een meerwaarde heeft ten opzichte van een ontwikkelingsgerichte interventie.

Met behulp van variantie-analyse is nagegaan in hoeverre er significante hoofdeffecten konden worden teruggevonden voor Groep, Stratum en Tijd en voor de interactie tussen Tijd enerzijds en Groep en Stratum anderzijds.

RESULTATEN ONTWIKKELINGSGERICHTE INTERVENTIE

In de onderstaande tabellen worden de gegevens vergeleken van de controlegroep en de experimentele groep die gedurende twee jaar een ontwikkelingsgerichte stimulering kreeg (ontwikkelingsgerichte interventie). Afzonderlijk worden de resultaten van Stratum 1, 2 en 3 vergeleken van de voormeting (1e afname) en de nameting (3e afname) na twee jaar. De resultaten worden weergegeven in vier clusters van vaardigheden geoperationaliseerd in steeds drie of vier toetsen. Aan de tabellen van iedere cluster is een variatieanalyse toegevoegd. Hierin worden P-waarden van de hoofdeffecten (Tijd, Stratum en Groep) weergegeven. Indien deze niet significant zijn wordt vermeld: N.S. Daar waar sprake is van interacties worden de waarden apart vermeld, per toets, aangegeven met 1), 2), 3), of 4).

Conceptuele vaardigheden

In Tabel 1 worden de gemiddelden en standaarddeviaties van de kinderen uit de onderzoeksstrata 1 (autochtoon, hoog milieu), 2 (autochtoon, laag milieu) en 3 (allochtoon) weergegeven op het cluster "conceptueel". Hieronder vallen woordenschat, receptief en actief, de definitievaardigheid en de vertelvaardigheid.

Uit de variantie-analyse bij het cluster conceptuele vaardigheden (Tabel 1), blijkt dat Tijd en Stratum steeds significant zijn en dat bij vertelvaardigheid ook het hoofdeffect Groep significant is. Hetzelfde geldt voor de interactie Tijd x Stratum en Tijd x Groep bij vertelvaardigheid. Wij kunnen concluderen dat er in de tijd beduidende verschillen zijn tussen de onderzoeksstrata, dat de verschillen tussen de strata geleidelijk kleiner worden en dat met betrekking tot vertelvaardigheid de vooruitgang in de tijd voor de experimentele groep groter is dan voor de controlegroep.

Woordbewustzijn

In Tabel 2 worden de gemiddelden en standaarddeviaties weergegeven op het cluster "Woordbewustzijn". Hieronder vallen rijmen, objectivatie en woordsegmentatie.

Variatie-analyse bij Tabel 2 laat zien dat bij het cluster woordbewustzijn de hoofdeffecten Tijd en Stratum significant zijn, behalve bij stratum woordsegmentatie. Voor objectivatie is ook het hoofdeffect Groep significant. Ook de interactie Tijd x Stratum is steeds significant en de interactie Tijd x Groep voor objectivatie en woordsegmentatie. Uit Tabel 4 kunnen we concluderen dat er verschillen zijn tussen de strata die geleidelijk kleiner worden en uit de variantie-analyse, dat de vooruitgang in de tijd voor de experimentele groep bij objectivatie en woordsegmentatie groter is dan voor de controlegroep.

Foneembewustzijn

In Tabel 3 worden gemiddelden en standaarddeviaties van de kinderen uit de drie onderzoeksstrata vermeld op het cluster Foneembewustzijn, bestaande uit lettergreepsynthese, foneemsynthese en foneemsegmentatie.

Uit de variatieanalyse bij Tabel 3 blijkt, dat het hoofdeffect Tijd significant is, behalve bij

Tabel 1. Conceptuele vaardigheden.

		Voormeting		Nameting	
		x	SD	x	SD
• Passieve woordenschat					
stratum 1:	experimentele groep	43.85	18.27	70.10	15.28
	controlegroep	42.78	17.20	73.74	13.13
stratum 2:	experimentele groep	32.93	13.62	58.32	17.57
	controlegroep	32.92	11.30	54.08	13.25
stratum 3:	experimentele groep	13.60	12.26	36.07	10.16
	controlegroep	18.60	7.04	41.60	19.83
• Aktieve woordenschat					
stratum 1:	experimentele groep	24.41	9.20	40.12	8.33
	controlegroep	27.36	11.23	40.71	8.79
stratum 2:	experimentele groep	19.89	6.97	33.03	7.10
	controlegroep	19.75	6.01	33.46	7.59
stratum 3:	experimentele groep	6.75	6.08	19.44	8.04
	controlegroep	6.75	4.54	19.58	8.40
• Definitievaardigheid					
stratum 1:	experimentele groep	22.75	11.74	36.58	9.57
	controlegroep	21.74	8.99	36.63	8.88
stratum 2:	experimentele groep	15.97	8.83	28.32	8.13
	controlegroep	16.57	8.95	29.52	8.32
stratum 3:	experimentele groep	3.87	5.95	17.40	10.09
	controlegroep	5.91	6.79	19.27	9.66
• Vertelvaardigheid					
stratum 1:	experimentele groep	5.79	2.88	11.50	2.18
	controlegroep	8.00	2.97	11.44	2.73
stratum 2:	experimentele groep	5.76	3.45	9.91	2.43
	controlegroep	6.86	3.12	9.82	2.24
stratum 3:	experimentele groep	3.33	2.97	9.53	2.92
	controlegroep	6.10	3.41	10.00	1.83

Variantie-analyse bij cluster
conceptuele vaardigheden

Hoofdeffecten

	Tijd	Stratum	Groep
Woordenschat passief	P:<.001	P:<.001	N.S.
Woordenschat actief	P:<.001	P:<.001	N.S.
Definitievaardigheid	P:<.001	P:<.001	N.S.
4) Vertelvaardigheid	P:<.001	P:<.001	P:<.05

4) Interactie Tijd x Groep = P:<.001
Interactie Tijd x Stratum = P:<.05

Tabel 2. Woordbewustzijn.

		Voormeting		Nameting	
		x	SD	x	SD
• Rijmen					
stratum 1:	experimentele groep	6.28	3.4	9.58	1.55
	controlegroep	6.32	3.64	9.57	1.00
stratum 2:	experimentele groep	3.21	2.59	8.68	2.51
	controlegroep	4.22	3.62	8.68	2.51
stratum 3:	experimentele groep	1.50	2.00	7.38	2.50
	controlegroep	2.25	2.05	8.00	2.76
• Woordobjectivatie					
stratum 1:	experimentele groep	5.58	1.54	7.23	1.95
	controlegroep	6.07	2.42	7.96	1.91
stratum 2:	experimentele groep	5.04	1.77	6.89	1.91
	controlegroep	6.39	1.37	7.26	2.40
stratum 3:	experimentele groep	3.27	2.49	6.53	2.42
	controlegroep	4.50	1.90	6.80	2.66
• Woordsegmentatie					
stratum 1:	experimentele groep	3.54	2.68	5.44	1.96
	controlegroep	4.93	2.37	6.07	2.06
stratum 2:	experimentele groep	3.49	1.87	6.29	1.96
	controlegroep	3.25	3.18	4.96	2.11
stratum 3:	experimentele groep	2.31	3.03	6.44	2.58
	controlegroep	3.17	2.89	4.92	2.19
Variantie-analyse bij cluster		Hoofdeffecten			
Woordbewustzijn		Tijd	Stratum	Groep	
1) Rijm		P:<.001	P:<.001	N.S.	
2) Objectivatie		P:<.001	P:<.001	P:<.01	
3) Woordsegmentatie		P:<.001	N.S.	N.S.	
1) Interactie	Tijd x Stratum = P:<.001				
2) Interactie	Tijd x Groep = P:<.05				
	Tijd x Stratum = P:<.05				
3) Interactie	Tijd x Groep = P:<.01				
	Tijd x Stratum = P:<.05				

lettergreepsynthese en dat alleen bij lettergreepsynthese significante verschillen bestaan bij Stratum en Groep. Hier is ook de interactie Tijd x Stratum en Tijd x Groep significant. Hieruit kunnen wij concluderen dat de verschillen tussen de strata geringer worden of afwezig zijn en dat de vooruitgang in de tijd voor de experimentele groep m.b.t. lettergreepsynthese groter is dan voor de controlegroep.

Tabel 3. Foneembewustzijn.

		Voormeting		Nameting	
		x	SD	x	SD
• Lettergreepsynthese					
stratum 1:	experimentele groep	16.18	5.68	19.18	2.56
	controlegroep	17.82	3.08	19.57	.88
stratum 2:	experimentele groep	11.69	6.94	17.97	3.37
	controlegroep	16.00	4.82	18.83	4.08
stratum 3:	experimentele groep	11.13	6.99	15.75	7.89
	controlegroep	10.08	7.4	15.92	6.78
• Foneemsynthese					
stratum 1:	experimentele groep	2.85	3.48	7.50	6.20
	controlegroep	3.39	3.83	8.07	6.24
stratum 2:	experimentele groep	2.03	2.77	6.83	5.06
	controlegroep	2.42	3.13	7.79	5.28
stratum 3:	experimentele groep	2.38	2.42	7.79	5.28
	controlegroep	1.50	1.17	5.17	5.46
• Foneemsegmentatie					
stratum 1:	experimentele groep	.68	2.14	4.93	5.92
	controlegroep	.93	1.94	5.07	5.66
stratum 2:	experimentele groep	.31	.72	2.14	4.39
	controlegroep	.21	.59	5.29	5.36
stratum 3:	experimentele groep	.55	1.55	3.31	4.39
	controlegroep	.58	1.73	4.83	4.90
Variantie-analyse bij cluster Foneembewustzijn		Hoofdeffecten			
		Tijd	Stratum	Groep	
1) Lettergreepsynthese		P:<.10 (NS)	P:<.001	P:<.05	
Foneemsynthese		P:<.001	N.S.	N.S.	
Foneemsegmentatie		P:<.001	N.S.	N.S.	

1) Tijd x Groep: P:<.05

Tijd x Stratum: P:<.05

Oriëntatie op boek en schrift

Tabel 4 laat de resultaten zien van de drie onderzoeksstrata op het cluster Oriëntatie op boek en schrift, bestaande uit schriftoriëntatie, leesbegrippen en klank-letterkoppeling.

Variantie-analyse bij Tabel 4 laat zien dat de hoofdeffecten Tijd en Stratum significant zijn evenals de interacties tussen Stratum en Tijd. Voor het hoofdeffect Groep zijn de verschillen niet significant. Wij kunnen concluderen dat er geleidelijk kleiner wordende verschillen zijn tussen de onderzoeks-strata en dat kinderen uit stratum 3 grote vooruitgang boeken in de tijd. Experimentele- en controlegroep gaan in gelijke mate vooruit op dit cluster van vaardigheden.

Tabel 4. Oriëntatie op boek en schrift.

		Voormeting		Nameting	
		x	SD	x	SD
• Schriftoriëntatie					
stratum 1:	experimentele groep	8.81	5.20	15.56	2.42
	controlegroep	9.11	5.59	14.54	3.04
stratum 2:	experimentele groep	5.80	3.39	14.06	4.01
	controlegroep	5.65	2.81	14.04	3.47
stratum 3:	experimentele groep	3.94	3.13	11.20	3.90
	controlegroep	3.82	2.79	10.90	3.39
• Leesbegrippen					
stratum 1:	experimentele groep	12.13	4.56	17.03	3.42
	controlegroep	10.78	5.87	17.85	3.02
stratum 2:	experimentele groep	7.45	3.35	15.10	3.89
	controlegroep	6.83	5.91	15.30	4.87
stratum 3:	experimentele groep	4.50	4.77	11.79	4.25
	controlegroep	4.00	4.09	11.83	5.51
• Klank-letter-koppeling					
stratum 1:	experimentele groep	1.65	2.86	8.75	7.88
	controlegroep	2.18	3.82	7.39	6.71
stratum 2:	experimentele groep	.50	1.11	3.91	6.48
	controlegroep	1.42	3.44	4.92	7.43
stratum 3:	experimentele groep	.56	1.03	7.13	6.83
	controlegroep	.46	.52	5.46	6.48
Variantie-analyse bij cluster		Hoofdeffecten			
Oriëntatie op boek en schrift		Tijd	Stratum	Groep	
1) Schriftoriëntatie		P:<.001	P:<.001	N.S.	
2) Leesbegrippen		P:<.001	P:<.001	N.S.	
3) Klank-letterkoppeling		P:<.001	P:<.01	N.S.	
1) Interactie Stratum x Tijd: P:<.05					
2) Interactie Stratum x Tijd: P:<.05					
3) Interactie Stratum x Tijd: P:<.05					

RESULTATEN GECOMBINEERDE INTERVENTIE

In onderstaande tabellen worden resultaten vergeleken van de gecombineerde interventie met die van de ontwikkelingsgerichte interventie. Nagegaan wordt in hoeverre de eerste een surplus biedt ten opzichte van de laatstgenoemde. Met andere woorden: draagt het toevoegen van programmagerichte elementen aan een ontwikkelingsgerichte benadering bij aan de verbetering van de resultaten. De vergelijking geldt uitsluitend voor de zwakste groepen: de strata 2 en 3.

Conceptuele vaardigheden

In Tabel 5 worden gemiddelden en standaarddeviaties vergeleken van de gecombineerde (G-groep) en de ontwikkelingsgerichte interventie (O-groep) van stratum 2 en 3 van het cluster Conceptuele vaardigheden.

Tabel 5. Conceptuele vaardigheden.

		Voormeting		Nameting	
		x	SD	x	SD
• Passieve woordenschat					
stratum 2:	G-groep	43.82	16.70	53.93	18.84
	O-groep	41.14	18.32	58.32	17.57
stratum 3:	G-groep	24.57	16.60	42.86	13.98
	O-groep	21.69	11.93	36.07	10.16
• Aktieve woordenschat					
stratum 2:	G-groep	25.89	8.44	34.52	9.20
	O-groep	25.88	6.85	33.03	7.10
stratum 3:	G-groep	10.43	8.34	20.64	8.61
	O-groep	10.00	7.21	19.07	8.18
• Definitievaardigheid					
stratum 2:	G-groep	17.82	10.66	28.30	10.89
	O-groep	19.91	9.45	28.32	8.13
stratum 3:	G-groep	6.64	8.22	17.57	10.58
	O-groep	6.47	6.64	17.40	10.09
• Vertelvaardigheid					
stratum 2:	G-groep	5.96	3.12	8.85	2.11
	O-groep	6.41	2.65	9.91	2.39
stratum 3:	G-groep	4.86	3.96	9.57	2.65
	O-groep	4.87	2.85	9.53	2.92

Variantie-analyse bij cluster
Conceptuele vaardigheden

Hoofdeffecten

	Tijd	Stratum	Groep
1) Passieve woordenschat	P:<.001	P:<.001	N.S.
Aktieve woordenschat	P:<.001	P:<.001	N.S.
Definitievaardigheid	P:<.001	P:<.001	N.S.
4) Verteltaak	P:<.001	N.S.	N.S.

1) Interactie Tijd x Stratum: P:<.05

4) Interactie Tijd x Stratum: P:<.05

Uit de variantie-analyse bij Tabel 5 blijkt dat de hoofdeffecten Tijd en Stratum significant zijn, behalve bij Stratum verteltaak. Het hoofdeffect Groep is niet significant. Interactie tussen Tijd x Stratum bij passieve woordenschat en verteltaak is wel significant. Hieruit kunnen wij concluderen dat de gecombineerde interventie geen extra effect heeft boven op de ontwikkelingsgerichte interventie, al geven de gemiddelden van Tabel 7 wel aan dat allochtone kinderen meer van de gecombineerde aanpak lijken te profiteren op passieve en actieve woordenschat. Dat geldt niet voor de kinderen in stratum 2.

Woordbewustzijn

In Tabel 6 worden de gemiddelden en standaarddeviaties weergegeven van de strata 2 en 3 op het cluster Woordbewustzijn, inhoudende rijm, objectivatie en woordsegmentatie.

Tabel 6. Woordbewustzijn.

		Voormeting		Nameting	
		x	SD	x	SD
• Rijmen					
stratum 2:	G-groep	5.41	3.41	8.15	3.10
	O-groep	5.24	3.73	8.77	2.51
stratum 3:	G-groep	3.29	3.31	8.07	2.27
	O-groep	1.07	1.28	7.20	2.48
• Woordobjectivatie					
stratum 2:	G-groep	6.15	2.01	7.48	1.72
	O-groep	5.65	2.15	6.91	1.87
stratum 3:	G-groep	5.57	3.06	7.93	2.56
	O-groep	4.80	1.78	6.67	2.50
• Woordsegmentatie					
stratum 2:	G-groep	3.56	1.87	4.96	2.08
	O-groep	4.83	1.95	6.29	1.96
stratum 3:	G-groep	3.71	2.46	5.36	1.98
	O-groep	4.67	2.66	6.40	2.67
Variantie-analyse bij cluster		Hoofdeffecten			
Woordbewustzijn		Tijd	Stratum	Groep	
1) Rijm		P:<.001	P:<.001	N.S.	
Objectivatie		P:<.001	N.S.	P:<.05	
Woordsegmentatie		P:<.001	N.S.	P:<.01	

1) Interactie Tijd x Stratum: P:<.001

Variantie-analyse bij Tabel 6 laat zien dat het hoofdeffect Tijd significant is, Stratum is alleen significant voor rijm. Het hoofdeffect Groep is significant, behalve bij rijm. Interactie Tijd x Stratum is wel significant voor rijm. Hieruit kunnen wij concluderen dat er aanzienlijke verschillen zijn tussen de G-groep en de O-groep, hoewel beide groepen vergelijkbaar vooruitgang boeken op woordobjectivatie en woordsegmentatie.

Foneembewustzijn

In Tabel 7 zijn de gemiddelden en standaarddeviaties vergeleken van de gecombineerde en ontwikkelingsgerichte interventie van strata 2 en 3 van het cluster Foneembewustzijn.

Tabel 7. Foneembewustzijn.

		Voormeting		Nameting	
		x	SD	x	SD
• Lettergreepsynthese					
stratum 2:	G-groep	12.19	8.25	18.48	2.38
	O-groep	14.29	6.97	17.97	3.37
stratum 3:	G-groep	12.93	6.80	18.71	2.38
	O-groep	14.20	8.49	15.47	8.08
• Foneemsynthese					
stratum 2:	G-groep	3.26	3.13	6.67	6.14
	O-groep	2.66	3.10	3.83	5.06
stratum 3:	G-groep	3.29	4.41	9.14	5.57
	O-groep	2.80	4.89	4.20	4.78
• Foneemsegmentatie					
stratum 2:	G-groep	1.00	1.98	3.30	4.94
	O-groep	.57	2.99	2.14	4.39
stratum 3:	G-groep	.93	1.73	4.36	4.77
	O-groep	1.33	2.99	3.20	4.52

Variantie-analyse bij cluster
Foneembewustzijn

Hoofdeffecten

	Tijd	Stratum	Groep
1) Lettergreepsynthese	P:<.001	N.S.	N.S.
2) Foneemsynthese	P:<.001	N.S.	P:<.05
Foneemsegmentatie	P:<.001	N.S.	N.S.

1) Tijd x Groep: P:<.05

2) Tijd x Groep: P:<.001

Variantie-analyse bij Tabel 7 laat zien dat het hoofdeffect Tijd significant is, maar Stratum en Groep niet, behalve Foneemsynthese bij Groep. Wel zijn de interacties Tijd x Groep van lettergreepsynthese en foneemsynthese significant. Dit duidt erop dat de gecombineerde aanpak een extra effect heeft ten opzichte van de ontwikkelingsgerichte conditie m.b.t. lettergreepsynthese en foneemsynthese. Bij deze laatste vaardigheid blijken de verschillen tussen de groepen ook groot te zijn. Bij foneemsegmentatie wijzen de verschillen in dezelfde richting. Als we Tabel 9 bezien lijken allochtone kinderen extra te profiteren van de gecombineerde aanpak boven de kinderen van stratum 2.

Oriëntatie op boek en schrift

In Tabel 8 geven wij de gemiddelden en standaarddeviaties van strata 2 en 3 van de gecombineerde en ontwikkelingsgerichte interventie van het cluster Oriëntatie op boek en schrift.

Uit de variantie-analyse bij Tabel 8 blijkt, dat de hoofdeffecten Tijd en Stratum significant zijn, behalve klank-letterkoppeling bij Stratum. Het hoofdeffect Groep is niet significant. Interactie Tijd x Stratum is significant bij leesbegrippen, evenals de interactie Tijd x Stratum x Groep.

Hieruit kunnen wij concluderen dat de verschillen tussen de onderzoeksstrata afnemen en dat

Tabel 8. Oriëntatie op boek en schrift.

		Voormeting		Nameting	
		x	SD	x	SD
• Schriftoriëntatie					
stratum 2:	G-groep	9.04	4.67	14.07	4.04
	O-groep	8.31	3.41	14.06	4.01
stratum 3:	G-groep	5.64	3.61	13.50	2.57
	O-groep	5.67	3.74	10.87	3.82
• Leesbegrippen					
stratum 2:	G-groep	10.68	5.49	13.80	5.22
	O-groep	10.39	4.82	15.10	3.89
stratum 3:	G-groep	6.21	5.55	14.14	5.42
	O-groep	6.92	4.19	12.00	4.34
• Klank-letter-koppeling					
stratum 2:	G-groep	1.41	2.53	5.00	6.40
	O-groep	1.41	3.56	3.91	6.48
stratum 3:	G-groep	.86	2.14	5.21	6.03
	O-groep	3.53	5.44	7.20	7.06
Variantie-analyse bij cluster		Hoofdeffecten			
Oriëntatie op boek en schrift		Tijd	Stratum	Groep	
Schriftoriëntatie		P:<.001	P:<.01	N.S.	
2) Leesbegrippen		P:<.001	P:<.01	N.S.	
Klank- en letterkoppeling		P:<.001	N.S.	N.S.	

2) Interactie Tijd x Stratum: P:<.05

Interactie Tijd x Stratum x Groep: P:<.05

bij leesbegrippen de ontwikkelingsgerichte aanpak voor stratum 2 kinderen effectief is maar voor de allochtone kinderen de gecombineerde aanpak. Ook bij schriftoriëntatie lijken allochtone kinderen meer te profiteren dan de kinderen van stratum 2. Met betrekking tot klank-letterkoppeling lijken de kinderen van de gecombineerde aanpak extra profijt te hebben ten opzichte van de ontwikkelingsgerichte interventie.

CONCLUSIES EN DISCUSSIE

Uit het onderhavige onderzoek komen enkele opvallende conclusies naar voren. Allereerst blijken kinderen op alle onderscheiden conceptuele en metalinguïstische onderdelen substantiële vooruitgang te boeken. Op de taken inzake fonologisch bewustzijn, uitgezonderd het onderdeel Lettergreepsynthese, blijkt de vooruitgang relatief gering.

Verder blijken de verschillen in gemiddelde scores tussen kinderen in de drie onderscheiden onderzoeksstrata op veel taken aanzienlijk. Vooral tussen autochtone en allochtone kinderen is sprake van grote verschillen. Dit geldt met name voor de woordenschat- en tekstuele taken en voor de onderdelen Rijm en Lettergreepsynthese. Op de overige taken op het terrein van objec-

tivatie en fonologisch bewustzijn blijken de verschillen in scores tussen de drie onderzoeksstrata beduidend geringer. In veel gevallen is in de loop van de kleuterperiode sprake van nivellering op deze onderdelen. Hieruit zou kunnen worden afgeleid dat allochtone kinderen veel hulp nodig hebben m.b.t. woordenschat en tekstuele taken, maar dat dit in veel mindere mate geldt voor metalinguïstische taken. Voorts blijkt er sprake van differentiële effecten van de onderscheiden interventies.

Als we ons richten op de eerste onderzoeksvraag dan blijkt de ontwikkelingsgerichte interventie ten opzichte van een controlegroep te leiden tot een significant grotere vooruitgang voor de onderdelen Vertelvaardigheid, Woordobjectivatie, Woordsegmentatie en Lettergreepsynthese.

Ten aanzien van de tweede onderzoeksvraag kunnen we concluderen dat een extra program-magerichte stimulering in aanvulling op de ontwikkelingsgerichte aanpak (de gecombineerde conditie), een positief effect heeft op de ontwikkeling van risico kinderen. De positieve effecten treden met name op bij de onderdelen Lettergreepsynthese en Foneemsynthese. Op foneemsegmentatie, grafeemkennis en actieve woordenschat leidt een gecombineerde interventie tot hogere scores. Bij passieve woordenschat en leesbegrippen geldt dat alleen voor de allochtone groep. Behalve voor het onderdeel Leesbegrippen blijken de gevonden verschillen echter niet significant.

Samenvattend kunnen we concluderen dat een ontwikkelingsgerichte interventie vooral een positief effect heeft op de tekstuele vaardigheid en op het vermogen tot objectivatie van gesproken taal.

De ontwikkelingsgerichte aanpak blijkt vooral dankzij het regelmatig houden van kringgesprekken en het beluisteren en dramatiseren van verhaalstructuren de vertelvaardigheid van kinderen te vergroten. Bovendien blijken spelvormen, die kinderen doen stilstaan bij vormaspecten van taal, behulpzaam bij het op gang brengen van het vermogen tot reflectie op gesproken taal.

De gecombineerde interventie blijkt vooral additionele effecten te leveren op taken die betrekking hebben op fonologisch bewustzijn. Blijkbaar is een gerichte instructie vooral succesvol bij het stimuleren van het bewustzijn dat woorden zijn opgebouwd uit klanken.

Als eerste discussiepunt merken we op dat we er niet van uitgaan dat de gekozen benaderingen beter zouden werken voor kinderen uit bepaalde strata. De enige aanname is dat de gecombineerde aanpak primair geschikt is voor risicokinderen, maar dat is reeds in het design ingebouwd. Verder gaan we ervan uit dat de ontwikkelingsgerichte benadering voor de twee overgebleven strata even goed werkt. Het uitblijven (op een na) van de TxGxS interacties ondersteunt dit beeld.

Opvallend is dat beide interventieprogramma's geen significante effecten te zien geven op woordenschatmaten. Hiervoor zijn tenminste twee mogelijke verklaringen te geven. Ten eerste geldt dat in de controleconditie evenzeer onderwijs wordt gegeven in het domein van vroege geletterdheid. Met name onderdelen als kringgesprek, voorlezen en vertellen zijn te beschouwen als klassieke werkvormen binnen het kleuteronderwijs. Als tweede mogelijke verklaring geldt het feit dat de gekozen maten onafhankelijk zijn ten opzichte van de aangeboden leerstof. De kans dat woorden uit de gekozen toetsen in de interventies zijn aangeleerd, is uitermate gering. Gegeven het feit dat het bij woordenschat om domeinspecifieke kennis gaat mag de kans op transfer van in een programma aangeleerde kennis naar andere kennisdomeinen die in de toets aan bod komen, gering worden geacht. Brainerd (1975) en Kingma (1981) onderscheiden in dit verband drie soorten transfer. De eerste is near-near transfer, waarbij de gekozen maten dezelfde zijn als de aangeboden leerstof. Bij near-far transfer zijn de gekozen maten onafhankelijk van aangeboden leerstof, zoals we dat bij de woordenschattaken aangeven. Far-Far transfer geeft een generalisatie naar andere dan de gestimuleerde vaardigheden. In feite geldt voor zowel de ontwikkelingsgerichte als de gecombineerde interventie een near-far transfer. Toetsen die gebaseerd zijn op de concrete inhoud, bijv. bepaalde aangeleerde woorden, klanken, bepaalde geanalyseerde woorden, zijn niet gebruikt. Vooral bij specifieke programmagerichte elementen

zou de interventie dan nauwkeuriger in kaart zijn gebracht. Daar tegenover staat dat er dan slechts sprake zou zijn geweest van near-near transfer. De gevonden effecten liggen nu in alle gevallen op het hogere near-far generaliseerbaarheidsniveau.

Ten aanzien van de gecombineerde benadering valt verder op te merken dat het programmagerichte deel als gevolg van financiële beperkingen alleen in het tweede leerjaar kon worden aangeboden. Hierdoor kon het programma in het eerste jaar niet worden voorbereid. Bovendien kon de hulp niet individueel worden uitgevoerd maar in groepjes van zes kinderen. Al vanaf het begin bleek dat de verschillen tussen de kinderen in de groepjes groot waren. Er waren grote verschillen in ontwikkeling en grote verschillen in taalontwikkeling. De aandacht ging daardoor ofwel sterk naar de zwakste kinderen of naar kinderen die al een redelijke taalontwikkeling hadden. Door de grote verschillen in de groepen konden de individuele kinderen onvoldoende profiteren van de inhoud van het programma. Groepjes met een meer homogene problematiek zouden wellicht beter gewerkt hebben, maar nog meer zou een individuele vorm van interventie te prefereren zijn geweest in de vorm van tutoring (individuele hulp zolang de achterstand duurt), zoals die is toegepast in het programma Success for All. Dit programma bleek blijvende effecten te hebben (Wasik & Slavin, 1993). Opgemerkt moet wel worden dat de tutoring werd toegepast in groep drie en niet in het kleuteronderwijs.

Ondanks beperkingen in het onderzoeksdesign blijkt een ontwikkelingsgerichte interventie effectief te zijn en blijkt een combinatie van ontwikkelings- en programmagerichte elementen een extra effect te bewerkstelligen. In deze zin kunnen de gestelde verwachtingen worden bevestigd. Uit meta-analyses van interventiestudies blijkt, dat interventies intensief en voldoende lang moeten zijn. Als vuistregel voor interventieduur wordt daarbij een periode van twee à drie jaar genoemd (Leseman, 1992). De ontwikkelingsgerichte interventie was door het vrijblijvende karakter niet intensief en de programmagerichte elementen in de gecombineerde conditie beperkt in tijd (20 weken, 2x20 minuten). Ondanks deze beperkte intensiteit en tijdsduur geven de resultaten een duidelijke richting aan. Gezien het beperkt aantal proefpersonen echter verdient het aanbeveling het onderzoek te repliceren. Gestreefd moet dan worden naar een meer ideale onderzoeksofzet waarin ook stratum 1 kinderen zijn opgenomen en waarin wordt voorzien in een grotere intensiteit en tijdsduur, met tutoring als strategie in de gecombineerde conditie.

LITERATUUR

- ARBO (1990). *Spelen en Jongleren. Advies over vernieuwing van het onderwijs aan jonge kinderen*. Arbo: Zeist.
- Ball, E.W., & Blachman, B.A. (1991). *Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling?*
- Boland, Th., & Mommers, M.J.C. (1991). Het relatieve belang van het beheersen van leesvoorwaarden. *Pedagogische Studiën* 68, 390-403.
- Bradley, L., & Bryant, P.E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature*, 30, 419-221.
- Brainerd, C.J. (1975). Structures of the Whole and elementary Education. *American Educational Research Journal*, 1 (3), 369-378.
- Cunningham, A.E. (1988). *A metacognitive approach to phonemic awareness*. University of Michigan: Dissertation.
- Elkind, D. (1986). *The hurried child*. Reading, MA: Addison Wesley.
- Elley, W. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, 24, (2), 174-187.
- Ellwein, M.C., Walsh, D.J., Eads II, G.M. & Miller, A. (1991). Using readiness tests to route kindergarten students: The snarled intersection of psychometrics, policy and practice. *Educational evaluation and policy analysis*. Vol. 13, (2) 159-176.
- Groenendaal, H.J. (1978). *Vroegtijdige hulpverlening aan zwakfunctionerende kleuters*. Verslag van een onderzoek. Amsterdam (proefschrift).

- Kingma, J. (1981). *De ontwikkeling van quantitative en relationele begrippen bij kinderen van 3-12 jaar*. Swets en Zeitlinger, Lisse.
- Kuyk, J.J. van (1985). *Ruimtelijke Oriëntatie. Ontwikkeling en validering van een observatie- en hulpprogramma voor kleuters in de basisschool*. Tilburg: Zwijsen.
- Kuyk, J.J. van (1991a). Een schets voor het onderwijs aan jonge kinderen (4-8 jaar). *Pedagogische Studiën* 68, 378-389.
- Kuyk, J.J. van (1991b). Effectief onderwijs in basisvaardigheden, een reactie. In J.J. van Kuyk & J. Claessen. *De Grensverleggende Basisschool*. Groningen: Wolters-Noordhoff.
- Kuyk, J.J. van (1992a). Visie op het onderwijs aan jonge kinderen. In B. van Oers en F. Janssen-Vos. *Visies op onderwijs aan jonge kinderen*, pp. 15-35. Assen: Van Gorcum.
- Kuyk, J.J. van (1992b). *Taal- en voorbereiding op het lezen. Hulpprogramma voor kleuters in de basisschool. Eerste conceptversie*. Arnhem: Cito.
- Lamb, M., & Easterbrooks, E. (1981). Individual differences in parental sensitivity: origins, components and consequences. In M.E. Lamb & L.R. Scherrod (red.). *Infant social cognition*. Hillsdale, N.J.: Lawrence Erlbaum.
- Leseman, P.P.M. (1992). Doel en strategie van voorschoolse stimuleringsprogramma's. In J. Rispens en W.F. van der Meulen. *Gezinsgerichte stimulering van kinderen in achterstandssituaties*. Lisse: Swets & Zeitlinger.
- Lundberg, I., Frost, J., & Petersen, O.P. (1988). Effects of an exclusive program for stimulativ phonological awareness in preschool children. *Reading Research Quarterly*, XXIII/3, 263-284.
- Mason, J., McCormick, C., & Bhavnagri, N. (1986). How are you going to help me learn? In: B. Yaden & S. Templeton (eds.). *Metalinguistic awareness and beginning literacy*. New York: Heinemann.
- Mommers, C., & Verhoeven, L. (1990). Kinderen oriënteren op geschreven taal. *School en Begeleiding*, 7, (27), 28-33.
- Morrow, L.M., & Smit, J.K. (1990). The effects of group setting on interactive storybook reading. In *Reading Research Quarterly*, 25, Vol. 3, pp. 196-212.
- Oloffson, A., & Lundberg, I. (1985). Can phonemic awareness be trained in kindergarten? *Scandinavian Journal of Psychology*, 24, 35-44.
- Parreren, C.F. van (1985). Van kleuter tot schoolkind: continuïteit in het leren? *Pedagogische Studiën*, 62, 174-183.
- Sixma, J. (1973). *Leesvoorwaarden*. Groningen: Tjeenk-Willink.
- Slavin, R.E., Madden, N.A., Karweit, N.L., Livermon, B.J., & Dolan, L. (1990). Success For All: first year outcomes of a comprehensive plan for reforming urban education. *American Educational Research Journal*, 27, (2), 279-300.
- Slavin, R.E., Karweit N.L., & Wasik, B.A. (1994). *Preventing early school failure*. Boston: Allyn and Bacon.
- Stevens, L.M. (1987). Vroegtijdige onderkenning van in hun ontwikkeling bedreigde kleuters. In K. Doornbos en L.M. Stevens. *De groei van het speciaal onderwijs*. 's Gravenhage: Staatsuitgeverij.
- Sulzby E., & Teale, W. (1991). Emergent Literacy. In R. Barr, M. Kamii, P. Mosenthal, a.o. (eds.). *Handbook of reading research. Vol 2*. pp. 727-757. New York.
- Teale, W.H., & Sulzby, E. (1986). *Emergent literacy, writing and reading*. Norwood, NJ: Ablex.
- Verhoeven, L. Th., & Vermeer, A. (1986). *Taaltoets Allochtone Kinderen*. Tilburg: Zwijsen.
- Verhoeven, L., & Mommers, C. (1989). De didactiek van het leren lezen. *School en begeleiding*, 21, 60-64.
- Verhoeven, L. (1990). Acquisition of reading in a second language. *Reading Research Quarterly*, 15, 90-114.
- Verhoeven, L., & van Kuyk, J.J. (1991). Peiling van conceptuele en metalinguïstische kennis bij de aanvang van het basisonderwijs. *Pedagogische Studiën* 68, 415-425.
- Verhoeven, L. (1992). Beginnende geletterdheid: een ontwikkelingsperspectief. *Pedagogische Studiën* 69, 332-351.
- Verhoeven, L. (1994). *Ontluikende geletterdheid*. Lisse: Swets en Zeitlinger.
- Verhoeven, L., & van Kuyk, J.J. (1995). *Ontwikkeling van beginnende geletterdheid bij kleuters*. In voorbereiding.
- Wasik, B.A., & Slavin, R.E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly Vol. 28, no. 2*, pp. 179-200.
- Williams, J. (1986). *The role of phonemic analysis in reading. Psychological and educational perspectives on learning disabilities*. New York: Academic Press.

Ontwikkeling van tweetaligheid bij allochtone kleuters

Ludo Verhoeven en Goretti Narain

Faculteit der Letteren, Katholieke Universiteit Brabant

ABSTRACT

In the present study the process of early bilingual development of Turkish, Antillean and Moroccan children living in the Netherlands is evaluated. Moreover, an attempt is made to explain the individual variation in the children's language development. A bilingual test is constructed, monitoring children's first and second language development in the age range from 4 to 7. The results show a large variation in the patterns of language development in the three groups of children. Turkish children turn out to be dominant in their mother tongue, Antillean children have a more balanced bilingual proficiency level, whereas the Moroccan children held an intermediate position. There is positive evidence for the interdependency hypothesis stating that in bilingual development a great deal of transfer of abilities from the dominant language to the dominated language is taking place. The variation in language proficiency can primarily be explained from sociocultural factors, such as language input and cultural orientation.

INLEIDING

Een belangrijke vraag is hoe kinderen van etnische minderheden in de onderbouw van het basisonderwijs het beste kunnen worden begeleid, gezien hun taalsituatie. Differentiatie en individualisering lijken noodzakelijk om op adequate wijze op de behoeften van allochtone leerlingen te kunnen inspelen. Om op deze complexe onderwijsleercontext beter zicht te krijgen zijn longitudinale onderzoeksgegevens wenselijk die de individuele variatie in aard en mate van tweetaligheid in de onderbouwperiode van het basisonderwijs blootleggen en verklaren. Om het tempo van tweede-taalverwerving van allochtone kinderen te kunnen inschatten is het van belang dat er tevens gegevens beschikbaar komen omtrent de taalontwikkeling van autochtone leeftijdgenoten (Verhoeven & Vermeer 1989). Met het oog op beleidsontwikkeling inzake meertaligheid en meertalig onderwijs is het verder van belang een beeld te krijgen van de sociolinguïstische en onderwijskundige variabelen die het tempo van eerste- en tweede-taalontwikkeling van allochtone kinderen mede bepalen.

Onderzoek naar de ontwikkeling van vroege tweetaligheid van allochtone kinderen kan om om een aantal redenen worden gemotiveerd. In de eerste plaats is van belang zicht te krijgen op het tempo van eerste- en tweede-taalverwerving binnen uiteenlopende etnische groepen. De mate van behoefte om naast het Nederlands als tweede taal ook de eigen etnische groepstaal te blijven gebruiken blijkt in de diverse groepen sterk verschillend (vgl. Extra & Verhoeven, 1993). Daarnaast is er behoefte aan inzicht in de aard en mate van transfer van vaardigheden (vgl. Van Els, et al., 1985). Het leren van een tweede taal kan worden gezien als een herstructureringsproces waarbij eerder verworven declaratieve en procedurele kennis van de moedertaal als uitgangspunt kan worden genomen voor het opslaan van kennis omtrent de tweede taal (McLeod & McLaughlin, 1986). In welke mate transfer optreedt in het taalontwikkelingsproces van kinderen uit verschillende etnische groepen is vooralsnog onduidelijk. Ten slotte is het van belang inzicht te verkrijgen in de factoren die de mate van individuele variatie in de ontwikkeling van tweetaligheid verklaren. In sociolinguïstisch opzicht blijken de mate van taalcontact en de sociaal-culturele oriëntatie van kinderen relevante variabelen (McLaughlin, 1985; Vermeer,

1986; Boers, et al., 1987; Verhoeven, 1987). Belangrijke onderwijskundige variabelen blijken de mate van instructie in de eerste en de tweede taal en de vormgeving van het curriculum (Teunissen, 1986).

In het huidige onderzoek is de tweetalige ontwikkeling van een groot aantal allochtone kleuters van Turkse, Antilliaanse en Marokkaanse afkomst gevolgd gedurende de onderbouwperiode van het basisonderwijs in Nederland. Alvorens de opzet en resultaten van dit onderzoek te bespreken gaan we kort in op de theoretische achtergronden van het onderzoek naar tweetalige ontwikkeling van allochtone kleuters.

ACHTERGRONDEN VAN HET ONDERZOEK

Afhankelijkheden tussen eerste- en tweede-taalverwerving

Cummins (1979, 1980) vond empirische evidentie voor een meerdimensionale opvatting van taalvaardigheid. Hij vond dat de prestaties van leerlingen op een breed scala van taaltaken konden worden samengenomen in twee clusters: *cognitive/academic language proficiency* (CALP) en *basic interpersonal communicative skills* (BICS). Het eerste cluster verwijst naar algemene cognitieve of academische kennis en vaardigheden, zoals woordenschat en complexe syntax; in het tweede cluster gaat het om pragmatische vaardigheden die gerelateerd zijn aan taalgebruik in interactie.

In latere publicaties (Cummins, 1984; Cummins & Swain, 1986) wordt het onderscheid tussen CALP en BICS in een theoretisch raamwerk geïntegreerd. Daarin wordt uitgegaan van twee dimensies waarop taaltaken te schalen zouden zijn. De eerste dimensie heeft betrekking op de cognitieve inspanning die taalgebruik vereist. Daarbij wordt aangenomen dat naarmate iemand taalvaardiger is, het taalgebruik automatischer gaat verlopen en bijgevolg minder cognitieve inspanning zal vergen. De tweede dimensie betreft de mate waarin taalgebruik is ingebed in een concrete context. Daarbij wordt onderscheid gemaakt tussen aan de ene kant communicatie in een concrete situatie waarin naar kenmerken van die situatie wordt verwezen en aan de andere kant taalgebruik waarvoor uitsluitend een linguïstische context aanwezig is. In het laatste geval is alles in taal gecodeerd zonder dat er voor de taalgebruiker aanknopingspunten zijn in de onmiddellijke realiteit.

Met betrekking tot tweetaligheid en tweetalig onderwijs gaat Cummins ervan uit dat interpersoonlijke communicatieve taalvaardigheidsaspecten (BICS) min of meer gescheiden tot ontwikkeling komen in een eerste taal (T1) en tweede taal (T2). Cognitief/academische vaardigheidsaspecten zouden daarentegen over de twee talen een gemeenschappelijke basis hebben. Juist die gemeenschappelijke basis van cognitief/academische taalvaardigheid zou een positieve transfer van de ene taal naar de andere taal mogelijk maken, zoals verwoord in Cummins' afhankelijkheidshypothese (Cummins, 1983):

"To the extent that instruction in Lx is effective in promoting proficiency in Ly will occur, provided there is adequate exposure to Ly (either in school or in the environment) and adequate motivation to learn Ly".

Volgens Cummins zou transfer steeds optreden vanuit de dominante taal naar de gedomineerde taal. Met andere woorden, kennis en vaardigheden opgedaan vanuit de taal die kinderen het best beheersen zijn overdraagbaar naar de taal die kinderen in mindere mate beheersen. Evidentie voor het optreden van taaltransfer is gevonden in een groot aantal studies (vgl. Van Els, et al, 1985; Cummins & Swain, 1986; Verhoeven, 1987, 1991a/b, 1994). Transfer kan daarbij niet worden verklaard vanuit de onderliggende intelligentie van kinderen (vgl. Verhoeven, 1992). Volgens Cummins zou transfer vooral optreden bij niet-geautomatiseerde vaardigheden in een linguïstische context. Vandaar dat in het onderzoek de nadruk zal worden gelegd op de operationalisering van cognitief/academische taalkennis. Met name de beheersing van cognitieve begrippen die tot de "linguïstische universalia" worden gerekend lijken in dit opzicht van groot belang (Clark & Clark, 1977; Cummins, 1983).

Determinanten van tweetalige ontwikkeling

Eerder onderzoek naar mogelijke determinanten van de tweetalige ontwikkeling van kinderen van etnische minderheden heeft zich toegespitst op variabelen in de volgende drie domeinen: het kind, het gezin en de institutionele zorg (kindercentra, kleuteropvang) (vgl. Verhoeven, 1991a).

Kindvariabelen

In een aantal studies staat de culturele oriëntatie van het kind centraal. Culturele oriëntatie is te definiëren als de attituden die allochtone kinderen hebben ontwikkeld tegenover de eigen taal en cultuur en de dominante taal en cultuur. Deze attituden bepalen de mate van identificatie met de twee culturen (vgl. Lambert, 1978).

Onderzoek naar de invloed van kindvariabelen op tweetalige ontwikkeling heeft zich vooral gericht op de predictie van T2-leersucces. Onderzoek laat zien dat leeftijd een gevoelige factor is met betrekking tot tweede-taalverwerving (vgl. Larsen-Freeman & Long 1991). Hoewel oudere leerders sneller vooruitgang boeken, zijn jongere leerders uiteindelijk succesvoller. Alleen jonge kinderen zijn in staat een accent-vrije T2-vaardigheid op te bouwen. Voorts blijkt motivatie een rol van betekenis te spelen. Wong-Fillmore (1976, 1979) toont aan dat de integratieve motivatie van kinderen de T2-ontwikkeling positief beïnvloedde. Kinderen die zich vooral identificeerden met leeftijdgenoten die de T2 spraken boekten de meeste vooruitgang met T2-leren. Vermeer (1986) en Bosch, Broers & Verhoeven (1987) vinden dat culturele oriëntatie naast taalcontact belangrijke voorspellers van T2-ontwikkeling vormden. Echter, in de studies van Genesee & Hamayan (1980) en Appel (1986) bleek het verband tussen attitudinele variabelen en T2-vaardigheid gering.

Gezinsvariabelen

Onderzoek naar de invloed van gezinsvariabelen heeft zich vooral gericht op de effectiviteit van het gezin als leeromgeving. Wells (1985) vindt dat de mate waarin ouders erin slagen hun taalgebruik af te stemmen op het niveau van het kind van invloed is op de taalontwikkeling van kinderen. De mate van intersubjectiviteit van aandacht, de uitdrukking van begrijpelijke proposities, het toezicht op het succes van de communicatie en de stimulering van verdere interacties blijken daarbij de belangrijkste factoren.

Een andere belangrijke gezinsfactor vormt het taalaanbod. Fantini (1985) laat zien dat het taalcontact binnen het gezin van grote invloed is op de taalontwikkeling van allochtone kinderen. Tosi (1979, 1984) toont aan dat het aanbod van T1 over generaties een steeds belangrijkere factor wordt met het oog op verwerving van die taal door immigrantenkinderen. Binnen de eerste generatie immigranten is de taal ingebed in een rijke infrastructuur. In latere generaties wordt de eigen taal meestal alleen nog gebruikt in een beperkter aantal domeinen. Een aantal studies laat het belang zien van de mate van contact van gezinsleden met de T2 voor de ontwikkeling van die taal (vgl. Schumann, 1978).

Een andere belangrijke gezinsvariabele vormt de attituden van ouders tegenover T1-behoud en T2-leren. Gardner (1968) liet in de Canadese context zien dat de attituden van Engels-sprekende ouders tegenover het leren van Frans als tweede taal de motivatie van kinderen om Frans te leren beïnvloedt. Lalleman (1986) vindt eveneens een significant verband tussen de attituden van ouders en de T2-vaardigheid van jonge Turkse kinderen in Nederland.

Institutionele variabelen

Onderzoek laat zien dat institutionele opvang in kinderdagverblijven en in het kleuteronderwijs de taalontwikkeling van jonge kinderen in sterke mate beïnvloedt. Wong-Fillmore (1982) constateert dat binnen institutionele contexten de T2-verwerving van allochtone kinderen wordt bevorderd, doordat kinderen in staat worden gesteld te profiteren van frequente interactie met T2-sprekende leeftijdgenoten en van de positieve feedback van groepsleiders. In een groot aantal studies komt naar voren dat de aard en mate van interactiepatronen in institutionele

omgevingen uitermate belangrijk is voor de T2-ontwikkeling (vgl. McLaughlin, 1985: 145-163).

Voorts is de kwaliteit van T2-instructie op school een belangrijke variabele. Larsen-Freeman en Long (1991:312-321) rapporteren over een groot aantal studies over het effect van T2-instructie op het leren van een tweede taal.

Met betrekking tot de institutionele opvang van met name jonge allochtone kinderen wordt vaak een tweetalige opvang bepleit. Een groot aantal studies in Canada wijst erop dat kinderen binnen de context van een tweetalig programma een positief zelfbeeld ontwikkelen (vgl. Harley & Lapkin, 1984; Cummins & Swain, 1986). Er is echter sprake van aanzienlijke controverse ten aanzien van de effectiviteit van tweetalige programma's. Overigens blijken uit empirisch onderzoek de effecten van tweetalig onderwijs in de Verenigde Staten overwegend positief te zijn. Willig (1985) vatte de resultaten van 28 experimenten met tweetalig onderwijs samen en herleidde deze met behulp van meta-analyse tot een serie onderwijseffecten. Tweetalig onderwijs blijkt in vergelijking met uitsluitend T2-onderwijs een positief effect te hebben op vrijwel alle vaardigheden in de moedertaal en op de ontwikkeling van het zelfbeeld als lid van de betreffende etnische groep. Daarnaast blijkt tweetalig onderwijs de ontwikkeling van de lees- en rekenvaardigheid in de tweede taal te bevorderen en een positieve invloed te hebben op de ontwikkeling van het zelfbeeld als lid van een multiculturele samenleving.

Tweetaligheid en metalinguïstische ontwikkeling

Tweetaligheid heeft geen negatieve invloed op de ontwikkeling van algemene intellectuele vaardigheden van kinderen. Volgens Hakuta (1986) is er juist een positief effect van tweetaligheid op aan taal gerelateerde vaardigheden. Dat geldt met name voor het metalinguïstisch bewustzijn. Onder metalinguïstisch bewustzijn wordt impliciete kennis over taal begrepen die expliciet wordt gemaakt. Bij dit soort kennis verschuift de aandacht van de communicatieve inhoud van taal naar de grammatische vormgeving. Minimaal kunnen tot metalinguïstisch bewustzijn gedragingen worden gerekend die een zekere mate van abstractie van gebruikaspecten van taal veronderstellen; maximaal betreft het een expliciete formulering van linguïstische kennis. Het begrip metalinguïstisch bewustzijn wordt vooral in relatie gebracht met leerprocessen binnen het aanvankelijk leesonderwijs (vgl. Tunmer, et al., 1984; Downing & Valtin, 1984). Ten aanzien van de groei van cognitieve taalkennis bij allochtone kinderen is het van belang zicht te krijgen op het effect van de aard en mate van tweetaligheid van allochtone kinderen op de ontwikkeling van metalinguïstisch bewustzijn. Er zijn aanwijzingen dat het min of meer tegelijkertijd verwerven van twee talen bevorderlijk is voor de ontwikkeling van het metalinguïstisch bewustzijn (Cummins, 1983, 1984), voor het leren hanteren van cognitieve strategieën (Hakuta, 1986) en voor het vermogen om auditieve informatie te verwerken (Albert & Obler, 1978).

Onderzoek naar taalverwerving onder etnische minderheden in Nederland

Onderzoek naar de taalontwikkeling van allochtone kinderen in Nederland wordt gekenmerkt door een sterke groei. Het begrip 'taal' wordt in de meeste studies echter gereduceerd tot Nederlands (als tweede taal). In haar studie onderzoekt Van Helvert (1985) gedurende 9 maanden met een interval van een week de aanvankelijke T2-ontwikkeling van 5 zesjarige Turkse kinderen in interactie met Nederlandstalige leeftijdgenoten en met de onderzoekster zelf. Bij de data-analyse lag het accent op de verwerving van syntactische structuren en op die van de woordenschat. Tevens werd een aantal toetsen geconstrueerd en afgenomen, en werden kenmerken van het taalaanbod ten opzichte van de Turkse kinderen onderzocht. Primair is getracht inzicht te geven in de volgorde van taalverwerving en in daarbij optredende verschillen tussen de kinderen. Bij Vermeer (1986) staat de verwerving centraal van een reeks grammatische morfemen door 16 Turkse en 16 Marokkaanse kinderen die 2-7 jaar in Nederland woonden. Het tempo van T2-verwerving vertoonde grote verschillen, afhankelijk van patronen van taalkeuze en taalcontact van de kinderen.

Studies over de verwerving van etnische groepstalen door jonge allochtone kinderen zijn tot

nu toe beperkt in aantal. Verhoeven & Boeschoten (1986) vergelijken in een cross-sectionele studie de eerste-taalontwikkeling van Turkse kinderen in Turkije en Nederland. Zij vinden dat het tempo van het taalontwikkelingsproces van de kinderen in Nederland beduidend langzamer was dan dat van leeftijdgenoten in het land van herkomst.

Een longitudinale micro-analyse van de verwerving van het Turks door 12 in Nederland geboren Turkse kinderen in de leeftijd van 4-6 jaar is ondernomen door Boeschoten (1990). De data-verzameling vond plaats op 3 meetmomenten, bij een gemiddelde leeftijd van 4.5, 5.2 en 6.2 jaar. Bij de data-analyse lag het accent op de verwerving van het lexicon, nominale groepen, en temporele referentie. Hij constateerde dat het Turks van de kinderen in Nederland stagneerde in vergelijking met kinderen in Turkije.

Wagenaar (1990) doet verslag van een tweetalig onderwijsexperiment met Marokkaanse kleuters. In de eerste twee jaar van de basisschool worden er 15 uur in het Marokkaans-Arabisch onderwezen en 8 uur in het Nederlands. In het derde jaar worden deze tijden gewisseld. In groep 4 wordt dan tenslotte omgeschakeld naar volledig Nederlandstalig onderwijs. Wagenaar bestudeert ook de effecten van het experiment op de Berberkinderen in de groep. Na twee jaar van het experiment blijkt dat wat taalbegrip en taalproductie van het Marokkaans-Arabisch betreft de experimentele groep op een gelijk niveau blijft en de controlegroep achteruitgaat. De achterstand in het Nederlands blijft gelijk in beide groepen. De conclusie van Wagenaar is dan ook dat tweetalig onderwijs de ontwikkeling van het Nederlands niet verstoort.

In een onderzoek van Pels (1991) is het (gerapporteerd) taalgebruik onderzocht van 19 Marokkaanse kinderen in de periode voor zij naar de basisschool gaan. 35% spreekt bijna altijd Nederlands thuis, 40% gebruikt het Nederlands nu en dan en slechts 25% sprak thuis geen Nederlands. Pels merkt daarbij op dat de ouders van de kinderen van de eerste categorie naast de eigen taal het Nederlands gebruiken en dat deze kinderen bovendien tot de jongsten van het gezin behoren. Daarmee verklaart zij dan in zekere zin de gevonden patronen.

Samenvattend kan worden geconstateerd dat internationaal onderzoek naar de ontwikkeling van vroege tweetaligheid laat zien dat er transfer optreedt van de dominante taal naar de andere taal. Taalaanbod en culturele oriëntatie van kinderen en hun ouders blijken belangrijke factoren die de mate van individuele variatie in taalontwikkeling bepalen. Daarnaast blijken ook de kwaliteit en kwantiteit van het taalaanbod in een institutionele context belangrijke determinanten van taalontwikkeling. Voor zover het de Nederlandse context betreft heeft het meeste onderzoek naar de taalontwikkeling van jonge allochtone kinderen zich gericht op de verwerving van vaardigheden in het Nederlands (als tweede taal). Daarbij zijn sterke overeenkomsten tussen kinderen geconstateerd in de volgorde van tweede-taalverwerving en grote verschillen in het tempo van tweede-taalverwerving. Genoemde verschillen doen zich voor binnen en tussen groepen, waarbij vooral voor Turkse en Marokkaanse kinderen relatief lage taalvaardigheidsniveaus worden geconstateerd. Voorzover onderzoek is uitgevoerd naar tweetalige ontwikkeling, blijken allochtone kleuters beter te presteren in etnische groepstalen dan in het Nederlands. Van de aard en mate van tweetaligheid van allochtone kinderen bij instroom in het basisonderwijs bestaat echter geen duidelijk beeld. Ook is niet duidelijk wat de rol is van taaltransfer en welke determinanten vooral het tempo van tweetalige ontwikkeling bepalen. De kerndoelstelling van het onderhavige onderzoek is om daarover bij twee Mediterrane groepen (Turkse en Marokkaanse kinderen) en één ex-koloniale groep (Antilliaanse/Arubaanse kinderen) vergelijkend onderzoek te verrichten.

OPZET VAN HET ONDERZOEK

Vraagstellingen

In het huidige onderzoek stonden de volgende vragen centraal:

1. Hoe ontwikkelt zich de aard en mate van tweetaligheid, met nadruk op cognitieve taalkennis, bij Turkse, Marokkaanse en Antilliaanse kinderen in de eerste twee jaargroepen van het basisonderwijs?

2. In welke mate verschilt de ontwikkeling van het Nederlands bij allochtone kinderen met die van autochtone leeftijdgenoten?
3. In welke mate is binnen het proces van tweetalige ontwikkeling sprake van afhankelijkheden tussen eerste- en tweede-taalverwerving, zoals verondersteld in de door Cummins geformuleerde afhankelijkheidshypothese?
4. In welke mate bepalen sociolinguïstische variabelen, gerelateerd aan het taalaanbod thuis en de culturele oriëntatie van minderheidsgroepen, en onderwijskundige variabelen, gerelateerd aan de mate van taalaanbod in de (pre)school de individuele variatie van tweetalige ontwikkeling?
5. Wat is op het eind van groep 2 de relatie tussen het niveau van tweetaligheid en de mate van metalinguïstisch bewustzijn?

Informanten

Om de onderzoeksvragen te kunnen beantwoorden zijn in totaal 4 cohorten van informanten geselecteerd: twee allochtone cohorten (cohort 1 en 2) en twee autochtone cohorten (cohort 3 en 4) die als controlegroepen dienden. Bij de dataverzameling onder allochtone kinderen is uitgegaan van een longitudinaal design. In cohort 1 is gemeten aan het begin en het eind van groep 1 en aan het eind van groep 2; in cohort 2 aan het begin en aan het eind van groep 2. Er is gekozen voor opdeling van de allochtone groepen in twee cohorten, omdat rekening is gehouden met uitval van proefpersonen in de loop van de tijd. Bij de autochtone cohorten is sprake van twee cross-sectionele metingen in het Nederlands als moedertaal, aan het begin van groep 1 en aan het eind van groep 2.

De procedure van steekproeftrekking was als volgt. Aan een steekproef van 100 scholen verspreid over kleine en (middel)grote steden over het hele land werd verzocht om aan het onderzoek deel te nemen. Als selectiecriteria voor de allochtone onderzoekspopulatie dienden nationaliteit/geboorteland van het kind of de ouders, leeftijd van het kind en thuistaal. De scholen werden verzocht om uit hun schooladministratie kleuters te selecteren van Turkse, Marokkaanse en Antilliaanse/Arubaanse oorsprong die thuis respectievelijk Turks, Marokaans-Arabisch en Papiamentu spreken. Kinderen uit gemengde huwelijken alsmede Berbertalige Marokkaanse en Engelstalige Antilliaanse kinderen vielen af als informant. Uiteindelijk namen 74 scholen deel aan het onderzoek. Daarbij bleek sprake van een overrepresentatie van scholen uit de vier grote steden. In totaal namen 91 Turkse kinderen (waarvan 60 uit Cohort 1), 104 Antilliaanse/Arubaanse kinderen (waarvan 64 uit Cohort 1) en 111 Marokkaanse kinderen (waarvan 69 uit Cohort 1) deel aan het onderzoek. Daarnaast participeerden twee cohorten van 60 autochtone kleuters in het onderzoek. De autochtone kinderen, hadden een vergelijkbare sociaal-economische achtergrond, waren aselekt gekozen uit dezelfde scholen en zaten in dezelfde jaargroepen als de allochtone kinderen. Van een volledige matching van de achtergrondsituatie van allochtone en autochtone groepen is uiteraard geen sprake.

Instrumenten

Toetsing tweetaligheid

Voor het meten van de vaardigheid in de eerste en tweede taal van allochtone kinderen is gebruik gemaakt van een in een vooronderzoek (vgl. Verhoeven, et al., 1991) geconstrueerde Toets Tweetaligheid Kleuters (voortaan TTK). Het betreft taken voor klankmanipulatie, passieve en actieve woordenschat, cognitieve begrippenkennis, auditief geheugen van zinnen en tekstbegrip in de talen Turks, Marokkaans-Arabisch, Papiamentu en Nederlands. In het toetsinstrumentarium ligt het accent op die aspecten van tweetalige ontwikkeling die nauw samenhangen met de cognitieve ontwikkeling van kinderen (CALP). Bij het ontwikkelen van de toets is uitgegaan van een beschrijvingsmodel waarbij taalvaardigheid wordt opgevat als een hiërarchisch geordend systeem waarin fonologische, lexicale, syntactische en tekstuele vaardigheden samengaan (Verhoeven & Vermeer 1989; Verhoeven, 1992).

Uitgaande van een hiërarchische structuur van taalvaardigheid, zijn deeltaken ontwikkeld

Tabel 1. Deeltaken tweetalige toets voor kleuters.

Linguistisch niveau	Deeltaken	Aantal items
Fonologisch	Klankmanipulatie	25
Lexicaal	1. Inhoudswoorden passief	60
	2. Inhoudswoorden actief	40
	3. Cognitieve begrippen	65
Syntactisch	Zinsimulatie	40
Tekstueel	Tekstbegrip	20
Totaal aantal items		250

waarmee taalvaardigheid in bovengenoemde domeinen kan worden getoetst. Tabel 1 geeft een overzicht van de deeltaken van de volledige toets per linguïstisch niveau, inclusief het aantal items waaruit de taak bestaat.

Op fonologisch niveau zijn taken ontwikkeld waarin kinderen moeten kunnen laten zien dat ze in hun eerste en tweede taal klanken kunnen onderscheiden en produceren. Elke taak bestaat uit 25 opgaven. Bij elke opgave krijgen kinderen twee woorden aangeboden die qua klankopbouw minimaal van elkaar verschillen. Zo kunnen woorden verschillen in één klank (bijvoorbeeld: *zoon* - *zoen*), in de positie van klanken (bijvoorbeeld: *kurk* - *kruk*), of in aantal klanken (*kaars* - *kaas*). Na de aanbieding van een woordpaar wordt van kinderen gevraagd de woorden achter elkaar te herhalen. Alleen als beide woorden door het kind correct zijn nagesproken wordt de opgave goed gerekend.

Om de kennis van inhoudswoorden te meten zijn passieve- en actieve-woordenshattaken ontwikkeld. Deze taken gaan in de vier talen uit van dezelfde concepten. Met de passieve-woordenshattaken wordt de receptieve kennis van inhoudswoorden gemeten. De taak bestaat uit 60 opgaven waarbij steeds uit vier aangeboden plaatjes één moet worden gekozen dat de betekenis van een aangeboden woord aangeeft. De actieve-woordenshattaak bestaat uit 40 opgaven waarbij kinderen plaatjes moeten benoemen.

In de cognitieve-begrippentaak wordt de kennis van functiewoorden getoetst. Omdat functiewoorden naast inhoudswoorden belangrijke betekenisvolle elementen van taal zijn waarvan de kennis in nauwe relatie staat tot de denkontwikkeling van kinderen krijgt dit onderdeel veel aandacht. In de cognitieve-begrippentaak zijn aparte deeltaken onderscheiden waarin begrippen worden getoetst die verwijzen naar kleur (voorbeelden: *blauw*, *licht groen*), vorm (voorbeelden: *cirkel*, *even groot*), hoeveelheid (voorbeelden: *alle*, *middelste*), ruimte (voorbeelden: *boven*, *tussen*) en relaties tussen gebeurtenissen (voorbeelden: *terwijl*, *nadat*). In alle vier de talen worden weer dezelfde begrippen getoetst. Bij elke opgave krijgt het kind enkele plaatjes te zien waarbij van het kind gevraagd wordt aan te geven welk plaatje bij een bepaald begrip hoort. Om inzicht te krijgen in de kennis van de zinsbouw van kinderen is per taal een zinsvormingstaak ontwikkeld. Elke taak bestaat uit 20 zinnen die kinderen moeten naspreken. Van elke nagesproken worden twee aspecten gescoord: een functiewoord en een grammaticaal patroon. Zo wordt van de zin *deze groene tas is groter dan die van jouw vader* het functiewoord *groter* en het patroon *deze groene tas* gescoord.

Mondeling tekstbegrip wordt getoetst door kinderen in elke taal vier korte verhaaltjes voor te lezen. Na elk voorgelezen verhaal krijgen kinderen vijf vragen voorgelegd die betrekking hebben op betekenisrelaties en verwijsrelaties die erin voorkomen.

Bij de uitwerking van deze taken zijn enkele belangrijke condities in acht genomen. In de eerste plaats is het taalgebruik binnen de instructies zo eenvoudig mogelijk gehouden, zodat zwakke toetsresultaten niet in termen van misverstaan van de instructie kunnen worden uitgelegd. Daar-

naast zijn cultuurspecifieke contexten in de taken zoveel mogelijk vermeden. Een complicerende factor daarbij is dat de componenten van de toetsen in alle vier betrokken talen vergelijkbaar moeten zijn. Dit betekent dat bij de samenstelling van de taken rekening moest worden gehouden met vier culturen. Hoewel culturele bias in gestandaardiseerde toetsen bijna onvermijdelijk is, is zoveel mogelijk gepoogd de verdisconteerde betekenisaspecten in de gekozen taalbouwsels en tekeningen voor kinderen met uiteenlopende culturele achtergronden een vergelijkbare herkenbaarheid te geven. Tenslotte zijn de taken zo geconstrueerd dat zij niet alleen aangeven op welk niveau binnen het eerste- en tweede-taalverwervingsproces allochtone kinderen zich bevinden, maar eveneens enig zicht geven op de mate van taaldominantie, oftewel de relatieve sterkte in beide talen, in de diverse domeinen. Hierbij moet worden opgemerkt dat dominantie in één van de getoetste domeinen nog geen taaldominantie hoeft te impliceren.

In een vooronderzoek is de psychometrische kwaliteit van de toetstaken onderzocht. Voor een gedetailleerde uitwerking van de betrouwbaarheid en validiteit van de ontwikkelde toetsen wordt verwezen naar Verhoeven, Extra, Konak, Narain en Zerrouk (1991). Met het oog op interne consistentie zijn de item-totaalscores berekend en is Cronbach's alpha bepaald. De alpha-coëfficiënt bereikte waarden variërend van .83 tot .95, terwijl de item-totaalscores vrijwel geen lage waarden kenden. Met het oog op inhoudsvaliditeit is de homogeniteit van toetstaken onderzocht. Daartoe is per toetstaak nagegaan in hoeverre het algemene eenparametermodel van

Tabel 2. Cronbach's alfa, gemiddelde Rit en toetsingsgrootheid eenparameter itemresponsmodel per groep en per toets.

	alpha	Rit	X ²	df	p
TURKS					
Klankdiscriminatie	.90	.56	25.7	24	.40
Cognitieve begrippen	.90	.38	44.94	64	.92
Passieve woordenschat	.90	.47	36.35	39	.63
Aktieve woordenschat	.85	.39	9.63	37	.99
Zinnenimitatie	.95	.58	22.88	39	.99
Tekstbegrip	.89	.57	15.37	19	.70
PAPIAMENTU					
Klankdiscriminatie	.92	.63	21.53	24	.61
Cognitieve begrippen	.92	.42	54.86	64	.62
Passieve woordenschat	.91	.48	33.04	38	.75
Aktieve woordenschat	.92	.57	27.73	26	.40
Zinnenimitatie	.96	.62	29.44	39	.90
Tekstbegrip	.93	.64	5.66	19	.99
MAROKKAANS					
Klankdiscriminatie	.87	.50	20.08	24	.69
Cognitieve begrippen	.83	.30	80.85	64	.06
Passieve woordenschat	.90	.47	37.25	39	.62
Aktieve woordenschat	.89	.53	23.18	22	.41
Zinnenimitatie	.94	.52	31.67	39	.82
Tekstbegrip	.87	.53	15.28	19	.70
NEDERLANDS					
Klankdiscriminatie	.87	.50	24.84	24	.42
Cognitieve begrippen	.89	.36	75.43	64	.11
Passieve woordenschat	.95	.52	67.39	56	.07
Aktieve woordenschat	.90	.63	9.04	15	.89
Zinnenimitatie	.95	.59	23.44	39	.97
Tekstbegrip	.91	.60	19.16	19	.60

Verhelst & Eggen (1989) op de scores paste. In geen enkel geval hoefde de aanname van het model te worden verworpen. In Tabel 2 staan Cronbach's alpha, gemiddelde Rit en toetsingsgrootte van het eenparameter itemresponsmodel per groep en per toets.

Toetsen voor metalinguïstisch bewustzijn en leesvaardigheid

Om het metalinguïstisch bewustzijn van de kinderen te meten is gebruik gemaakt van instrumentarium dat in het kader van het Cito-leerlingvolgsysteem is ontwikkeld (Van Kuyk & Verhoeven, 1991). Uit dit toetsinstrumentarium zijn vijf onderdelen geselecteerd waarbij de analyse in fonemen en woorden en de synthese van klanken aan bod komen, alsmede rijm en objectivatie.

De Toets voor Analyse in Fonemen wil nagaan hoe goed kinderen woorden kunnen uiteenleggen in hun samenstellende fonemen. Het kind moet een aangeboden woord mondeling analyseren in klanken. De taak bestaat uit 20 items die qua orthografische complexiteit geleidelijk aan oplopen: v-c (vocaal-consonant), cv, cvc, ccvc, cvcc, ccvcc, cccvc en cvccc. De taak werd afgebroken bij 5 opeenvolgende fouten. De score werd bepaald door het totaal aantal goed geanalyseerde woorden op te tellen.

De Toets voor Analyse in Woorden is bedoeld om na te gaan of kinderen de woorden in een zin van elkaar kunnen onderscheiden. Door met een potlood te tikken kan het kind aangeven hoeveel woorden een mondeling aangeboden zin bevat. Dit toetsonderdeel bestaat uit 10 zinnen met oplopende lengte. De score wordt bepaald door het aantal correct geanalyseerde zinnen.

De toets voor Synthese van Woorden beoogt na te gaan hoe goed kinderen uit afzonderlijke klanken woorden kunnen samenstellen. Aan het kind wordt door de proefleider een woord in zijn samenstellende elementen aangeboden, waarop het kind moet aangeven om welk woord het gaat. De toets bestaat uit 20 items die net als de fonemische analysetaak oplopen qua orthografische complexiteit. Er is wederom een afbreeknorm van 5 gehanteerd.

De Rijmtoets is bedoeld om na te gaan of het kind op een aangeboden woord een rijmwoord kan vinden. Het rijmwoord kan worden gevonden aan de hand van een aantal aangeboden plaatjes. In deze taak worden 10 woorden auditief aangeboden waarbij van het kind gevraagd wordt de referent van het juiste rijmwoord terug te zoeken in een serie plaatjes.

De Objectivatietoets heeft tot doel na te gaan of een kind uitsluitend kan letten op de lengte van een woord en daarbij kan afzien van de betekenis. Bij mondelinge aanbieding van twee woorden, moet het kind steeds aangeven wat het langste woord is. Een voorbeeld: *paddestoel/boom*. De toets bestaat uit 10 woordparen.

Met behulp van de Grafemetoets (Verhoeven, 1992) wordt de letterkennis van kinderen in het Nederlands bepaald. De toets gaat uit van een leeskaart waarop alle Nederlandse grafemen geïsoleerd in rijtjes onder elkaar worden aangeboden. De score bestaat uit het aantal correct verklankte grafemen.

Achtergrondgegevens

Allereerst zijn enkele sociobiografische gegevens van de kinderen bepaald, zoals de leeftijd, de verblijfsduur van vader, moeder en het kind zelf en de sociaal-economische status (gemeten aan de hand van de vooropleiding van de ouders).

Vervolgens zijn instrumenten ontwikkeld om de taalkeuze en het taalcontact van het kind vast te stellen. In een direct interview met het kind werden vragen gesteld omtrent de taalkeuze in interactie met vader, moeder, jongere broer/zus, oudere broer/zus, boezemvriend en overige vrienden. Daarbij kon steeds een keuze gemaakt worden tussen punten op een vijf-puntsschaal lopend van alleen eigen taal (1) tot alleen Nederlands (5). Op basis van de uitkomsten op deze schalen zijn vervolgens twee scores berekend: Taalkeuze (T1) en Taalkeuze (T2). Voor berekening van de eerstgenoemde variabele werden de schaa scores van 1-5 omgezet, zodat een hoge score overeenkomt met een relatief sterke keuze voor de eigen taal. Als maat voor taalcontact zijn vragen gesteld aan de leerkracht omtrent contact binnen en buiten schooltijd, de nationaliteit van de boezemvriend en het taalgebruik binnen en buiten schooltijd. Aan de hand van de vragenlijst konden aparte scores worden berekend met betrekking tot contact met de eigen taal

en contact met het Nederlands. De interne consistentie van genoemde instrumenten bleek redelijk (Cronbach's α bleek in alle gevallen groter dan .70).

Voorts zijn procedures gehanteerd om de culturele oriëntatie van de kinderen en hun ouders en de factor gezinstaal te bepalen (zie bijlage 2). Deze procedures zijn ontleend aan Verhoeven (1987). Als maat voor de culturele oriëntatie van het kind werd de oriëntatie tegenover het dagelijks leven in het land van herkomst en in Nederland onafhankelijk vastgesteld door middel van een gestructureerd interview met het kind, waarin de houdingen van kinderen tegenover 10 aspecten van het culturele leven in het land van herkomst en Nederland werden bepaald: school, vrienden op school, eten, leraar, voorlezen, televisie, muziek, vrienden thuis, taal en wonen. De oriëntaties van de kinderen werden gescoord volgens een vier-punt-schaal, variërend van extreem positief tot extreem negatief, gebaseerd op de antwoorden van de kinderen (bijvoorbeeld: "Zou je graag in Turkije op school zitten?", naast de vraag "Vind je het leuk om in Nederland op school te zitten"). De inter-rater betrouwbaarheid van deze schaal bleek hoog ($>.95$).

Voorts werden in een semi-gestructureerd interview met de ouders van de kinderen het taalcontact in het gezin (de factor 'gezinstaal') en de culturele oriëntatie van de ouders bepaald. Vragen omtrent het taalcontact betroffen het gebruik van de eigen taal en het Nederlands in gezinsinteractie, het contact met T1- en T2-sprekende personen in de thuisomgeving en het gebruik van media in de etnische groepstaal en het Nederlands (televisie, radio, krant). Vragen omtrent de culturele oriëntatie van de ouders hadden betrekking op de viering van eigen en Nederlandse feestdagen en ideeën omtrent OET- en Nederlandstalig onderwijs. In alle gevallen zijn op basis van sommering scores verkregen waarbij zowel voor de eigen taal/cultuur als voor de Nederlandse taal/cultuur geldt dat een hoge score overeenkomt met een grote mate en een lage score met een geringe mate van aanbod, c.q. oriëntatie. Voor een bespreking van de psychometrische deugdelijkheid van beide laatstgenoemde vragenlijsten wordt verwezen naar Verhoeven (1987:205-206).

Als eerste maat voor institutionele opvang is het aantal maanden bepaald dat kinderen voorschoolse opvang hebben genoten (preschool). Verder is het aantal uren OET binnen en buiten schooltijd bepaald. Eveneens is het aantal uren vastgesteld dat kinderen extra lessen Nederlands als tweede taal hebben genoten. Ten slotte is door middel van documentenanalyse onderzocht welke specifieke methodes binnen de kleuterschool worden gehanteerd ten behoeve van het tweede-taalonderwijs aan allochtone leerlingen. Daarbij is als maat gehanteerd: 0 (geen methode), 1 (gebruik één methode) of 2 (gebruik meerdere methodes).

Alle achtergrondgegevens zijn op het einde van de kleuterperiode verzameld. Tabel 3 geeft een overzicht van de verzamelde achtergrondgegevens en hun operationalisering.

Procedure

Alle taken werden afgenomen door moedertaalsprekers van de betreffende talen. De allochtone kinderen kregen eerst alle toetstaken in de eigen taal voorgelegd. Na een interval van twee weken werden bij dezelfde kinderen de Nederlandstalige toetsen afgenomen. Gegevens over de sociolinguïstische achtergrond van de kinderen werden verzameld met behulp van voorgestructureerde interviews met de kinderen, hun ouders en hun leerkrachten. Kinderen en ouders werden in de eigen taal aangesproken terwijl de leerkrachten vragenlijsten kregen voorgelegd die zij zelf invulden.

Met behulp van variantie-analyse is bepaald in welke mate de resultaten van allochtone kinderen (in zowel Cohort 1 als Cohort 2) op de onderscheiden taaltaken differentiëren van die van Nederlandse kinderen. Op vergelijkbare wijze is nagegaan in hoeverre allochtone kinderen in de beide cohorten beter (dominant) zijn in een van beide talen en in welke mate er in de loop van de tijd vooruitgang te zien is in de twee talen.

Om de mate van taaltransfer te onderzoeken is LISREL toegepast bij de kinderen uit het eerste cohort. Gegeven de beperkte omvang van de allochtone groepen is het meetmodel vooraf gespecificeerd. Daartoe zijn met behulp van factoranalyse de scores op de taaltoetsen gereduceerd tot een beperkter aantal factoren. Daarbij was de verwachting dat op elk meetmoment per taal één onderliggende factor zou worden teruggevonden. Vervolgens is met behulp van LIS-

Tabel 3. Achtergrondvariabelen en hun operationalisering.

Variabele	Operationalisering
<i>Kindkenmerken</i>	
Taalcontact kind	Interview leerkracht over T1/T2-gebruik binnen/buiten school
Taalkeuze kind	Interview kind over taalkeuze in en buiten het gezin
Culturele oriëntatie	Interview kind over attitudes t.o.v. eigen/Nederlandse cultuur
Verblijfsduur	Verblijf in Nederland in jaren
<i>Gezinskenmerken</i>	
Gezinstaal	Interview ouders over T1/T2-gebruik in het gezin
Culturele oriëntatie ouders	Interview ouders over attitudes t.o.v. eigen/Nederlandse cultuur
Verblijfsduur vader/moeder	Verblijf in Nederland in jaren
SES	Vooropleiding ouders
<i>Onderwijskenmerken</i>	
Preschool	Aantal jaren
OET binnen/buiten de klas	Aantal uren
Extra T2-onderwijs	Aantal uren
T2-methodes	Aantal methodes

REL-analyse nagegaan in hoeverre bij tweetalige ontwikkeling gesproken kan worden van afhankelijkheden tussen eerste- en tweede-taalverwerving. Daarbij is uitgegaan van een structureel model waarin causale relaties waren toegestaan in de tijd en vanuit de dominante (sterkste) taal naar de gedomineerde (zwakkere) taal.

Voorts zijn de factorscores voor taalvaardigheid van de kinderen uit beide cohorten op het eind van groep 2 gerelateerd aan achtergrondkenmerken met behulp van correlatie- en stepwise multiple-regressie-analyse. De beperkte steekproefomvang in het eerste cohort liet niet toe om voor dit onderdeel waarbij een groot aantal variabelen in het geding is, eveneens gebruik te maken van LISREL.

Om de relatie tussen tweetaligheid en metalinguïstisch bewustzijn te onderzoeken zijn per etnische groep correlaties berekend tussen enerzijds eerste- en tweede-taalvaardigheid en anderzijds aspecten van metalinguïstisch bewustzijn. Bovendien is met behulp van variantie-analyse nagegaan in hoeverre een gebalanceerd niveau van tweetaligheid samengaat met een hoge mate van metalinguïstisch bewustzijn. Daartoe is per etnische groep een driedeling gemaakt tussen kinderen die boven gemiddeld scoren in beide talen (vaardig in twee talen), kinderen die boven gemiddeld scoren in slechts één taal (vaardig in één taal) en kinderen die beneden gemiddeld scoren in beide talen (vaardig in geen van beide talen). Vervolgens is nagegaan in hoeverre er verschillen zijn in groeps-gemiddelden op de afgenomen metalinguïstische taken op het eind van groep 2.

RESULTATEN

Profielen van tweetalige ontwikkeling

Turken

Figuur 1 geeft de gemiddelde scores van de Turkse kleuters (cohorten 1 en 2) op de toetstaken in de eigen taal en in het Nederlands.

Er valt af te lezen dat de scores van de Turkse kinderen op de 6 toetsonderdelen in beide talen toenemen, naarmate de kinderen ouder worden. Variantie-analyse laat zien dat voor elke afzonderlijke taak de hoofdeffecten Taal en Tijd significant zijn (in alle gevallen $p < .001$). We kunnen

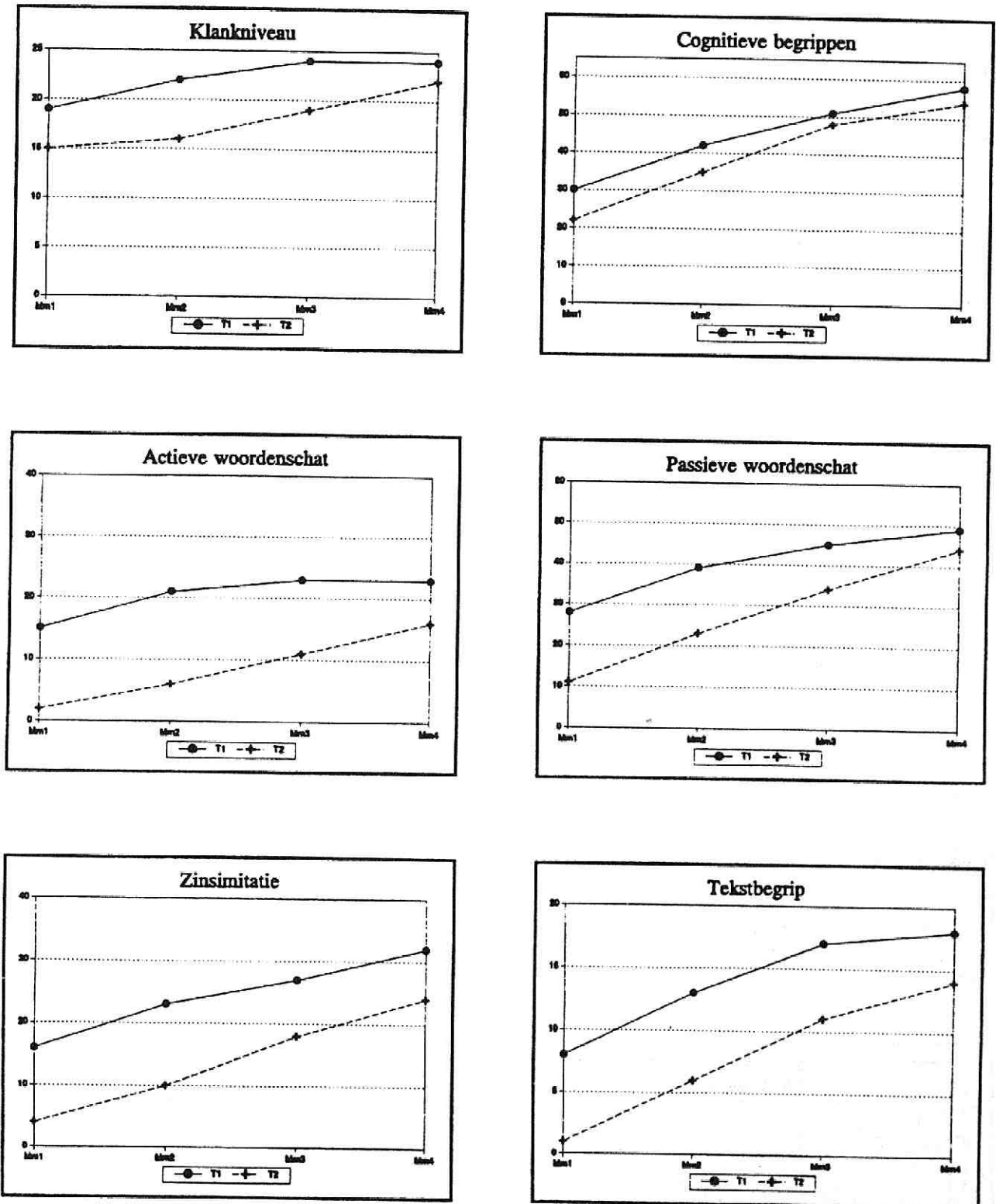


Fig. 1. Gemiddelde scores van de Turkse kinderen op de toetstaken in het Turks en het Nederlands bij de aanvang en op het eind van groep 1 en op het eind van groep 2.

dus concluderen dat de Turkse kinderen dominant Turkstalig zijn en blijven in de leeftijdperiode van 4-7 jaar. In beide talen is sprake van een significante vooruitgang in de loop van de tijd. Voor de onderdelen Klankmanipulatie, Cognitieve Begrippen en Passieve woordenschat bleek de Taal x Tijd interactie eveneens significant, hetgeen erop duidt dat de verschillen tussen de

vaardigheid in de eigen taal en die in het Nederlands in de loop van de tijd kleiner worden.

Antillianen/Arubanen

Figuur 2 geeft een overzicht van de gemiddelde toetsscores van de Antilliaanse kinderen (cohorten 1 en 2) per taak, taal en meetmoment.

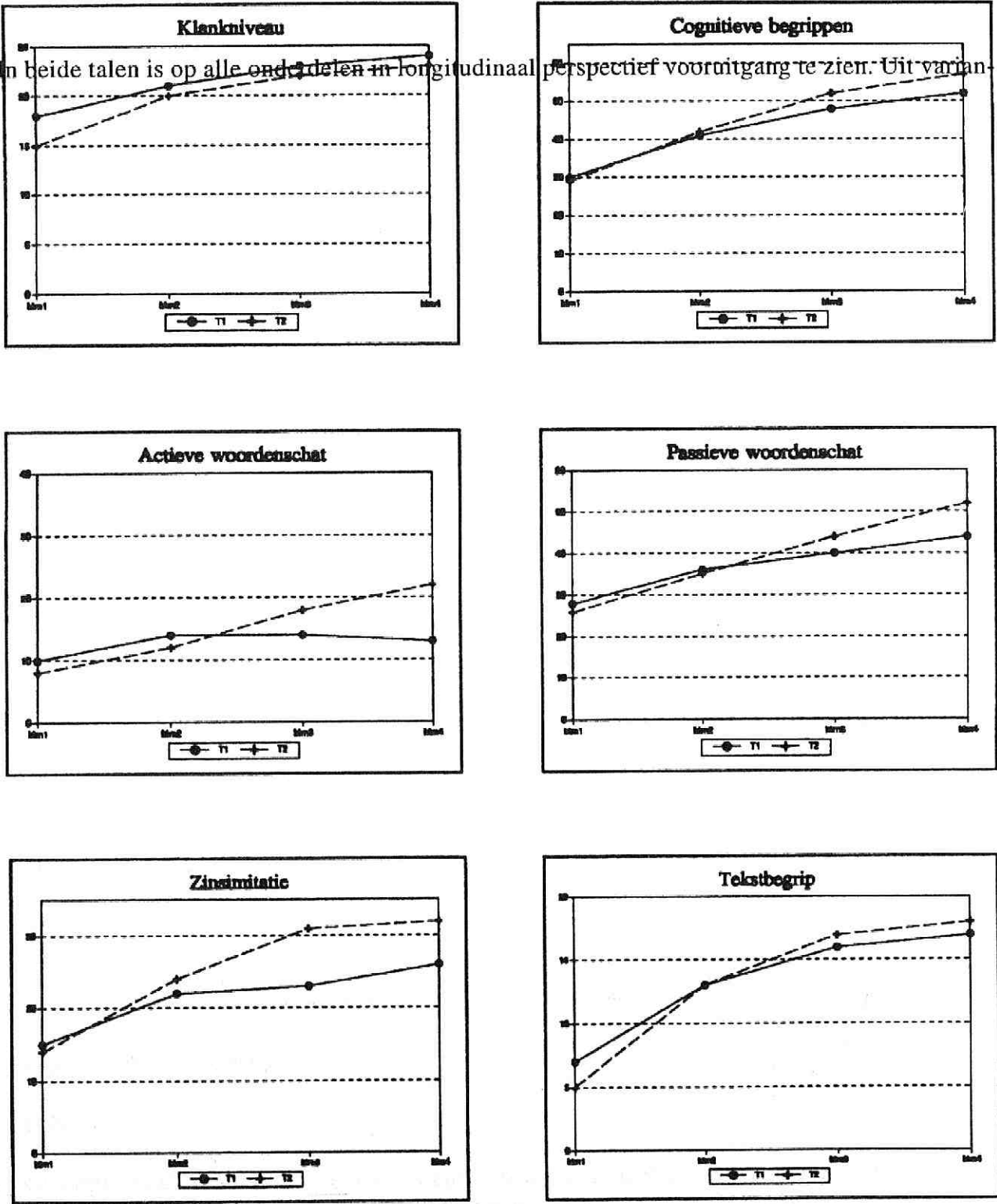


Fig. 2. Gemiddelde scores van de Antilliaanse/Arubaanse kinderen op de toetstaken in het Turks en het Nederlands bij de aanvang en op het einde van groep 1 en op het einde van groep 2.

tie-analyse blijkt dat het hoofdeffect Tijd in alle gevallen significant is ($p < .001$). De verschillen in scores tussen talen blijken gering. Slechts voor het onderdeel Klankmanipulatie blijkt het hoofdeffect Taal significant ($p < .05$). Het niveau van tweetaligheid is met andere woorden bij Antilliaanse/Arubaanse kinderen min of meer gebalanceerd te noemen. Verder valt af te lezen

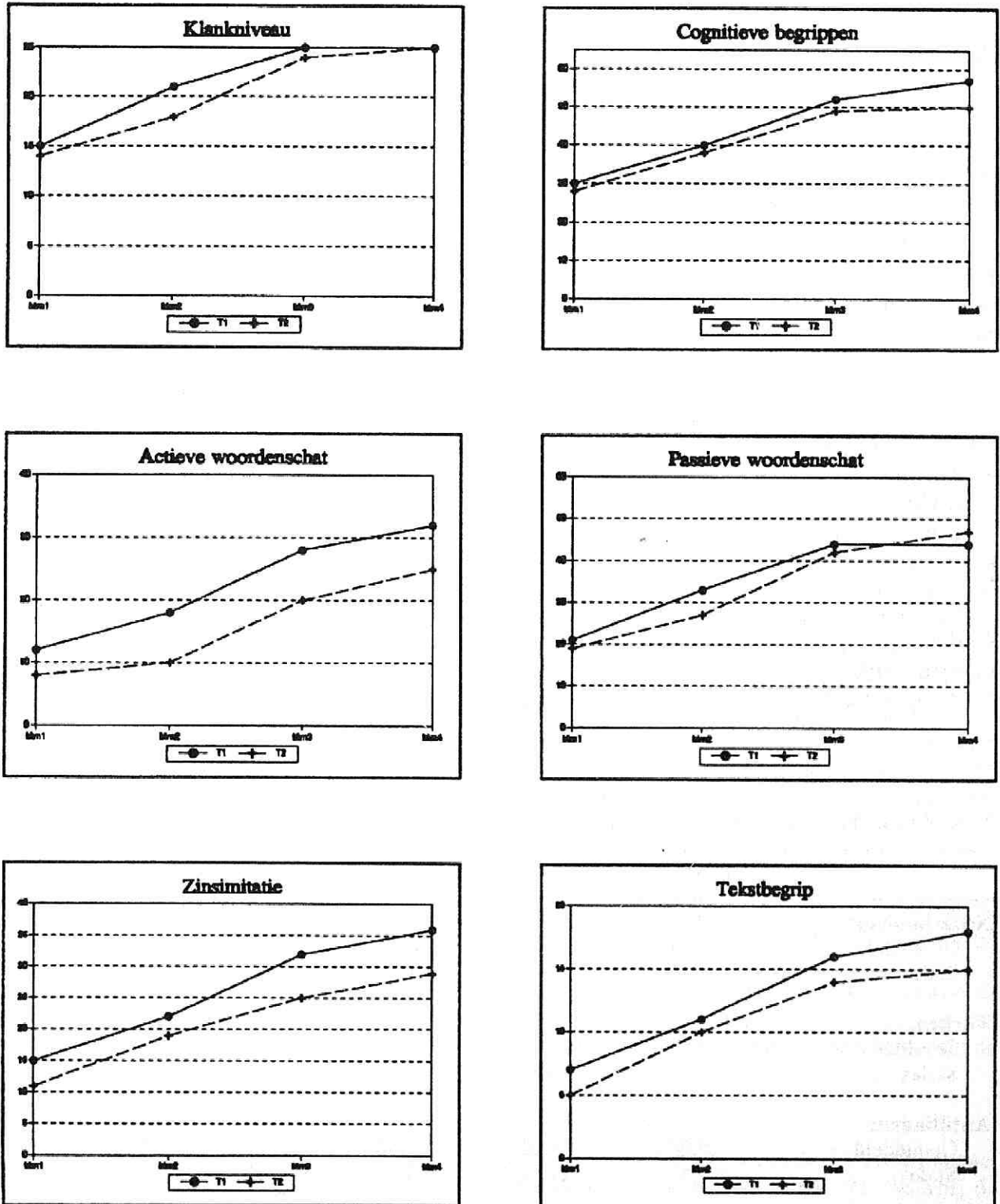


Fig. 3. Gemiddelde scores van de Marokkaanse kinderen op de toetstaken in het Turks en het Nederlands bij de aanvang en op het eind van groep 1 en op het eind van groep 2.

dat met uitzondering van Klankmanipulatie op alle taken een tendentie tot dominantie in de eigen taal plaatsmaakt voor Nederlandstalige dominantie. De Taal x Tijd interactie bleek voor alle taken, behalve Passieve woordenschat significant.

Marokkanen

Figuur 3 geeft een overzicht van de gemiddelde toetsscores van de Marokkaanse kinderen (cohorten 1 en 2) per taak, taal en meetmoment.

Op het onderdeel Klankmanipulatie bereiken de kinderen al op het derde meetmoment een plafond in beide talen. Op alle onderdelen blijken de verschillen tussen de toetsresultaten op de diverse meetmomenten significant te zijn ($p < .001$). Ook de verschillen tussen de T1 en de T2 van de Marokkaanse kinderen blijken voor alle taken significant (voor Klankmanipulatie en Passieve Woordenschat $p < .05$; in de overige gevallen $p < .001$). Op de meeste onderdelen is er een significant sterkere groei waar te nemen in het Nederlands als tweede taal. Behalve voor Klankmanipulatie blijkt de Tijd x Taal interactie steeds significant ($p < .001$).

Vergelijking van de resultaten NT2 en NT1

Tabel 4 geeft de gemiddelden en standaarddeviaties van de Turkse, Antilliaanse en Marokkaanse kinderen (cohorten 1 en 2) en Nederlandse kinderen (cohort 1) op de zes toetstaken in het Nederlands op het eerste meetmoment (begin groep 1).

De scores van de autochtone kinderen in het Nederlands als moedertaal blijken op alle onderdelen beduidend hoger dan de scores van allochtone kinderen in het Nederlands als tweede taal. De verschillen tussen de allochtone groepen enerzijds en de autochtone groep anderzijds blijken voor alle taken significant ($p < .001$). Binnen de allochtone groepen blijken de scores van de Antilliaanse kinderen relatief hoog, en die van de Turkse kinderen relatief laag, terwijl de Marokkaanse kinderen een middenpositie innemen.

Tabel 5 geeft de gemiddelden en standaarddeviaties van de Turkse, Antilliaanse en Marokkaanse kinderen (cohorten 1 en 2) en Nederlandse kinderen (cohort 2) op de zes toetstaken in het Nederlands op het derde meetmoment (eind groep 2).

Op dit meetmoment scoren de Nederlandse kinderen op alle onderdelen, uitgezonderd Klankmanipulatie significant hoger dan de allochtone kinderen ($p < .001$). Net als bij het eerste meetmoment blijken de scores van de Turkse kinderen het laagst en die van de Antilliaanse kinderen het hoogst. De Marokkaanse kinderen nemen wederom een tussenpositie in.

Tabel 4. Gemiddelden en standaarddeviaties op de Nederlandse toetsen aan het begin van Groep 1.

	Klankman.	Cogn.begr.	Akt.ws.	Pas. ws.	Zinsim.	Tekst
Nederlanders:						
Gemiddeld	19.26	40.00	16.20	35.85	23.95	12.83
St.dev.	3.55	7.14	4.04	9.54	8.78	3.74
Turken:						
Gemiddeld	14.55	22.03	2.08	10.81	4.32	1.37
St.dev.	6.98	10.34	2.23	8.23	5.86	2.28
Antillianen:						
Gemiddeld	15.04	29.01	7.82	25.92	13.74	5.20
St.dev.	5.51	11.92	4.37	11.18	11.34	4.24
Marokkanen:						
Gemiddeld	14.07	28.52	7.50	18.64	11.00	5.32
St.dev.	6.53	10.82	3.87	12.23	10.32	4.06

Tabel 5. Gemiddelden en standaarddeviaties op de Nederlandse toetsen aan het eind van Groep 2.

	Klankman.	Cogn.begr.	Akt.ws.	Pas. ws.	Zinsim.	Tekst
Nederlanders:						
Gemiddeld	22.20	52.93	25.24	50.98	32.76	16.93
St.dev.	2.51	5.98	5.84	5.65	4.60	2.67
Turken:						
Gemiddeld	19.15	47.61	10.56	34.08	18.24	10.76
St.dev.	3.39	6.36	4.33	8.19	7.88	4.39
Antillianen :						
Gemiddeld	22.34	52.32	17.92	44.48	31.00	16.75
St.dev.	2.42	6.64	5.86	6.04	5.71	3.34
Marokkanen:						
Gemiddeld	23.79	49.33	20.12	42.35	25.40	14.53
St.dev.	2.10	5.41	6.79	8.19	4.71	2.88

Rol van afhankelijkheden tussen eerste- en tweede-taalverwerving

Met behulp van principale-componenten-analyse is gezocht naar de centrale tendenties van eerste- en tweede-taalvaardigheid. Gegeven het feit dat taalvaardigheid hier is geoperationali-seerd in termen van grammaticale competentie, en pragmatische aspecten van taalgebruik niet zijn verdisconteerd, was de verwachting dat in alle gevallen van een ééndimensionaal construct sprake zou zijn. Dit bleek inderdaad het geval.

Vervolgens is nagegaan of, en in hoeverre, er bij ontwikkeling van tweetaligheid sprake is van afhankelijkheden tussen eerste- en tweede-taalverwerving. Daarbij wordt empirische evi-dentie gezocht voor Cummins' afhankelijkheidshypothese. Voor de drie etnische groepen zal deze hypothese worden getoetst voor de kinderen uit Cohort 1. Toetsing van de hypothese vindt plaats door de passendheid van het LISREL-model na te gaan. Bij toepassing van het LISREL-model is uitgegaan van de factorscores van taalvaardigheid. Binnen de structurele component van het model is nagegaan in hoeverre naast longitudinale verbanden binnen eerste- en tweede-taalvaardigheden over tijd eveneens sprake is van (cross-sectionele) causale relaties tussen eerste en tweede-taalvaardigheden. Daarbij is steeds gezocht naar het best passend structurele model.

Turken

In Figuur 4 staat het gestandaardiseerde structurele LISREL-model beschreven voor de 60 Turkse kinderen uit het eerste cohort. De goodness-of-fit van het beschreven model bedraagt .94. X²-toetsing gaf aan dat het model niet verworpen hoeft te worden ($p=.32$).

Het model laat zien dat sprake is van sterke longitudinale trends binnen eerste- en tweede-taalverwerving. Met name vanaf het eerste meetmoment is er daarnaast evidentie voor een afhankelijkheidsrelatie van T1 naar T2. De beta-coëfficiënten worden kleiner bij de overgang van het eerste naar het tweede meetmoment.

Antillianen/Arubanen

Figuur 5 presenteert de gestandaardiseerde LISREL-modeloplossing voor de 64 Antilliaanse/Arubaanse kinderen uit het eerste cohort. De gevonden goodness-of-fit bleek .91, waarbij de gevonden modelaannname niet verworpen hoefde te worden ($p=.17$).

Het model postuleert sterk longitudinale verbanden binnen eerste- en tweede-taalverwer-ving. Daarnaast blijkt enige evidentie voor transfer van taalvaardigheden. Op meetmoment 1 is

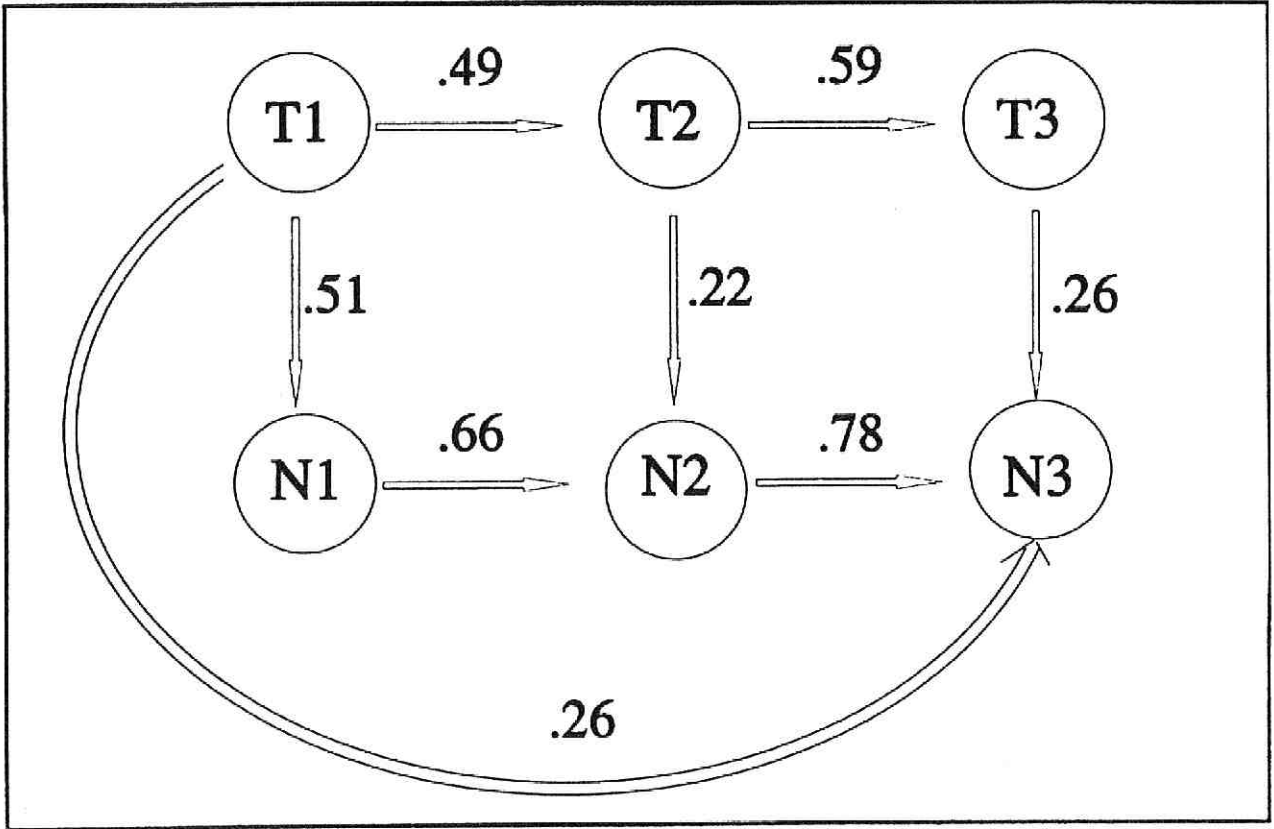


Fig. 4. Gestandaardiseerde oplossing voor het structurele LISREL-model dat de relaties specificceert tussen Turkse (T) en Nederlandse (N) taalverwerving bij de Turkse kinderen op de drie meetmomenten.

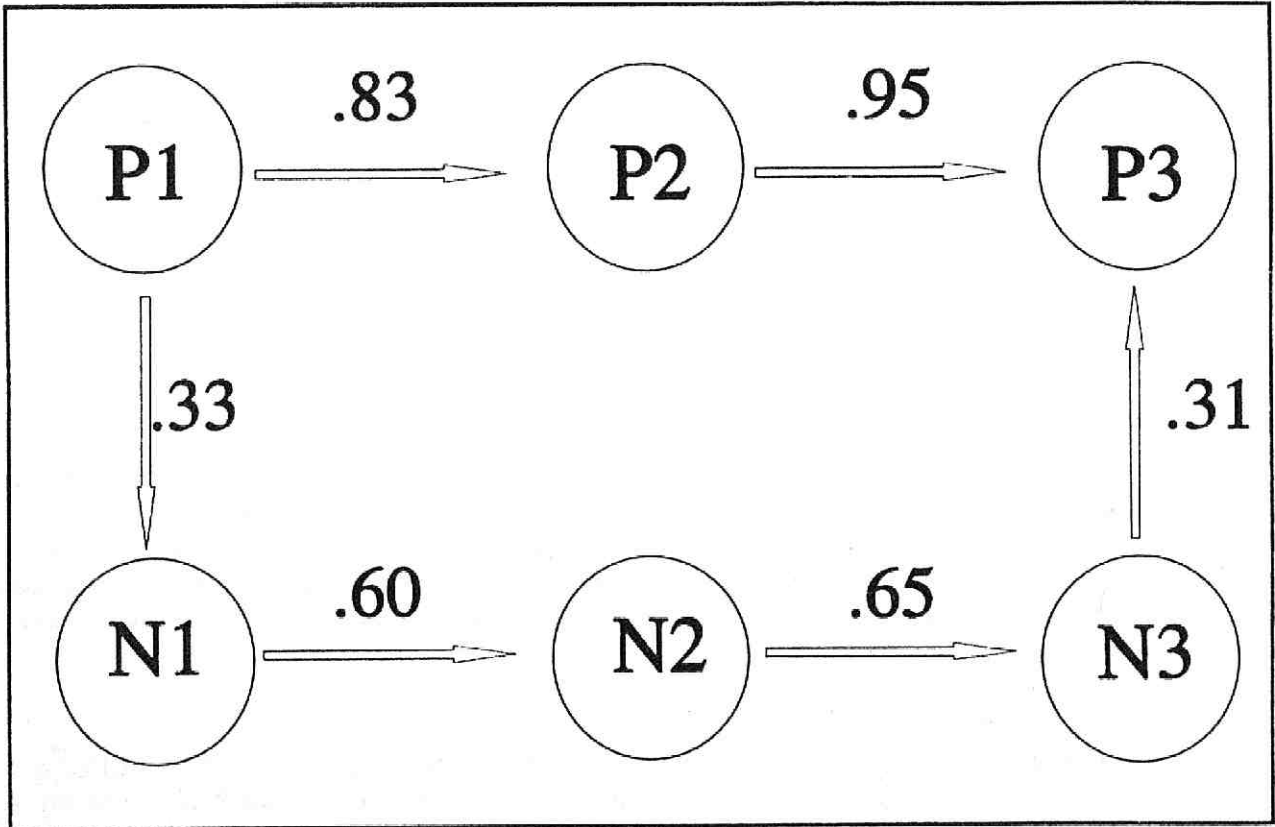


Fig. 5. Gestandaardiseerde oplossing voor het structurele LISREL-model dat de relaties specificceert tussen Papiamentu (P) en Nederlandse taalverwerving bij de Antilliaanse/Arubaanse kinderen op de drie meetmomenten.

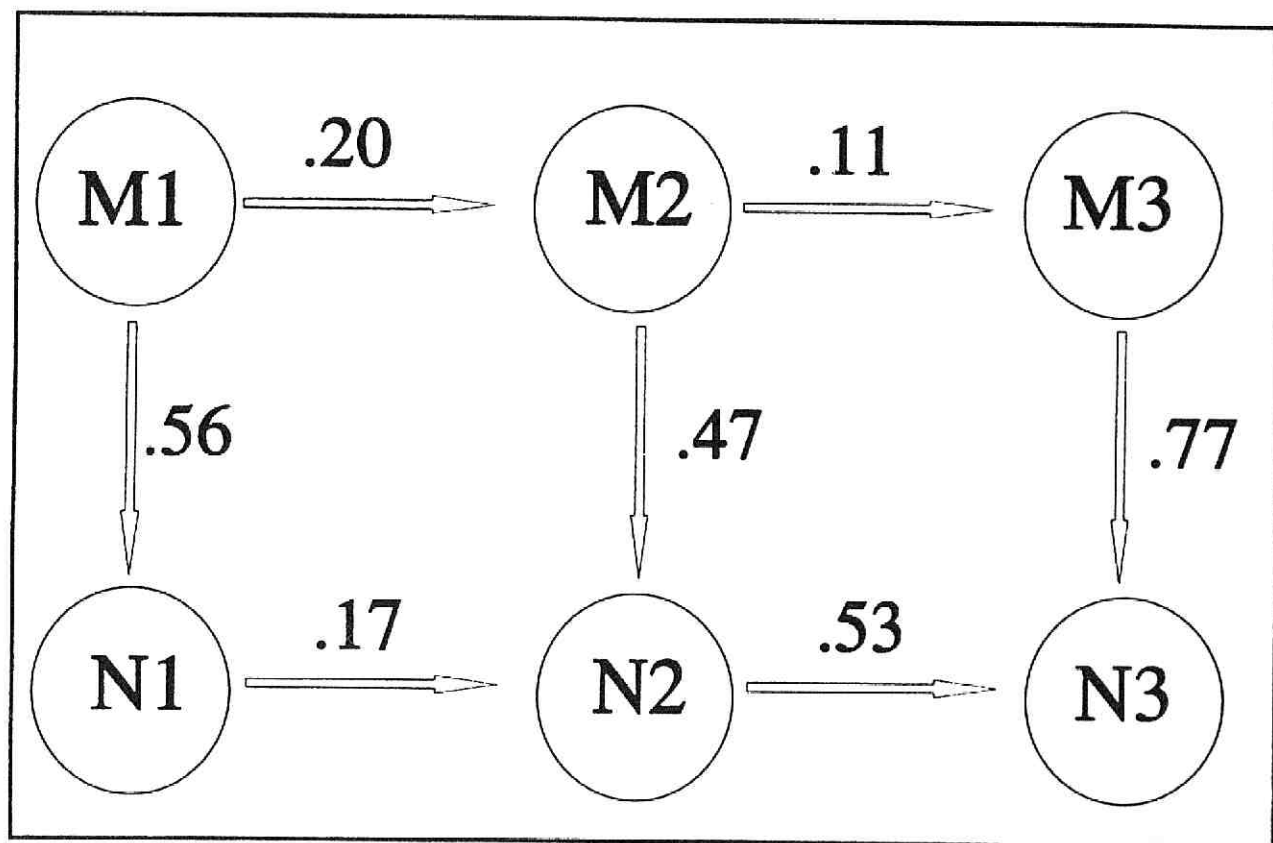


Fig. 6. Gestandaardiseerde oplossing voor het structurele LISREL-model dat de relaties specificceert tussen Marokkaanse (M) en Nederlandse (N) taalverwerving bij de Marokkaanse kinderen op de drie meetmomenten.

sprake van transfer van T1 naar T2, op meetmoment 3 is de relatie juist andersom. Het lijkt erop dat de vorderingen die de kinderen met name in het Nederlands maken eveneens de ontwikkeling van het Papiamentu bevorderen.

Marokkanen

Figuur 6 geeft de gestandaardiseerde LISREL-modeloplossing voor de 69 Marokkaanse kinderen uit het eerste cohort. De goodness-of-fit bleek .91; het model behoefde niet te worden verworpen ($p=.42$).

Het gevonden structurele model gaat uit van relatief zwakke longitudinale verbanden binnen eerste- en tweede-taalontwikkeling. Het longitudinale verband binnen tweede-taalverwerving wordt geleidelijk aan groter. Daarnaast blijkt de transfer van T1 naar T2 op alle meetmomenten groot.

Samenvattend kunnen we stellen dat in de aanvang van het basisonderwijs er evidentie is voor transfer van taalvaardigheden en wel van de dominante, sterkste taal naar de zwakkere taal. Bij de Turkse en Marokkaanse kinderen zien we transfer optreden van de eigen taal naar de tweede taal. Bij de Antilliaanse leerlingen zien we eerst transfer van T1 naar T2 en later in de tijd, vanaf het moment dat het Nederlands de sterkste taal is, juist andersom. Deze bevindingen zijn in overeenstemming met Cummins' hypothese die taaltransfer in twee richtingen toestaat.

Determinanten van tweetaligheid

In het onderzoek naar determinanten van tweetaligheid is uitgegaan van de factorscores van eerste- en tweede-taalvaardigheid bij de aanvang en op het eind van het kleuteronderwijs. Met

Tabel 6. Correlaties tussen kind-, gezins- en onderwijskenmerken en T1- en T2-vaardigheid op het einde van de kleuterperiode.

	Turken		Antillianen		Marokkanen	
	T1	T2	T1	T2	T1	T2
<i>Kindkenmerken</i>						
Taalcontact kind	.27*	.11	.23	.38*	.22	.01
Taalkeuze kind	-.10	.02	.18	.11	.09	.06
Culturele oriëntatie kind	.03	.13	.23	.22	.30	.13
Verblijfsduur kind	-.27*	.02	-.09	.01	.16	.17
<i>Gezinskenmerken</i>						
Gezinstaal	.08	.25*	.52*	.01	.29*	.26
Culturele oriëntatie ouders	.18	.09	.25	.16	-.08	.43**
Verblijfsduur vader	.10	.02	-.40	.20	.20	.10
Verblijfsduur moeder	-.10	.27	-.28*	.09	.26*	.13
SES	.02	.16	.01	.33*	-.20	-.23
<i>Onderwijskenmerken</i>						
Preschool	.09	.10	-.13	.15	.09	.05
OET binnen de klas		-.12	-	-	-	.43**
OET buiten de klas	-.15	-	-	-	.16	-
Extra T2	-	-.08	-	-.18	-	-.27
T2-methodes	-	-.12	-	-.22	-	.32*

behulp van correlatie- en regressie-analyse is per etnische groep gezocht naar de samenhang tussen taalvaardigheidsscores en kind-, gezins- en onderwijskenmerken van allochtone leerlingen uit de cohorten 1 en 2. Tabel 6 geeft de correlaties tussen kind- en gezinskenmerken en T1- en T2-vaardigheid op het eind van groep 2.

Binnen de Turkse groep blijkt taalcontact van het kind positief samen te hangen met T1-vaardigheid, terwijl de factor gezinstaal positieve samenhang vertoont met T2-vaardigheid. Binnen de Antilliaanse groep blijken de correlaties tussen T2-vaardigheid, SES en taalcontact significant. Opvallend binnen deze groep is de negatieve correlatie tussen extra T2-onderwijs en T2-vaardigheid. Dit resultaat kan worden verklaard vanuit het feit dat juist kinderen met een lage T2-vaardigheid in aanmerking komen voor extra T2-onderwijs. In de Marokkaanse groep zijn er significante correlaties tussen T1-vaardigheid enerzijds en gezinstaal, verblijfsduur moeder en onderwijs in eigen taal en cultuur. De variabelen culturele oriëntatie ouders en gebruikte NT2-methodes blijken significant samen te hangen met de T2-vaardigheid. Opvallend is ook hier de redelijk hoge samenhang tussen gezinstaal en T2-vaardigheid.

In Tabel 7 worden de resultaten van multiple-regressie-analyse gegeven met T1- en T2 als criteriumvariabelen en kind- gezins- en institutionele factoren als predictoren.

Binnen de Turkse groep vormen de verblijfsduur en het taalcontact van het kind als belangrijkste voorspellers van T1-vaardigheid en de gezinstaal en het taalcontact van het kind die van T2-vaardigheid. Binnen de Antilliaanse/Arubaanse groep blijkt de T1-vaardigheid te worden geprediceerd vanuit drie variabelen: gezinstaal, culturele oriëntatie ouders en verblijfsduur moeder. Regressie-analyse met taalvaardigheid Nederlands als afhankelijke variabele leverde twee predictoren op: gezinstaal en taalcontact. Binnen de Marokkaanse groep wordt de T1-vaardigheid geprediceerd vanuit drie variabelen: gezinstaal, culturele oriëntatie ouders en OET binnen de school. Regressie-analyse met taalvaardigheid Nederlands als afhankelijke variabele leverde eveneens drie predictoren op, te weten culturele oriëntatie ouders, gezinstaal en het gebruik van NT2 methodes op school.

Tabel 7. Multiple-regressie van eerste- en tweede-taalvaardigheid op predictorvariabelen aan het eind van Groep 2.

Predictor	Beta	R
<i>Turken</i>		
T1:		
Verblijfsduur kind	-.26	
Taalcontact kind	.24	.31**
T2:		
Gezinstaal	.32	
Taalcontact kind	.27	.28*
<i>Antillianen/Arubanen</i>		
T1:		
Gezinstaal	.30	
Culturele oriëntatie ouders	.24	
Verblijfsduur moeder	-.19	.43***
T2:		
Taalcontact kind	.41	
Gezinstaal	.17	.37***
<i>Marokkanen</i>		
T1:		
Culturele oriëntatie ouders	.32	
Gezinstaal	.27	
OET binnen school	.22	.48***
T2:		
Culturele oriëntatie ouders	.37	
Gezinstaal	.19	
NT2-methodes	.19	.50***

Tweetaligheid en metalinguïstisch bewustzijn

In Tabel 8 zijn de gemiddelden en standaarddeviaties opgenomen van de Turkse, Antilliaanse/Arubaanse, Marokkaanse en Nederlandse kinderen op de vijf metalinguïstische taken, afgenomen aan het eind van groep 2.

Tabel 8. Gemiddelde scores metalinguïstische taken aan het einde van groep 2.

	Turken		Antill.		Marokkanen		Nederl.	
	Gem	Sd	Gem	Sd	Gem	Sd	Gem	Sd
Taak 1	5.37	1.96	5.85	2.10	6.21	1.34	5.60	1.95
Taak 2	2.90	4.71	2.15	4.01	5.11	1.19	3.39	5.21
Taak 3	4.91	5.71	3.03	4.19	5.05	1.87	5.32	5.13
Taak 4	5.88	3.35	7.45	2.83	7.93	1.70	8.91	2.37
Taak 5	6.29	3.02	6.79	2.40	7.32	1.26	7.05	2.29

Taak 1= Analyse in woorden (10) Taak 2= Analyse in fonemen (20)

Taak 3= Synthese van klanken (20) Taak 4= Rijm (10) Taak 5 = Objectivatie (10)

De verschillen tussen autochtone en allochtone kinderen op de taken zijn relatief gering. De Marokkaanse leerlingen scoren op vrijwel alle taken hoger dan de Nederlandse leerlingen, waarbij er bovendien sprake is van een geringe spreiding. Op de taken 2 (analyse in fonemen) en 3 (synthese van klanken) scoren de Antilliaanse/Arubaanse kinderen zeer zwak.

Tabel 9 geeft de groepsgegevens voor de kinderen op het eind van de kleuterperiode waarbij onderscheid is gemaakt tussen kinderen die vaardig zijn in beide talen, in één van beide talen of in geen enkele taal.

Er is een duidelijk waarneembare tendens dat de kinderen die over een gebalanceerd hoog niveau van tweetaligheid beschikken het hoogst uitkomen op metalinguïstische taken, dat de kinderen die vaardig zijn in één taal een tussenpositie innemen en dat de kinderen die in geen van beide talen vaardig zijn het laagst scoren. Binnen de Turkse groep komt dit patroon het duidelijkst naar voren. Voor analyse in woorden blijkt het onderscheid significant ($F(2,77)=3.24$, $p<.05$). Voor analyse in fonemen is sprake van een duidelijk tendentie tot significantie ($F(2,77)=2.84$, $p<.10$). Ook voor klanksynthese blijkt het verschil in scores significant ($F(2,77)=4.74$, $p<.01$). Voor rijmen en objectivatie is in beide gevallen sprake van een tendentie tot significantie (resp. ($F(2,77)=2.43$, $p<.10$ en $F(2,77)=2.38$, $p<.10$). Ook binnen de Antilliaanse/Arubaanse groep zien we dat in alle gevallen de kinderen die vaardig zijn in zowel T1 als T2 hoger uitkomen op de metalinguïstische taken dan de kinderen die vaardig zijn in één taal, en dat op hun beurt laatstgenoemde kinderen hoger scoren dan de kinderen die in geen van beide talen vaardig zijn. Voor analyse in woorden blijkt het verschil niet significant. Voor analyse in fonemen is sprake van een tendentie tot significantie ($F(2,59)=2.41$, $p<.10$). Wel duidelijk significant blijkt het verschil voor klanksynthese ($F(2,59)=3.42$, $p<.05$), rijmen ($F(2,59)=3.97$, $p<.05$) en objectivatie ($F(2,59)=4.12$, $p<.05$).

Voor de Marokkaanse groep valt ten slotte af te lezen dat kinderen met een gebalanceerd hoog niveau van tweetaligheid, behalve op klanksynthese gemiddeld het hoogst uitkomen op metalinguïstische taken. Met name voor woordanalyse, fonemische analyse en objectivatie is sprake van duidelijke verschillen. De kinderen met een laag niveau op zowel T1 als T2 scoren het zwakst op alle metalinguïstische taken. De gevonden verschillen tussen deelgroepen zijn

Tabel 9. Gemiddelden op metalinguïstische taken op het eind van groep 2, onderscheiden naar kinderen die vaardig zijn in twee talen, één taal of geen taal.

	Vaardig in 2 talen	Vaardig in 1 taal	Vaardig in geen taal
<i>Turken</i>			
Analyse in woorden	5.63	5.67	4.50
Analyse in fonemen	4.67	2.53	1.55
Synthese van klanken	7.67	4.24	2.95
Rijmen	6.41	4.94	6.41
Objectivatie	6.92	6.59	5.14
<i>Antillianen/Arubanen</i>			
Analyse in woorden	6.25	5.68	5.50
Analyse in fonemen	3.21	2.23	0.44
Synthese van klanken	4.46	2.91	1.06
Rijmen	8.54	7.23	6.13
Objectivatie	7.75	6.55	5.69
<i>Marokkanen</i>			
Analyse in woorden	6.81	6.08	5.57
Analyse in fonemen	5.65	5.08	4.48
Synthese van klanken	5.12	5.85	4.46
Rijmen	8.35	8.38	7.14
Objectivatie	7.75	6.55	5.69

significant voor woordanalyse ($F(2,57)=5.85$, $p<.01$), fonemische analyse ($F(2,57)=6.76$, $p<.01$), rijm ($F(2,57)=3.80$, $p<.05$) en objectivatie ($F(2,57)=4.58$, $p<.01$).

CONCLUSIES EN DISCUSSIE

Profielen van tweetalige ontwikkeling

Uit het in de vorige hoofdstukken gerapporteerd onderzoek mogen we concluderen dat de aard en de mate van tweetaligheid zich per etnische groep verschillend ontwikkelen.

Turkse kinderen blijken vanaf de aanvang van het basisonderwijs tot aan het eind van groep T3 op alle taalvaardigheidsniveaus dominant te zijn in de eigen taal. Dit gegeven komt overeen met de studie van Verhoeven (1987) waaruit bleek dat Turkse kinderen tot aan het eind van groep 4 dominant bleken in de eigen taal. Het huidige onderzoek laat zien dat er tussen deelvaardigheden sprake is van geringe verschillen in ontwikkelingspatronen. Op het niveau van klankmanipulatie ontwikkelt het Turks zich in de eerste fase van het kleuteronderwijs (groep 1 en 2) sneller dan het Nederlands. In de periode daarna bereikt deze ontwikkeling op het Turkse klankniveau een plafond terwijl er in het Nederlands nog sprake is van enige groei. Op het niveau van cognitieve begrippenkennis blijken de Turkse kinderen een betere beheersing van de begrippen in het Turks dan in het Nederlands te hebben gedurende de hele onderbouwperiode van de basisschool (groep 1, 2 en 3). De ontwikkeling van begrippenkennis in het Nederlands verloopt sneller dan die in de eigen taal. De actieve woordenschat blijkt over alle meetmomenten het sterkst ontwikkeld te zijn in de eigen taal van de Turkse kinderen. Het niveau van de actieve woordenschat blijft relatief zwak in de hele onderbouwperiode, zowel in de eigen taal als in het Nederlands. In de kleuterperiode is de ontwikkeling van de actieve woordenschat in beide talen even groot. Geleidelijk aan gaat het Nederlands harder vooruit. Ook op het gebied van de passieve woordenschat is het Turks de dominante taal gedurende de hele periode en groeit het Nederlands in sterkere mate dan het Turks. Op syntactisch niveau presteren de Turkse kinderen beter in de eigen taal dan in het Nederlands. Aan het begin van de kleuterperiode is er sprake van gelijke progressie in beide talen, terwijl aan het eind van de kleuterperiode de ontwikkeling van het Nederlands zelfs sterker is dan die in het Turks. Op het niveau van tekstbegrip is er aanvankelijk sprake van een sterkere positie van de moedertaal. Aan het eind van de kleuterperiode gaat het Nederlands relatief sterker vooruit.

Bij de *Antilliaanse/Arubaanse kinderen* blijkt er bij de aanvang van het basisonderwijs sprake te zijn van een meer gebalanceerd niveau van tweetaligheid. Al vanaf het eerste jaar in het onderwijs blijkt echter een sterkere groei van het Nederlands plaats te vinden dan van het Papiamentu. Dit lijkt vooral het gevolg van het feit dat in Antilliaanse gezinnen meer Nederlands wordt gesproken dan in Turkse of Marokkaanse gezinnen. Geleidelijk aan is er zelfs sprake van een tendens naar dominantie van het Nederlands, met name op het gebied van begrippenkennis. Op het niveau van klankmanipulatie bereikt het Papiamentu in de loop van de tijd een plafond, terwijl het Nederlands nog kan uitlopen. Ten aanzien van kennis van cognitieve begrippen blijken de Antilliaanse kinderen aanvankelijk een vergelijkbaar niveau in beide talen te hebben. In de loop van het basisonderwijs blijkt deze kennis beduidend sterker te worden in het Nederlands dan in de eigen taal. Ook op het gebied van de actieve en passieve woordenschat blijkt aanvankelijk sprake van een gebalanceerd niveau van tweetaligheid. Geleidelijk aan geeft het Nederlands een sterkere groei te zien, terwijl er bij het Papiamentu sprake lijkt van een stagnatie. Eenzelfde patroon is terug te vinden op syntactisch niveau waar het Nederlands een dominante positie inneemt na een redelijk gebalanceerd niveau van de twee talen bij de aanvang van het onderwijs. Op tekstueel niveau loopt de ontwikkeling van beide talen vrijwel parallel en is er gedurende de hele periode in de onderbouw sprake van een redelijk gebalanceerd niveau van de twee talen.

De *Marokkaanse kinderen* zijn bij de aanvang van het basisonderwijs op vrijwel alle taalvaardigheidsniveaus dominant in de eigen taal en blijven dat gedurende de hele periode in de onderbouw van het basisonderwijs. Anders dan Wagenaar (1990) in haar onderzoek concludeerde blijken de Marokkaanse kinderen gedurende de jaargroepen 1-3 vooruitgang te boeken in zowel T1 als T2. Het Marokkaans-Arabisch ontwikkelt zich in deze periode in vergelijkbare mate als het Nederlands. Een uitzondering daarop is de passieve woordenschat waar het Nederlands aan het eind van de benedenbouw (eind groep 3) zich sterker ontwikkelt en het Marokkaans-Arabisch een stagnatie te zien geeft. De ontwikkeling van cognitieve begrippenkennis loopt aanvankelijk parallel in de eerste en de tweede taal, maar aan het einde van de benedenbouw treedt er een stagnatie op in de ontwikkeling van deze kennis in het Nederlands. Aan het eind van de onderbouwperiode loopt de ontwikkeling op syntactisch niveau in de twee talen parallel, met een dominante moedertaal.

Nederlandse taalvaardigheid

Voor wat de Nederlandse taalvaardigheid betreft, blijken er grote verschillen te bestaan tussen de allochtone kinderen enerzijds en de autochtone informanten anderzijds. De ontwikkeling op het niveau van klankmanipulatie vormt hierop een uitzondering. De Nederlandse kinderen scoren op vrijwel alle onderdelen en op ieder meetmoment hoger dan de allochtone kinderen, op de voet gevolgd door de Antilliaanse informanten. De ontwikkeling van het Nederlands verloopt bij de Turkse en Marokkaanse kinderen relatief langzaam en bij de Antilliaanse kinderen relatief snel. De gegevens omtrent de T2-ontwikkeling van Turkse en Marokkaanse kinderen komen overeen met die van Verhoeven & Vermeer (1989) naar aanleiding van het normeringsonderzoek van de Taaltoets Allochtone Kinderen en die van Boogaard, e.a. (1990) in een replicatieonderzoek.

Afhankelijkheden binnen eerste- en tweede-taalverwerving

Gegeven de beperkte omvang van de steekproeven kinderen, enkele voorzichtige conclusies worden getrokken met betrekking tot de binnen het kader van de "afhankelijkheidshypothese" van Cummins veronderstelde verbanden tussen T1- en T2-verwerving. Binnen de Turkse en de Antilliaanse/Antilliaanse groep bleek sprake van sterke longitudinale trends binnen zowel T1- als T2-ontwikkeling. Verder kon bij de aanvang van het basisonderwijs enige evidentie worden teruggevonden voor transfer van de dominante, sterke taal naar de zwakkere taal. Bij de Turkse en Marokkaanse kinderen is T1 de sterkste taal gedurende de gehele kleuterperiode. Bij de Antilliaanse/Arubaanse kleuters is aanvankelijk T1 de sterkste taal, maar zien we op het derde meetmoment dat de T2 dominant geworden is. Uitgaande van de veronderstelling dat transfer optreedt van de dominante taal naar de gedomineerde taal zien we bij de Antilliaanse/Arubaanse groep dat op het eind van het kleuteronderwijs de transfer andersom gaat werken: van T2 naar T1. Binnen de Marokkaanse groep bleken de longitudinale verbanden binnen de T1- en T2-ontwikkeling beduidend zwakker. Dit valt voor zowel de T1 als de T2 mogelijk te verklaren door de invloed van onderwijsfactoren, een gegeven dat exclusief voor de Marokkaanse groep is gevonden. Het is echter ook mogelijk dat binnen de Marokkaanse groep, ondanks de selectieprocedure, kinderen participeerden die van huis uit Berbertalig zijn. Pas in de loop van groep 2 blijkt de T2-vaardigheid van de kinderen een stabiele factor te worden.

Determinanten van tweetaligheid

Voor de *Turkse groep* blijkt de vaardigheid in zowel de eigen taal als het Nederlands bij de aanvang van het basisonderwijs positief samen te hangen met de verblijfsduur van ouders en kind. Het lijkt erop dat de algehele cognitieve ontwikkeling van het kind positief wordt beïnvloed door een langer verblijf in Nederland. Op het einde van de kleuterperiode blijkt de factor taalcontact de belangrijkste voorspeller van T1-vaardigheid. De verblijfsduur van het kind die mede bepalend geacht kan worden voor taalcontact blijkt op dat tijdstip juist negatieve samenhang met de eerste-taalvaardigheid van Turkse kinderen te vertonen. Het is evident dat een langer verblijf in het land van herkomst samengaat met een betere T1-beheersing. Wat het

Nederlands op het eind van de kleuterperiode betreft blijken de factoren gezinstaal en taalcontact van het kind de belangrijkste voorspellers.

Binnen de *Antilliaanse/Arubaanse groep* blijkt de vaardigheid in Papiamentu bij het begin van het kleuteronderwijs positief samen te hangen met sociaal-culturele factoren, zoals gezinstaal, taalkeuze en culturele oriëntatie van het kind en negatief met verblijfsduur van ouders en kind en met de mate van preschoolervaring. De beide negatieve verbanden kunnen worden verklaard vanuit het feit dat een lange verblijfsduur en een frequent bezoek aan een preschool in Nederland een verminderd contact met Papiamentu impliceert. Op het eind van de kleuterperiode blijken vooral de factoren gezinstaal, culturele oriëntatie van ouders en de verblijfsduur van de moeder belangrijke voorspellers van T1-vaardigheid. Wat het Nederlands van Antilliaanse/Arubaanse kleuters betreft zien we een brede range van sociaal-culturele factoren positieve samenhang vertonen. De culturele oriëntatie van de ouders, het taalcontact en de taalkeuze van het kind en de verblijfsduur van de moeder blijken de belangrijkste indicatoren. Op het eind van de kleuterperiode kunnen de sociaal-economische achtergrond en taalcontact van het kind als belangrijkste voorspellers worden aangemerkt.

Binnen de *Marokkaanse groep* blijkt bij de aanvang van het kleuteronderwijs sprake van een positieve samenhang tussen T1-vaardigheid en de factoren gezinstaal en culturele oriëntatie van de ouders. Op het eind van de kleuterperiode blijken beide factoren nog steeds belangrijke voorspellers van T1-vaardigheid. Daarnaast komt onderwijs in eigen taal en cultuur als relevante predictor naar voren. Wat het Nederlands als tweede taal betreft blijken taalcontact en taalkeuze van het kind naast de culturele oriëntatie van ouders belangrijke voorspellers. Op het eind van de kleuterperiode blijken vooral gezinstaal en culturele oriëntatie naast de inzet van NT2-methodes relevante voorspellers van T2-vaardigheid.

Samenvattend kunnen we stellen dat de gezinstaal en het taalcontact van het kind belangrijke indicatoren vormen van zowel de ontwikkeling van T1- als T2-vaardigheid. Daarnaast blijken de verblijfsduur van ouders en kinderen en de culturele oriëntatie van ouders belangrijke additionele voorspellers van met name de T1-vaardigheid van kinderen. Deze resultaten komen overeen met bevindingen uit eerder onderzoek.

Opvallend is de geringe invloed van preschoolervaring op de ontwikkeling van T2-vaardigheid bij de aanvang van het basisonderwijs. Wellicht dat de aard van het taalaanbod in de voorschoolse periode bepalender is dan de kwantitatieve maat van het aantal maanden dat kinderen institutionele opvang in die periode hebben genoten. Daarnaast springt ook de inconsistente invloed van onderwijsfactoren in het kleuteronderwijs in het oog. Alleen bij de Marokkanen blijkt de mate van OET een relevante voorspeller van T1-vaardigheid en de inzet van T2-methodes van T2-vaardigheid op het eind van de kleuterperiode. Binnen de Antilliaanse/Arubaanse groep is zelfs sprake van een negatieve samenhang van extra T2-onderwijs en T2-vaardigheid. Dit gegeven kan worden verklaard vanuit het feit dat mogelijk juist de kinderen met een zwakke T2-vaardigheid voor extra T2-onderwijs werden uitgekozen. Nader onderzoek naar het effect van uiteenlopende organisatieprincipes van T1- en T2-onderwijs is gewenst. Het lijkt aannemelijk dat daarbij veeleer etnografische methoden nodig zijn dan de hier gehanteerde methode van interviews en documentenanalyse (zie ook Damhuis, e.a., 1989:265).

Metalinguïstisch bewustzijn

Ten aanzien van tweetaligheid en metalinguïstisch bewustzijn laat het onderzoek zien dat de verschillen tussen allochtone en autochtone kinderen in metalinguïstisch bewustzijn relatief gering zijn. De scores van Marokkaanse kinderen op enkele metalinguïstische taken zijn op het eind van groep 2 zelfs hoger dan die van Nederlandse kinderen. Allochtone leerlingen maken in groep 3 bovendien een vergelijkbare progressie in metalinguïstische ontwikkeling als autochtone leerlingen. Deze uitkomsten komen min of meer overeen met de bevindingen uit onderzoek naar de ontwikkeling van metalinguïstisch bewustzijn onder allochtone en autochtone kinderen

van Verhoeven & Van Kuijk (1991). Zij vonden dat op taken die gerelateerd zijn aan fonologisch bewustzijn er vrijwel geen verschillen zijn tussen allochtone en autochtone kleuters.

Verder kunnen we concluderen dat er een duidelijke samenhang is tussen de aard en mate van tweetaligheid enerzijds en metalinguïstisch bewustzijn anderzijds. Kinderen die aan het eind van de kleuterperiode een min of meer gebalanceerd hoog niveau van tweetaligheid bezitten scoren hoger op metalinguïstische taken dan kinderen die slechts één taal goed beheersen. De kinderen die laag scoren op zowel T1 als T2 scoren ook het laagst op metalinguïstische taken. Blijkbaar leidt een hoog niveau van tweetaligheid tot een goed inzicht in de structuur van taal als zodanig. Dit gegeven komt overeen met eerder onderzoek (vgl. Cummins, 1983, 1984; Diaz, 1983; Hakuta, 1986; Albert & Obler, 1978; Ben-Zeev, 1977).

Aanbevelingen voor de onderwijspraktijk

Het huidige onderzoek toont aan dat allochtone kinderen uit gezinnen waar thuis de eigen taal wordt gesproken de basisschool binnenkomen met een dominantie van de eigen taal. Hoe het patroon van tweetalige ontwikkeling verder verloopt is afhankelijk van het acculturatieproces van de etnische groep waar het kind deel van uitmaakt en van uiteenlopende kind-, gezins- en schoolkenmerken. Turkse en Marokkaanse kinderen blijven gedurende de kleuterperiode dominant in de eigen taal. In beide talen gaan zij vooruit, in sterkere mate geldt dit echter voor het Nederlands. Antilliaanse/Arubaanse kinderen blijken in de loop van de onderbouw van het basisonderwijs geleidelijk aan dominant te worden in het Nederlands. Van belang is het gegeven dat er bij alle drie de groepen sprake is van transfer van de dominante taal naar de gedomineerde taal. Bij Turkse en Marokkaanse kinderen gaat de transfer steeds van T1 naar T2; bij de Antilliaanse/Arubaanse groep is de richting van taaltransfer aan het eind van de kleuterperiode juist andersom.

Het onderzoek laat verder zien dat er tussen allochtone kinderen sprake kan zijn van aanzienlijke verschillen in profielen van tweetalige ontwikkeling. Voor kinderen die (nog) dominant zijn in de eigen taal kan het van belang zijn om de cognitieve ontwikkeling te stimuleren vanuit die taal. Tegelijkertijd dient echter het Nederlands te worden bevorderd. Op het moment dat kinderen een gebalanceerd niveau van tweetaligheid bereiken kan de cognitieve stimulering even goed vanuit het Nederlands worden verzorgd. Het onderwijs in de eigen taal is vanaf dat moment niet zozeer vanuit onderwijskundige, dan wel vanuit culturele redenen te motiveren.

Om in te kunnen spelen op de ontwikkeling van kinderen vanaf hun entree in de basisschool hebben zowel Nederlandse als allochtone leerkrachten behoefte aan evaluatietechnieken. Differentiatie en individualisering binnen het taalonderwijs zijn pas goed mogelijk wanneer leerkrachten instrumenten in handen hebben waarmee een concreet beeld wordt verkregen van de aard en mate van tweetaligheid. Vandaar dat besloten is om het in dit project ontwikkelde toetsinstrumentarium onder te brengen in het Cito leerlingvolgsysteem. Behalve voor de onderdelen op klankniveau blijkt voor de toetstaken geen sprake van plafondeffecten. Het ontwikkelde toetsinstrumentarium kan daarmee zeer bruikbaar worden geacht voor het volgen van de tweetalige ontwikkeling van allochtone kinderen in de onderbouw van het basisonderwijs.

LITERATUUR

- Albert, M., & L. Obler, L. (1987). *The bilingual brain: neuropsychological and neurolinguistic aspects of bilingualism*. New York: Academic Press.
- Appel, R. (1984). *Immigrant children learning Dutch. Sociolinguistic and psycholinguistic aspects of second language acquisition*. Dordrecht: Foris.
- Ben-Zeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. *Child Development*, 48, 1009-1018.
- Boers, M., van den Bosch, L., & Verhoeven, L. (1987). Taalaanbod en culturele oriëntatie als voorspellers van tweede-taalvaardigheid. *Tijdschrift voor Taalbeheersing*, 9 (2), 166-180.
- Boeschoten, H. (1990). *Acquisition of Turkish by immigrant children. A multiple case study of Turkish children in the Netherlands aged 4 to 6*. Wiesbaden: Harrassowitz.

- Boogaard, M., Damhuis, R., de Glopper, K., & v.d. Bergh, H. (1990). *De Nederlandse taalvaardigheid van allochtone en Nederlandse kleuters*. Lisse: Zwets & Zeitlinger.
- Clark, H., & Clark, E. (1977). *Psychology and language. An introduction to psycholinguistics*. New York : Harcourt Brace and Jovanovich.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 221-251.
- Cummins J. (1980). The cross-lingual dimensions of language proficiency. *TESOL Quarterly*, 14, 175-187
- Cummins, J. (1983). Language and literacy in bilingual instruction. *ERIC*, ED245-575.
- Cummins, J. (1984). Wanted: a theoretical framework for relating language proficiency to academic achievement among bilingual students. In C. Rivera (ed.). *Language proficiency and academic achievement*. Clevedon: Multilingual Matters.
- Cummins J., & Swain, M. (1986). *Bilingualism in education*. London: Longman.
- Diaz, R.M. (1983). Thought and two languages: The impact of bilingualism on cognitive development. In E.W. Gordon (ed.). *Review of research in education*, vol X. Washington, DC: AERA.
- Downing, J., & Valtin, R. (1984). *Language awareness and learning to read*. New York: Springer Verlag.
- Els, T. van e.a. (1985). *Applied linguistics and the learning and teaching of foreign languages*. London: Edward Arnold.
- Extra, G., & Verhoeven, L. (1993). A bilingual perspective on Turkish and Moroccan children and adults. In G. Extra & L. Verhoeven (eds.). *Immigrant languages in Europe*. Clevedon : Multilingual Matters.
- Fantini, A. (1985). *Language acquisition of a bilingual child*. Clevedon: Multilingual Matters.
- Gardner R. (1968). Attitudes and motivations: Their role in second language acquisition. *TESOL Quarterly*, 141-150.
- Genesee, F., & Hamayan, E. (1980). Individual differences in second language learning. *Applied Psycholinguistics*, 1, 95-110.
- Hakuta, R. (1986). *Mirror of language: The debate on bilingualism*. New York: Basic Books.
- Harley, B., & Lapkin, S. (1984). *The effects of early bilingual schooling on first language development*. Toronto: OISE.
- Helvert, K. van (1985). *Nederlands van en tegen Turkse kinderen*. Nijmegen: dissertatie KUN.
- Lalleman, J. (1986). *Dutch language proficiency of Turkish children born in the Netherlands*. Dordrecht: Foris.
- Lambert, W. (1978). Psychological approaches to bilingualism, translation and interpretation. In D. Gerver & H. Sinaiko (eds.). *Language interpretation and communication*. New York: Plenum Press.
- Larsen-Freeman, D., & Long, M. (1991). *An introduction to second language acquisition research*. London: Longman.
- McLaughlin, B. (1985). *Second language acquisition in childhood*. Hillsdale: Erlbaum.
- McLeod, B., & McLaughlin, B. (1986). Restructuring or automaticity ? Reading in a second language. *Language Learning*, 36, 109-123.
- Pels, T. (1991). *Marokkaanse kleuters en hun culturele kapitaal; opvoeden en leren in het gezin en op school*. Lisse: Swets & Zeitlinger.
- Ruiter, J.J. de (1989). *Young Moroccans in the Netherlands. An integral approach to their language situation and acquisition of Dutch*. Utrecht : dissertatie RUU.
- Schumann, J. (1978). The acculturation model for second language acquisition. In R. Gingras (ed.). *Second language acquisition and foreign language teaching*. Arlington: Center for Applied Linguistics.
- Tosi, A. (1979). Mother tongue teaching for the children of migrants. *Language Teaching and Linguistics: Abstracts*. 12, 213-231.
- Tosi, A. (1984). *Immigration and Bilingual Education*. Oxford: Pergamon Press.
- Teunissen, F. (1986). *Een school, twee talen*. Utrecht : Vakgroep Onderwijskunde.
- Verhelst Tunmer, W., Pratt, C., & Herriman, M. (1984). *Metalinguistic awareness in children. Theory, research and implications*. Berlin: Springer Verlag.
- Veltman, C. (1983). *Language shift in the United States*. Berlin: Mouton.
- Verhelst, N., & Eggen, T. (1989). *Psychometrische en statistische aspecten van peilingsonderzoek*, PPON-rapport 4. Arnhem: Cito.
- Verhoeven, L. (1987). *Ethnic minority children acquiring literacy* (Published Ph. D. dissertation). Dordrecht: Foris.
- Verhoeven, L. (1990). Acquisition of reading in Dutch as a second language. *Reading Research Quarterly*, 25 (2), 90-114.
- Verhoeven, L. (1991a). Predicting minority children's bilingual proficiency: Child, family and institutional factors. *Language Learning*, 41 (2), 205-233.
- Verhoeven, L. (1991b). Acquisition of biliteracy. *AILA Review*, 8, 61-74.

- Verhoeven, L. (1992). Assessment of bilingual proficiency. In J. de Jong & L. Verhoeven (eds.), *The construct of language proficiency*. Amsterdam/Philadelphia: John Benjamins.
- Verhoeven, L. (1994). Transfer in bilingual development: The linguistic interdependence hypothesis revisited. *Language Learning*, 44, (3), 381-415.
- Verhoeven, L., & Boeschoten, H. (1986). First language acquisition in a second language environment. *Journal of Applied Psycholinguistics* 7, (3), 241-255.
- Verhoeven, L., & van Kuijk, J. (1991). Conceptuele en metalinguïstische kennis bij jongste kleuters. *Pedagogische Studiën*, 68, (9), 60-65.
- Verhoeven, L., & Vermeer, A. (1987). *Taaltoets Allochtone Kinderen*. Tilburg: Zwijssen.
- Verhoeven, L., & Vermeer, A. (1989). *Diagnose van kindertaal*. Tilburg: Zwijssen.
- Verhoeven, L., Extra, G., Konak, O., Narain, G., & Zerrouk, R. (1990). *Vroege tweetaligheid bij Turkse, Marokkaanse en Antilliaanse/Arubaanse kleuters*. Tilburg: Tilburg University Press.
- Vermeer, A. (1986). *Tempo en structuur van tweede-taalverwerving bij Turkse en Marokkaanse kinderen*. Tilburg: dissertatie KUB.
- Wagenaar, E. (1990). *Kleuterklas en moedertaal: een onderzoek naar de effecten van tweetalig onderwijs aan Marokkaanse kleuters*. UVA: Vakgroep Orthopedagogiek.
- Wells, G. (1985). *Language development in the preschool years*. Cambridge: University Press.
- Willig, A. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research*, 55, 269-317.
- Wong-Fillmore, L. (1976). *The second time around: Cognitive and social strategies in second language acquisition*. Dissertation Stanford University.
- Wong-Fillmore, L. (1979). Individual differences in second language learning. In C.J. Fillmore, W.S. Wang & D. Kempler (eds.), *Individual differences in language ability and language behavior*. New York: Academic Press.
- Wong-Fillmore, L. (1982). Instructional language as linguistic input: Second Language learning in classrooms. In L. Cherry-Wilkinson (ed.), *Communication in the Classroom*. New York: Academic Press.

Manuscript ontvangen 6-9-1994

Definitieve versie ontvangen 5-4-1995

Actuele opvattingen over onderwijs aan jonge kinderen

Een onderzoek onder leerkrachten van groep 2 in het basisonderwijs

L.M. Stevens¹ en E.G. Harskamp²

¹*Vakgroep Pedagogiek, Universiteit Utrecht*

²*GION, Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling, Rijksuniversiteit Groningen*

ABSTRACT

In 1985, Dutch primary schools (6 to 12 year old children) and nursery schools (4 to 6 year old children) were integrated into "new primary schools". This article deals with the attitudes of teachers in grade 2, the educational goals they find important for young children in school and the nature of the decisions they make in ambiguous situations. The results show that teachers with a positive attitude towards direct education favour cognitive objectives and offer more cognitive activities to their pupils than teachers with an attitude towards indirect education. However, all teachers agree upon the great value of non-cognitive objectives and activities. There is no difference between directive and non-directive teachers in pedagogical decision making in ambiguous situations. Recommendations are made for an instrument to make explicit the objectives, pedagogical approach and educational activities in grade 1 and 2.

1. INLEIDING

De survey die hier wordt gerapporteerd maakt deel uit van een aantal voorstudies ten dienste van de ontwikkeling van een interventieprogramma voor de verbetering van individuele leerkracht-leerling interacties in het onderwijs aan jonge kinderen. Het interventieprogramma heeft als doel leerkrachtgedrag te bevorderen dat responsief is ten opzichte van onderwijsbehoeften van kinderen en dat de zelfstandigheid en de speel- en werkcompetentie van kinderen stimuleert. Het programma wordt gebaseerd op een reeds bestaand en geëvalueerd interventieprogramma voor leraren van oudere kinderen in het basisonderwijs (Stevens e.a., 1992; Jager, 1993). De survey is voortgekomen uit de vraag in hoeverre leerkrachten van de groepen 1 en 2 toegankelijk zullen zijn voor een interventie die de dagelijkse omgang met hun leerlingen direct raakt en die hen confronteren zal met hun opvattingen over de taken van onderwijs aan kleuters. Sinds de voorbereiding en vaststelling van de nieuwe Wet op het Basisonderwijs is de discussie over doel, inhoud en werkwijze van het onderwijs aan kleuters in een stroomversnelling geraakt. Ondanks de angst om de zogenaamde verworvenheden van het kleuteronderwijs (Van Erp & Van Tuyl, 1981) te moeten prijsgeven en de angst om de in de opleidingen aanwezige expertise te moeten verliezen (De Frankrijker & De Kler, 1988; Lindt, 1987, 1988) werd de discussie geleidelijk sterker gericht op (de wenselijkheid van) een meer beredeneerd onderwijsaanbod voor het jonge kind in de zin van meer reflectief en doordacht handelen (ARBO, 1990; Janssen-Vos, 1990; Van Kuyk, 1987).

De onderwijspraktijk adapteert deze ontwikkeling echter langzaam. In onderzoek dat tot dusver is uitgevoerd komt naar voren dat vele leerkrachten in de onderbouw vooral nog zogenaamd kind-gecentreerd of kind-volgend onderwijs geven, hetgeen betekent dat men zich terughoudend opstelt ten aanzien van een bewust aanbod, zoals een vastgesteld programma voor

spel- en werkactiviteiten. Men maakt weinig gebruik van beschikbare programma's voor voorbereidend rekenen (Harskamp & Willemsen, 1991, constateren dat 67% van de leerkrachten in groep 2 een programma gebruikt, maar dat in dit verband gemiddeld niet meer dan twee lesactiviteiten per maand voorkomen) of voor taal en voorbereidend lezen (Harskamp e.a., 1993, constateren dat 85% van de leerkrachten in groep 2 gemiddeld acht keer per maand een lesactiviteit uit een taalprogramma gebruikt en dat 55% van de leerkrachten regelmatig aandacht besteedt aan herkennen van letters of woorden in dagelijkse situaties). De leerkrachten in de onderbouw zijn vooral gericht op het aanbieden van een voorbereide speel- en werkomgeving waarin de kinderen veel keuzevrijheid hebben en slechts in beperkte mate gestuurd worden. Er zijn echter verschillen tussen leerkrachten in de mate waarin ze keuzevrijheid toelaten en de mate waarin zij sturend optreden (zie van Gennip & Ruijs, 1992; Harskamp, Pijl & Snippe, 1991; Reezigt, Guldemon & Ros, 1993).

In deze context is een gerichte exploratie wenselijk van actuele opvattingen van leerkrachten en de mate waarin leerkrachten in het kleuteronderwijs sturend wensen op te treden, alvorens bovenbedoeld interventieprogramma te ontwikkelen. Opvattingen vormen waarschijnlijk niet alleen het referentiekader voor het dagelijkse handelen, maar ook voor de beoordeling van (nieuwe) handelingsvoorstellen of interventies en zijn daarmee essentiële condities voor het slagen van een innovatie (Clark & Petersen, 1986; Van den Berg & Vandenberghe, 1988). Het is echter niet eenvoudig om de onderwijsopvattingen van leerkrachten te meten, omdat leerkrachten niet altijd dezelfde betekenis geven aan termen als kindsturend of kindvolgend onderwijs. Bovendien is het zo dat opvattingen van leerkrachten niet altijd in overeenstemming zijn met het handelen van leerkrachten (zie Wolbert e.a., 1986).

2. ACHTERGRONDEN, PROBLEEMSTELLING EN ONDERZOEKSVRAGEN

Opvattingen van leerkrachten kan men samenvatten in wat een onderwijsvisie wordt genoemd, door Van Oers (1992) gedefinieerd als "... het stelsel van fundamentele (theoretische en normatieve) persoonlijke uitgangspunten die voor een persoon de grondslag vormen om antwoord te geven op de vraag wat goed onderwijs, c.q. "good practice" is" (p.5). De geschiedschrijving van het kleuteronderwijs en zijn opleidingen (zie bijvoorbeeld Nijkamp, 1963, 1973; Van Vroonhoven, 1985) en de hedendaagse praktijk, zoals deze verschijnt in systematische beschrijvingen (Van Gennip & Ruijs, 1992; Harskamp e.a., 1991; Reezigt e.a., 1993; Wolbert e.a., 1986) lijken te wijzen op een over tijd en situaties heen zeer stabiele en consistente visie in ons land op het onderwijs aan kleuters. De kern van de visie behelst de overtuiging dat kinderen van nature actief en leergierig zijn en derhalve de belangrijkste actor in hun eigen ontwikkeling. De leerkracht dient in deze visie de oorspronkelijke, actieve toewending van het kind tot de wereld en het vertrouwen in zichzelf en de omgeving dat hieruit spreekt niet te verstoren, doch te respecteren. De leerkracht moet het kind verder geleiden in een emotioneel ondersteunende en interesse-wekkende omgeving. De leerkracht heeft dus een faciliterende rol in het kleuteronderwijs. De onderwijsomgeving, respectievelijk het onderwijsaanbod dat hierbij past wordt verwoord in termen van "verworvenheden van het kleuteronderwijs" (Van Erp & Van Tuyl, 1981; Van Erp & De Visch Eibergen, 1983; Van Kuyk, 1987). Kenmerkend hiervoor is een (kleuter)school die kindgericht is (het primaat van het kind-initiatief), individugericht (verschillen tussen kinderen worden geaccepteerd en gehonoreerd) en die een kind in zijn totaliteit ziet. Voorts een school die niet programmagebonden is en een pluriforme onderwijssituatie biedt: een school waar prestatie en ontwikkeling niet aan normen gebonden zijn. Kinderen ontwikkelen zich naar eigen tempo en belangstelling.

Deze onderwijsvisie en onderwijspraktijk hebben een lange geschiedenis (zie bijvoorbeeld Nijkamp, 1973; Phillipi-Siewertz van Reesema, 1954). We beperken ons voor het onderwijs aan kleuters tot zeer invloedrijke denkers en onderwijsontwerpers als Fröbel, Montessori, Dewey en Decroly. Zij allen benadrukten, hoewel in verschillende definities, het "eigen wezen" en de "eigen wereld" van het jonge kind en de daaraan verbonden eigengeaarde behoeften. Dit "ei-

gen" verwijst vooral naar spontane behoefte aan exploratie door handelen, manifest in spel, eventueel te richten en te geleiden door passende materialen. In de overdracht van het gedachtegoed van met name Fröbel en Montessori hebben de opleidingen voor kleuterleidsters (KLOS) in Nederland vanaf de vorige eeuw een actieve bemiddelende rol gespeeld (De Haan, 1953; Nijkamp, 1966, 1973; Philippi-Siewertz van Reesema, 1954).

Tegen deze achtergrond mag men overwegend een kindvolgende benadering van de leerkracht verwachten, waarin geen specifieke leerdoelen of leervoorwaarden worden nagestreefd (zie Laevers, 1992).

Anderzijds, zoals in de inleiding reeds werd gesuggereerd, worden leerkrachten in groep 1 en 2 mede door de druk vanuit de hogere leerjaren, uitgedaagd een meer beredeneerd onderwijsaanbod te ontwikkelen en gebruik te maken van daartoe ontworpen onderwijsprogramma's (zie bijvoorbeeld De Baar, 1992; Van Kuyk, 1987, 1992) of van meer gestructureerde, doordachte algemene benaderingen (zie bijvoorbeeld Janssen-Vos, 1990). Deze ontwikkeling, die een combinatie inhoudt van kindvolgend en kindsturend onderwijs, is inmiddels een klein decennium oud. Men zou als effect bij leerkrachten in groep 1 en 2 minstens een groeiend bewustzijn mogen waarnemen van het belang van een meer beredeneerd onderwijsaanbod, naast de bekende brede basisdoelstellingen (zie Janssen-Vos, 1987). De onderwijsinspectie heeft in 1992 een onderzoek op tachtig scholen uitgevoerd naar de aard van het aanbod in groep 1 en 2 en de actualiteit en breedte van deskundigheid van onderwijsgeveenden. De conclusies zijn dat 39% van de leerkrachten in groep 1 en 2 voldoende kennis van recente inzichten in het leren en de ontwikkeling van jonge kinderen heeft en dat 55% in groep 1 en 2 voldoende brede didactische vaardigheden heeft voor het bevorderen van het leren en de ontwikkeling van jonge kinderen (Inspectie van het Onderwijs, 1994). Op grond van deze resultaten worden geen grote veranderingen verwacht in aandacht voor andere zaken dan brede basisdoelstellingen. Deze verwachting komt in de uitwerking van het onderzoek aan de orde.

Voorts wordt ingegaan op de vraag of de opleidingsachtergrond (KLOS enerzijds en PA of PABO anderzijds) een differentieel effect heeft op opvattingen over onderwijs aan kleuters. De PA en PABO opleiding hebben altijd sterk de nadruk gelegd op het programmagerichte onderwijs in de bovenbouw van de basisschool, terwijl de KLOS sterk gericht was op het kindgerichte onderwijs voor de onderbouw. Op de huidige PABO rest nog een zeer bescheiden expertise van oud KLOS docenten op het gebied van het onderwijs aan kleuters na de ingrijpende reorganisaties van PA en KLOS (De Frankrijker & De Kler, 1988; Lindt, 1987, 1988; Visitatiecommissie PABO, 1993). De veronderstelling van sommigen dat het kleuteronderwijs verschoolst is, wordt niet gesteund door onderzoek (zie Reezigt, e.a., 1993). Het kleuteronderwijs als organisatie en cultuur weerstaat kennelijk vooralsnog zonder problemen eventuele druk van buitenaf om het onderwijsaanbod meer in overeenstemming te brengen met het aanbod in de rest van de basisschool. Zo vonden Harskamp e.a. (1991) tussen leerkrachten met een KLOS of een PA/PABO opleiding geen verschil in aanbod op het gebied van speel-werk en leeractiviteiten in groep 2 van de basisschool en geen verschil in opvattingen over het onderwijs.

De vragen waarop dit onderzoek een antwoord probeert te vinden luiden als volgt:

1. Welke opvattingen hebben leerkrachten die met kleuters werken over het kleuteronderwijs, respectievelijk het sturen van activiteiten van kleuters? Is er verschil in opvatting tussen KLOS- en PABO-opgeleide leerkrachten?

Verwacht wordt dat de leerkrachten in de kleuterklassen overwegend kindvolgend denken en handelen en dat er, ondanks de verschillen in vooropleiding tussen leerkrachten met een KLOS-opleiding en leerkrachten met een PABO-opleiding, geen verschil in opvatting is.

2. Welke doelen stellen leerkrachten ten aanzien van de kinderen?

Verwacht wordt dat de meeste leerkrachten vooral algemene doelstellingen met betrekking tot de affectieve en cognitieve ontwikkeling van kinderen zullen nastreven (zogenaamde 'brede basisdoelstellingen') en dat de leerkrachten relatief weinig belang zullen hechten aan specifieke leerdoelen of voorwaarden zoals die in voorbereidende taal- en rekenprogramma's worden aangeboden.

3. Welk aanbod voor taal en voorbereidend lezen en rekenen geven leerkrachten aan kinderen? Verwacht wordt dat de meeste leerkrachten vooral activiteiten ten behoeve van algemene leervoorwaarden zullen aanbieden en relatief weinig aandacht aan specifieke leervoorwaarden voor taal, lezen en rekenen zullen besteden.

4. Welke relatie is er tussen de opvattingen van leerkrachten (zie vraag 1) en het handelen van leerkrachten in situaties waarin interacties tussen leerkracht en leerling vereist zijn?

Het is te verwachten dat leerkrachten met een sturende opvatting over activiteiten van kleuters meer sturende beslissingen zullen nemen in de klaspraktijk dan leerkrachten met een kindvolgende opvatting.

5. Welke relatie is er tussen de doeloriëntaties van leerkrachten (zie vraag 2) en het aanbod van leervoorwaarden (zie vraag 3)?

Verwacht wordt dat leerkrachten die meer belang hechten aan (algemene) leervoorwaarden ook vaker leervoorwaarden in hun aanbod opnemen dan de overige leerkrachten.

3. ONDERZOEKSOPZET

In maart 1992 is een a-selecte steekproef van 58 scholen getrokken uit de populatie scholen in de provincie Utrecht. Veertig kleuterleerkrachten van deze scholen waren bereid deel te nemen aan het onderzoek. De leerkrachten zijn geïnterviewd aan de hand van een vragenlijst. De vragen hebben betrekking op groep 2 van de basisschool. De lijst bevat gerichte vragen met een open antwoordmogelijkheid. Er is voor open antwoordvragen gekozen, omdat nog maar weinig onderzoek is gedaan naar het denken en handelen van kleuterleerkrachten (zie Reezigt e.a., 1993). De vragen hebben betrekking op vijf onderwerpen:

1. Achtergrondgegevens van de leerkrachten: de vooropleiding (KLOS, PA/PABO of andere vooropleiding), de leeftijd, het geslacht en het totaal aantal jaren ervaring in het kleuteronderwijs.
2. De opvattingen van de leerkrachten: gevraagd is 'Welke benadering spreekt u het meest aan, een kindvolgende benadering of een kindsturende benadering?' Er is een toelichting gegeven op dit onderscheid. Tevens is een open vraag gesteld of leerkrachten vinden dat er een typische eigen taak van de kleuterbouw in vergelijking met de midden- en bovenbouw is en zo ja, welke specifieke taken het onderwijs aan kleuters heeft.
3. Doeloriëntatie: leerkrachten krijgen een lijst met vaardigheidsgebieden voorgelegd die in het kleuteronderwijs aan bod kunnen komen. Als voorbeelden van dergelijke gebieden zijn genoemd de gebieden: motorische ontwikkeling, sociaal-emotionele ontwikkeling, taalontwikkeling, leervoorwaarden lezen, schrijven en rekenen, zelfstandigheid, concentratie en werkhouding. Leerkrachten werden verzocht ook zelf gebieden te noemen. Er is gevraagd 'Welke twee gebieden van vaardigheden vindt u het belangrijkste om een goede start te maken in groep 3? Voor elk van de twee door hen genoemde gebieden moesten de leerkrachten een voorbeeld van een belangrijke doelstelling geven.
4. Aanbod van leervoorwaarden: leerkrachten is gevraagd of ze tenminste één keer per week aan leervoorwaarden werken met de kinderen en zo ja, aan welke leervoorwaarden vooral; aangegeven is dat het om specifieke voorwaarden met betrekking tot lezen, schrijven of rekenen of om meer algemene leervoorwaarden kan gaan.
5. Pedagogisch-didactische beslissingen: er is een achttal probleemsituaties voorgelegd waarin pedagogisch/didactisch handelen vereist is (zie Tabel 3); gevraagd is: 'Hoe reageert u in zo'n situatie?' De spontane antwoorden van de leerkrachten zijn opgeschreven en achteraf gescoord van sterk naar matig tot niet sturend handelen. Door berekening van de coëfficiënt alpha is nagegaan in hoeverre met de acht situaties, op consistente wijze, het begrip 'sturend handelen in probleemsituaties' kan worden gemeten.

De procedure is als volgt verlopen. De 40 leerkrachten zijn a-select verdeeld over vier medewerkers. De interviews zijn op geluidsband opgenomen en duurden 1 tot 1½ uur, afhankelijk van het verloop van het gesprek. De antwoorden zijn verwerkt door de respons

van leerkrachten binnen achteraf onderscheiden antwoordcategorieën te scoren. De scoring is verricht door de vier onderzoeksmedewerkers onder leiding van een ervaren onderzoeker, die alle antwoorden gezamenlijk hebben beoordeeld.

Het onderzoek is te typeren als een exploratieve studie met een ex post facto opzet. Er is gebruik gemaakt van de half open interviewmethode om een betrouwbaar beeld te krijgen van de opvattingen en de handelwijze van leerkrachten. Een eenmalige observatie in de klas leek onvoldoende om een representatief beeld te krijgen van de dagelijkse handelwijze. Voor herhaalde leerkrachtobservaties waren geen financiële mogelijkheden. Deze opzet maakt dat de resultaten met voorzichtigheid moeten worden geïnterpreteerd.

4. RESULTATEN

De onderzoeksgroep bestaat geheel uit vrouwelijke leerkrachten met een gemiddelde leeftijd tussen 30 en 39 jaar. De meerderheid van hen (78%) heeft een KLOS-opleiding, de overigen (22%) hebben een opleiding op de PA of PABO gevolgd. Van de 40 leerkrachten hebben er 31 meer dan 5 jaar ervaring in het kleuteronderwijs en 9 hebben minder dan 5 jaar ervaring, maar minimaal 2 jaar. Vergelijkbare gegevens zijn gevonden in het onderzoek van Harskamp, Pijl en Snippe (1991): 82% KLOS-opleiding en 18% PA/PABO opleiding.

4.1. Opvattingen over het onderwijs in de kleuterfase

Met betrekking tot de opvattingen over kindvolgend of kindsturend onderwijs is gevraagd welke benadering van kleuters de leerkracht het meest aanspreekt. De vraag is als volgt gesteld. *Leerkrachten in de onderbouw denken genuanceerd over het onderwijs aan kleuters. Sommige benadrukken het belang van een kindvolgende benadering, waarbij het kind centraal staat, anderen benadrukken het belang van een kindsturende benadering, waarbij het programma centraal staat. Het één sluit het ander niet uit.*

- a. Welke benadering spreekt u het meest aan?
- b. Waarom kindvolgend?
- c. Waarom kindsturend?

De meeste leerkrachten (80%) kiezen voor kindvolgend onderwijs. Dit is in overeenstemming met onze verwachting vooraf. In vervolg hierop is de leerkrachten gevraagd of het onderwijs in groep 2 een eigen taak heeft in vergelijking met de andere groepen in de basisschool. Een ruime meerderheid (95%) ziet wel degelijk een typisch eigen taak voor de kleuterbouw in vergelijking met de midden- en bovenbouw van de basisschool. De door de leerkrachten zelf genoemde taken voor het kleuteronderwijs zijn in drie hoofdcategorieën ingedeeld:

1. ontwikkeling stimuleren (hieronder vallen taken als: onderwijs laten aansluiten bij het ontwikkelingsniveau van een kind, spelonderwijs leren of sociaal-emotionele ontwikkeling stimuleren);
2. schoolgedrag bevorderen (hieronder vallen taken als: het wennen van de kleuters aan en motiveren voor het schoolleven);
3. leervoorwaarden bevorderen (hieronder vallen taken als: voorbereiden op leren lezen, schrijven en rekenen).

Een leerkracht kon op meer dan één categorie scoren. Er zijn verschillende combinaties van antwoorden. De resultaten zijn vermeld in Tabel 1. Er zijn 40 respondenten. Vanwege het geringe aantal te verwachten respondenten, verdeeld over de 7 x 2 cellen van de tabel, zijn geen verschillen getoetst tussen kindvolgende en kindsturende leerkrachten. Het onderscheid tussen leerkrachten die kiezen voor een kindvolgende benadering en leerkrachten die kiezen voor een kindsturende benadering is gemaakt vanuit descriptief oogpunt.

Het stimuleren van de ontwikkeling scoort erg hoog bij de leerkrachten: 93% noemt dit, eventueel samen met andere taken, als typische taak voor het kleuteronderwijs. Het bevorderen van schoolgedrag wordt door 36% van de leerkrachten als taak gezien en het bevorderen van leervoorwaarden door 23%.

Leerkrachten met een kindsturende opvatting en met een kindvolgende opvatting zijn vergeleken ten aanzien van de eigen taken die zij voor het kleuteronderwijs noemen. Het blijkt dat het verschil in opvatting vooral tot uiting komt in het feit dat leerkrachten die kiezen voor kindsturend onderwijs vaker (50%) het bevorderen van leervoorwaarden als taak van het kleuteronderwijs noemen naast het stimuleren van de ontwikkeling en het bevorderen van schoolgedrag dan kindvolgende leerkrachten (16%). Het verschil is statistisch significant ($\chi^2 = 8.02$; $p = .01$). Kindvolgende leerkrachten zien leervoorwaarden minder als taak voor het kleuteronderwijs dan kindsturende leerkrachten, terwijl ze ongeveer even vaak het belang van het stimuleren van de algehele ontwikkeling van kinderen benadrukken.

Een groot deel van de leerkrachten in de kleuterbouw is opgeleid op de KLOS. Deze opleiding was specifiek voor het kleuteronderwijs. De overige leerkrachten zijn opgeleid op de PA (Pedagogische Academie) of de PABO (Pedagogische Academie Basis-Onderwijs). Men zou kunnen verwachten dat leerkrachten met een KLOS-opleiding andere opvattingen hebben over de taken van de kleuterbouw dan leerkrachten met een PA- of PABO-opleiding. Eerder onderzoek weerspreekt echter deze verwachting.

Tabel 2, waarin de door de leerkrachten genoemde eigen taken van de kleuterbouw zijn uitgesplitst naar opleiding van de leerkrachten, bevestigt eerdere onderzoeksuitkomsten. Leerkrachten afkomstig van de KLOS verschillen niet duidelijk in antwoorden van leerkrachten met een PA- of PABO-opleiding. Dit is ook in overeenstemming met onze verwachting.

4.2. Doeloriëntatie

In de vorige paragraaf is ingegaan op de belangrijkste taken die leerkrachten onderkennen voor het onderwijs in groep 2. Deze paragraaf behandelt de specifieke doelen die leerkrachten voor kinderen stellen.

Tabel 1. Ideeën over typisch eigen taak onderbouw naar benadering.

Taak	Benadering		Totaal (n=40)
	Kindvolgend (n=32)	Kindsturend (n=8)	
Ontwikkeling stimuleren	59%	50%	57%
Ontwikkeling stimuleren en bevorderen schoolgedrag	16%	0%	13%
Ontwikkeling stimuleren en leervoorwaarden en schoolgedrag bevorderen	16%	50%	23%
Alleen schoolgedrag of alleen leervoorwaarden bevorderen	0%	0%	0%
Geen eigen taak/geen specificatie	9%	0%	7%

Tabel 2. Meningen over de typisch eigen taak van het kleuteronderwijs, onderscheiden naar vooropleiding van respondenten..

Taak	Opleiding KLOS (n=31)	PA/PABO (n=9)
Ontwikkeling stimuleren	61%	44%
Ontwikkeling stimuleren en schoolgedrag bevorderen	13%	11%
Ontwikkeling stimuleren en leervoorwaarden en schoolgedrag	20%	33%
Geen eigen taak/geen specificatie	7%	11%

($\chi^2 = .49$; $p = .70$)

Tabel 3. De vaardigheden die het belangrijkste genoemd worden voor groep 2 met voor iedere vaardigheid enkele voorbeelden die genoemd zijn.

Vaardigheden	Percentage
Taalontwikkeling	
kunnen luisteren	
in volledige zinnen kunnen praten	50%
Zelfstandigheid	
zelf kunnen aan- en uitkleden	
zelfstandig aan opdracht kunnen werken	40%
Werkhouding	
taakgericht en geconcentreerd kunnen werken	
zelfstandig tot 30 minuten met een werkje bezig kunnen zijn	30%
Motorische ontwikkeling	
kunnen hanteren van knip- en schrijfgerei en pengreep	
kunnen huppelen, hinkelen en steppen	30%
Aanpassen in nieuwe situaties	
geen problemen hebben bij het aanpassen in nieuwe situaties	
flexibel zijn	22%
Sociaal-emotionele ontwikkeling	
zelfstandig een probleem of conflict kunnen oplossen	
samen kunnen spelen en werken	12%
Ruimtelijke oriëntatie	
ruimtelijke begrippen kennen en begrijpen	
kunnen oriënteren in de ruimte	10%
Beheersing specifieke leervoorwaarden	
rekenvoorwaarden beheersen	
leesvoorwaarden beheersen	3%
Spel	
rollenspellen kunnen spelen	
kunnen samenspelen	3%
Totaal	200%

Welke doelen stellen de leerkrachten ten aanzien van de vaardigheden die de kinderen in groep 2 moeten beheersen? De leerkrachten is gevraagd om de twee belangrijkste vaardigheden aan te geven en daarbij telkens voorbeelden te noemen. Als voorbeelden moeten verwachtingen worden genoemd van wat leerlingen aan het eind van groep 2 moeten beheersen. In Tabel 3 zijn de resultaten weergegeven. Per vaardigheid zijn de meest genoemde voorbeelden vermeld. Omdat de respondenten twee keuzes mochten aangeven, sommeren de percentages tot 200%. De 50% achter taalontwikkeling betekent dat 50% van de leerkrachten als één van de twee keuzes taalontwikkeling heeft genoemd.

Opmerkelijk maar niet onverwacht, is dat de prioriteiten wat betreft het belang van vaardigheden overeen komen met de prioriteiten van het traditionele kleuteronderwijs (zie Van Oers & Janssen-Vos, 1992). Dit wijst erop dat de prioriteiten van de leerkrachten van de kleuterbouw die nu in de basisschool werkzaam zijn niet veel zullen verschillen van kleuterleerkrachten in het verleden.

Taalontwikkeling wordt door de helft van de leerkrachten de belangrijkste vaardigheid van groep 2 genoemd. Onder taalontwikkeling wordt vooral het spreken en luisteren verstaan en niet het ontwikkelen van voorbereidende vaardigheden voor lezen, schrijven of rekenen. Aan deze specifieke voorwaarden wordt weinig waarde toegekend. Op de tweede plaats komt de zelfredzaamheid en het zelfstandig werken van leerlingen. Dit wordt door 40% van de leerkrachten genoemd. Op de gedeelde derde plaats staan taakgericht werken en de ontwikkeling van de

Tabel 4. Verschillen in doeloriëntaties tussen groepen leerkrachten.

	Kindvolgend (n=32)	Kindsturend (n=8)	KLOS (n=31)	PA/PABO (n=9)
Cognitieve vaardigheden van belang genoemd (n=23)	50%	88%	55%	67%
Alleen niet cognitieve vaardigheden van belang genoemd (n=17)	50%	13%	45%	33%
	$\chi^2 = 3.68; p=.00$		$\chi^2 = .40; p=.53$	

grove en fijne motoriek. Deze worden door 30% van de leerkrachten genoemd. Opvallend is dat beheersing van specifieke leervoorwaarden (bijvoorbeeld voor rekenen en lezen) nauwelijks genoemd wordt (3%) bij de twee belangrijkste vaardigheden voor kinderen in groep 2. De resultaten zijn in overeenstemming met de verwachting vooraf dat leerkrachten vooral brede basisdoelstellingen na willen streven en geen specifieke leervoorwaarden.

De vaardigheden en doelen die worden genoemd zijn onder te verdelen in twee hoofdgroepen: 1. cognitieve vaardigheden (taalontwikkeling, ruimtelijke oriëntatie en beheersing specifieke leervoorwaarden), 2. niet cognitieve vaardigheden (zelfstandigheid, werkhouding, motorische ontwikkeling, aanpassing in nieuwe situaties, sociaal-emotionele ontwikkeling en spel).

Nagegaan is of er tussen leerkrachten met een KLOS of een PA/PABO opleiding of leerkrachten met een kindvolgende en leerkrachten met een kindsturende opvatting verschillen zijn ten aanzien van het belang dat aan deze twee groepen van vaardigheden wordt gehecht. De resultaten zijn in Tabel 4 vermeld.

De tabel laat zien dat 50% van de kindvolgende leerkrachten (n=32) cognitieve vaardigheden prioriteit geeft en 50% van deze leerkrachten niet. Van de leerkrachten met een kindsturende opvatting (n=8) geeft 88% cognitieve vaardigheden prioriteit en 13% niet. Er is een tendens waar te nemen dat kindsturende leerkrachten over het algemeen meer waarde hechten aan cognitieve vaardigheden. Verschillen tussen leerkrachten met een KLOS en PA/PABO-opleiding in het belang dat men hecht aan cognitieve vaardigheden zijn niet groot genoeg om van betekenis te zijn.

4.3. Aanbod van leervoorwaarden

Leerkrachten is gevraagd of zij tenminste één keer per week systematisch laten werken aan leervoorwaarden. Er is daarbij een onderscheid gemaakt in specifieke voorwaarden voor het leren lezen, schrijven of rekenen en algemene voorwaarden, bijvoorbeeld op het gebied van taalbeheersing of taakgericht werken.

Al eerder, in Tabel 3, is gebleken dat slechts 3% van de leerkrachten het beheersen van specifieke leervoorwaarden voor lezen en rekenen als belangrijke doelstelling noemt. Op de vraag of de leerkracht één keer per week of vaker systematisch aan leervoorwaarden werkt met de kinderen, antwoordt 90% van de leerkrachten dat zij aan leervoorwaarden werken. Ongeveer 10% zegt dit niet te doen.

Bij de groep die aan leervoorwaarden werkt, is nagegaan over welke voorwaarden het gaat (zie Tabel 5). De leerkrachten zijn ingedeeld in vier categorieën: zij die specifieke leervoorwaarden aanbieden, zij die algemene leervoorwaarden aanbieden, zij die beide aanbieden en zij die geen leervoorwaarden aanbieden in hun onderwijs, althans niet op systematische wijze. De voorwaarden die door ten minste 10% van de respondenten zelf zijn genoemd zijn onder de hoofdcategorieën vermeld.

De tabel laat zien dat 35% van de leerkrachten alleen activiteiten voor de specifieke leervoorwaarden aanbiedt, 18% biedt alleen activiteiten aan voor algemene leervoorwaarden en 38% biedt activiteiten op beide terreinen aan. Dit betekent dat 56% van de leerkrachten in hun

Tabel 5. Het aanbod van leervoorwaarden in groep 2.

Alleen aan specifieke leervoorwaarden rekenvoorwaarden leesvoorwaarden	18%
Alleen aan algemene leervoorwaarden voorbereidend schrijven taalontwikkeling begrippenkennis kleurenkennis	35%
Zowel specifieke als algemene leervoorwaarden	38%
Geen leervoorwaarden	10%

Tabel 6. Verschillen in aanbod van specifieke leervoorwaarden tussen groepen leerkrachten.

	Kindvolgend (n=32)	Kindsturend (n=8)	KLOS (n=31)	PA/PABO (n=9)
Specifieke leervoorwaarden in het aanbod (n=22)	44%	100%	61%	33%
Geen specifieke leervoorwaarden in het aanbod (n=18)	56%	0%	39%	67%
	$\chi^2 = 8.18$; $p = .004$		$\chi^2 = 2.20$; $p = .14$	

aanbod activiteiten biedt voor specifieke leervoorwaarden, terwijl men dit voor de leerlingen, blijkens Tabel 3, niet van grote prioriteit vindt. Dit resultaat is dan ook tegengesteld aan de vooraf geformuleerde verwachting dat de meeste leerkrachten vooral algemene en weinig specifieke leervoorwaarden zullen nastreven. Er is evenwel een samenhang tussen de mening over het belang van cognitieve doelen (taalontwikkeling, ruimtelijke oriëntatie en beheersing van leervoorwaarden; zie Tabel 3) en het wel of niet aanbieden van specifieke leervoorwaarden ($\chi^2 = 7.8$; $p = .01$). De analyse laat zien dat leerkrachten die cognitieve vaardigheden prioriteit geven vaker ook specifieke leervoorwaarden aanbieden dan leerkrachten die cognitieve doelen geen prioriteit geven.

Het is interessant om na te gaan of er verschillen zijn tussen leerkrachten met een kindvolgende of kindsturende opvatting en leerkrachten met KLOS of PA/PABO-opleiding, ook al is de verdeling van aantallen leerkrachten in beide groepsindelingen erg ongelijk. De resultaten zijn in Tabel 6 vermeld.

De tabel laat zien dat alle kindsturende leerkrachten specifieke leervoorwaarden aanbieden, terwijl van de kindvolgende leerkrachten slechts 44% dit doet. Tussen leerkrachten met een KLOS of PA/PABO-opleiding is geen significant verschil in het aanbieden van specifieke leervoorwaarden.

4.4. Beslissingen over het pedagogisch/didactisch handelen in probleemsituaties

Nu is het interessant om te kijken of opvattingen van leerkrachten samenhangen met beslissingen van leerkrachten in situaties waarin handelend moet worden opgetreden. De leerkrachten kregen acht probleemsituaties uit de praktijk van de kleuterklas voorgelegd met de vraag wat zij in de gegeven situatie zouden doen. De situaties zijn gekozen op basis van vermoede representativiteit voor voorkomende probleemgetinte situaties in de kleuterklas en op basis van vermoede validiteit van de voorgenomen interventie (zie inleiding), waarbij de leerkracht gevraagd wordt zich in het betekenisperspectief van de leerling te plaatsen (Castelijns e.a., 1992; Jager, 1993).

Eén item betreft de algemene orde in de klas. Eén item betreft de wederzijdse acceptatie van de leerlingen. Twee items verwijzen naar de afgesproken regels voor het werken door de kleuters. Vier items verwijzen naar individuele kwetsbaarheden van leerlingen en de problemen die deze oproepen. De antwoorden zijn in drie categorieën van handelen ingedeeld: sterk sturend, matig sturend en niet sturend.

De betrouwbaarheid van het instrument is redelijk ($\alpha = .74$). De acht items hebben een item-restcorrelatie van .29 tot .45. In Tabel 7 zijn de probleemsituaties weergegeven. Per situatie zijn de meest genoemde antwoorden van de leerkrachten vermeld op de vraag: 'Wat zou u doen in deze situatie?'. De tabel laat zien hoe deze antwoorden zijn verwerkt in 3 categorieën van pedagogisch/didactisch handelen: sterk sturend, matig sturend en niet sturend.

Het beeld is genuanceerd. De items die verwijzen naar geldende regels, in dit geval de algemene orde in de klas en de regels voor zelfstandig werken (Claudia en Kees) roepen in het algemeen sturend gedrag op. Het item dat verwijst naar de wederzijdse acceptatie van de kinderen (Tommy) doet dit heel sterk. Vermoedelijk moeten we hier rekening houden met het impliciete morele gehalte van dit item, er wordt een norm overschreden. De vier items die verwijzen naar individuele kwetsbaarheden van kinderen (Peter, Kim, Steef, Frits) roepen daarentegen typisch begrijpend, acceptierend (kindvolgend) gedrag bij de leerkracht op.

De groep leerkrachten met een kindsturende opvatting blijken niet vaker voor een sturende oplossing te kiezen, dan de groep leerkrachten met een kindvolgende opvatting ($F 1,38 = 8,7$; $p=.28$). De groep leerkrachten met een KLOS-opleiding verschilt niet van de groep leerkrachten met een PA/PABO-opleiding in de mate waarin een sturende oplossing voor de probleemsituaties wordt gekozen ($F 1,38 = 1,1$; $p=.71$). Het eerste resultaat ondersteunt de vooraf geformuleerde verwachting niet, het tweede resultaat ondersteunt wel de verwachting.

5. CONCLUSIES EN DISCUSSIE

Doel van het onderzoek is om actuele informatie te krijgen over onderwijsopvattingen en daarmee samenhangende doeloriëntaties en beslissingen in pedagogisch/didactische situaties van leerkrachten in de kleuterbouw. Met deze informatie kan een interventieprogramma voor kleuterleerkrachten gericht op onderwijsbehoeften van jonge kinderen vorm worden gegeven. Het onderzoek is uitgevoerd bij een kleine steekproef van scholen en baseert zich op uitspraken van leerkrachten en niet op observaties van leerkrachten en leerlingen. Voorzichtigheid bij de interpretatie van gegevens is geboden.

De onderzochte kleuterleerkrachten hebben volgens verwachting in grote meerderheid (80%) een kindgerichte onderwijsvisie. Vrijwel alle leerkrachten (95%) zijn van mening dat de kleuterbouw een typisch eigen taak heeft ten opzichte van de midden- en bovenbouw van de basisschool. De taken voor het kleuteronderwijs die leerkrachten noemen kunnen worden herleid tot drie categorieën, namelijk het stimuleren van de cognitieve en sociaal-emotionele ontwikkeling van het kind (93% ziet dit als taak), het bevorderen van schoolgedrag (36%) en het bevorderen van leervoorwaarden (23%). Het blijkt dat 58% van de leerkrachten het stimuleren van de ontwikkeling van de leerlingen als belangrijkste taak van de kleuterbouw ziet. Stimuleren houdt in dat men een voorbereide speel/werkomgeving aanbiedt zonder kinderen te sturen. Deze groep houdt vast aan de tradities van het kleuteronderwijs en voelt er weinig voor om jonge kinderen een vastgesteld onderwijsprogramma aan te bieden.

Leerkrachten in groep 2 kiezen in meerderheid voor een kindvolgende onderwijsopvatting waarin het stimuleren van kinderen en niet het onderwijsprogramma centraal staat. Leerkrachten die kiezen voor kindsturend onderwijs zien vaker dan leerkrachten met een kindvolgende opvatting, het bevorderen van leervoorwaarden als taak voor het kleuteronderwijs. Of men een KLOS-opleiding danwel een PA- of PABO-opleiding heeft gehad is, overeenkomstig de verwachtingen, niet van invloed op de taakopvattingen.

De belangrijkste doelstellingen voor het kleuteronderwijs liggen volgens de leerkrachten op

Tabel 7. Probleemsituaties in de klas.

De klas is erg onrustig en u wilt met de kring beginnen.	
Sterk sturend	
-in handen klappen	
-persoonlijk appèl (bijvoorbeeld namen noemen)	59%
Matig sturend	
-nadrukkelijk staan wachten	15%
Niet sturend	
-zingen om de aandacht te trekken	26%
Peter komt voor de zoveelste keer bij u klagen over kinderen die hem pesten.	
Sterk sturend	
-kinderen van Peter's groep ter verantwoording roepen	5%
Matig sturend	
-Peter terug sturen en Peter en zijn groep extra in de gaten houden	
-gesprek met Peter en zijn groep over de oorzaak	8%
Niet sturend	
-gesprek met kind (laten uitleggen)	
-geen aandacht aan geven	87%
Kim durft niet in het klauterraam, zelfs niet op de eerste sport. Wanneer de andere kinderen naar boven klauteren, staat Kim huilend beneden.	
Sterk sturend	
-samen met Kim proberen (erbij blijven, vasthouden, helpende hand)	18%
Matig sturend	
-stimulerende opmerkingen om zelfvertrouwen te verbeteren	41%
Niet sturend	
-accepteren (niet dwingen, wat anders laten doen)	41%
In de kring is het stil. Steef die altijd heel erg verlegen is en nooit in de kring praat, begint te praten. Dat is tegen de regels.	
Sterk sturend	
-niet laten praten (aan de regels houden)	5%
Matig sturend	
-laten praten, maar wel op regels wijzen	34%
Niet sturend	
-laten praten en stimuleren	61%
Tommie heeft een zwakke fijne motoriek. Toch doet hij erg z'n best met het knippen van een ster. De andere kinderen lachen om het resultaat.	
Sterk sturend	
-algemene terechtwijzing tegen die kinderen of de hele groep	
-Tommie prijzen en het gedrag van de andere kinderen afkeuren	90%
Matig sturend	
-taak aanpassen	
-Tommie's werkje even wat verbeteren	5%
Niet sturend	
-Tommie positief benaderen	5%
U laat een groepje zelfstandig een bootje vouwen. Iedereen werkt behalve Claudia. Wanneer het tijd is om op te ruimen, maakt Claudia snel en slordig een bootje.	
Sterk sturend	
-er wordt al ingegrepen voordat de les om is	
-het werkje afkeuren en over laten doen	71%
Matig sturend	
-andere kinderen een opmerking over het werkje laten maken	3%
Niet sturend	
-resultaat bespreken en vragen wat Claudia er zelf van vindt	
-resultaat accepteren	26%

Vervolg Tabel 7

Vervolg Tabel 7. Probleemsituaties in de klas.

Wanneer Frits zijn werkje inlevert lijkt het altijd precies op uw voorbeeld of op het werkje van zijn vriendje.	
Sterk sturend	
-apart zetten (zonder voorbeeld)	29%
Matig sturend	
-oorzaak achterhalen en daarop inspelen	13%
Niet sturend	
-stimuleren om zelf iets te maken	
-resultaat accepteren	58%
Nog voordat u de nieuwe opdracht heeft uitgelegd, is Kees al aan het werk. Uit het resultaat blijkt weer eens dat Kees te snel met zijn werk is begonnen.	
Sterk sturend	
-er wordt direct ingegrepen	
-zeggen dat Kees eerst moet luisteren, wachten op instructie	
-bestraffend toespreken	
-Kees het werkje laten overdoen	77%
Matig sturend	
-kind instructie laten herhalen	
-oorzaak achterhalen en daarop inspelen	15%
Niet sturend	
-resultaat accepteren	8%

het gebied van: taalontwikkeling (50%), zelfstandigheid (40%), werkhouding (30%) en motorische ontwikkeling (30%). Zowel de opvattingen over onderwijs als de doelstellingen die de meeste leerkrachten van belang vinden blijken nauw overeen te komen met die van het kleuteronderwijs van vóór de invoering van de Wet op het Basisonderwijs in 1985 (zie Janssen-Vos, 1991).

Leerkrachten achten specifieke leervoorwaarden niet van groot belang voor leerlingen in groep 2 (slechts 3% noemt dit als prioriteit). Daar staat tegenover dat 90% van de leerkrachten zegt op een of andere wijze aan leervoorwaarden te werken. Van deze groep blijkt 18% uitsluitend te werken aan specifieke voorwaarden voor lezen, schrijven of rekenen, 35% werkt aan algemene leervoorwaarden en 38% werkt aan beide typen voorwaarden.

Een groot deel van de leerkrachten (56%) werkt dus meer of minder intensief aan specifieke leervoorwaarden, maar beschouwt deze niet als doelstellingen met grote prioriteit voor leerlingen in groep 2. De specifieke leervoorwaarden waaraan wordt gewerkt liggen vooral op het gebied van het leren lezen. Recent inventariserend onderzoek naar het aanbod van Nederlandse taal in groep 2 laat zien dat 55% van de leerkrachten één keer per week of vaker activiteiten rond specifieke voorwaarden voor het leren lezen aanbiedt (Harskamp e.a., 1993). Dit resultaat komt goed overeen met het percentage van 56% leerkrachten dat in dit onderzoek aangeeft specifieke leervoorwaarden na te streven (zie Tabel 5).

Een verklaring voor de opvallende discrepantie in dit onderzoek tussen de lage prioriteit voor specifieke voorwaarden enerzijds en het grote aantal leerkrachten dat er niettemin aan werkt anderzijds, is dat leerkrachten zich door de eisen van buitenaf gedwongen voelen aan leervoorwaarden te werken, maar dit liever niet doen. Deze veronderstelling is te nuanceren. Over het geheel genomen is er, zoals verwacht, een positieve samenhang tussen de oriëntatie op cognitieve doelen (de cognitieve doelen zijn: taalontwikkeling, ruimtelijke oriëntatie en leervoorwaarden, zie Tabel 3) en het aanbieden van specifieke leervoorwaarden in groep 2. Veel leerkrachten (58%) vinden algemene cognitieve doelen van belang, terwijl ze weinig belang hechten aan specifieke leervoorwaarden. Vooral deze groep leerkrachten neemt, mogelijk onder aandrang van buitenaf, specifieke leervoorwaarden in hun aanbod op.

We kunnen ook aan de hand van deze gegevens constateren dat er verschillen in het aanbod van speel- en leeractiviteiten zijn tussen leerkrachten van jonge kinderen (zie verder Van Gennip en Ruijs, 1992; Harskamp, Pijl & Snippe 1991 en Reezigt, Guldemond & Ros 1993). Er is nog geen sprake van een algemeen geaccepteerd en duidelijk herkenbaar curriculum voor de groepen 1 en 2.

Wat het pedagogisch/didactisch handelen in probleemsituaties betreft, kunnen we het volgende opmerken. Wanneer het om orde, regels en normen voor de hele groep gaat kiezen de leerkrachten overwegend voor sturend handelen. Bij individu-gebonden problemen kiezen zij overwegend voor niet-sturend handelen. Zij maken deze keuzes onafhankelijk van hun opvatting over kindsturend of kindvolgend onderwijs. Dit is geheel tegen de verwachting in. De keuze om wel of niet sturend op te treden in probleemsituaties lijkt eerder bepaald door de kenmerken van de situaties dan door algemene opvattingen van de leerkrachten. Dit betekent dat voor het trainen van leerkrachten in het handelen in probleemsituaties het niet nodig is om onderscheid te maken in interventies die passen bij kindvolgende danwel kindsturende leerkrachten.

Een kindvolgende of kindsturende opvatting hangt dus **niet** samen met beslissingen in het pedagogisch-didactisch handelen, maar hangt **wel** samen met prioriteiten in onderwijsdoelen en met het aanbod van leervoorwaarden in groep 2. Verschillen in opvattingen zullen via de voorkeur voor bepaalde doelen op cognitief, sociaal, emotioneel, motorisch en creatief gebied (zie Tabel 3) doorwerken op het onderwijsaanbod. De verschillen in opvattingen van leerkrachten zijn minder relevant voor het samenstellen van een interventieprogramma waarin getracht wordt het pedagogisch/didactisch handelen beter af te stemmen op de onderwijsbehoeften van kinderen. Voor het programma is het, gezien de huidige onderzoeksresultaten, wel gewenst dat er duidelijkheid is over de doelen en het onderwijsaanbod die individuele leerkrachten in groep 1 en 2 nastreven, zodat interventies worden gericht op situaties tijdens speel-, werk- en leeractiviteiten die door leerkrachten van groot belang worden geacht (zie Jager, 1993). In dit onderzoek is gebleken dat leerkrachten aan kunnen geven op welke algemene doelen en welk type speel- en leeractiviteiten zij de nadruk leggen in hun onderwijs.

Meer in het algemeen is het voor verbetering van de kwaliteit van het onderwijs aan jonge kinderen nodig dat leerkrachten expliciteren welke onderwijsdoelen en pedagogische aanpak van kinderen zij na willen streven. Op basis van verder onderzoek kan worden nagegaan welke feitelijk nagestreefde doelen en pedagogische handelwijze in groep 1 en 2 kunnen worden geëxpliciteerd en in welke mate er overeenstemming is tussen groep 1, 2 en groep 3 in de overgang van spelen naar leren. De onderwijsinspectie constateerde in 1994 dat op slechts 40% van de scholen het schoolwerkplan voldoende aanwijzingen geeft over de wijze waarop het onderwijs in groep 1 en 2 is ingericht en het aanbod dat men heeft voor de leerlingen. Concrete handreikingen bij het expliciteren van pedagogische handelwijzen en onderwijsinhouden zijn nodig om leerkrachten bewust te maken van hetgeen zij doen met de kinderen in groep 1 en 2 en om de doelen en het aanbod zonodig aan te passen.

DANKBETUIGING

Met dank aan Tanja Derks, Laura Prinsen, Lidy Sibon en in het bijzonder aan Anne Marie van Muijen die de leerkrachten interviewden.

LITERATUUR

- ARBO (1990). *Spelen en jongleren*. Advies over de vernieuwing van het onderwijs aan jonge kinderen.
- Baar, K. de (1992). Leren lezen met jonge kinderen. Ervaringsgericht en ontwikkelingsgericht leesonderwijs. In B. van Oers & F. Janssen-Vos. *Visies op onderwijs aan jonge kinderen*. Assen/ Maastricht: Van Gorcum.
- Berg, van den, R., & Vandenbergh, R. (1988). *Onderwijsvernieuwing op een keerpunt*. Tilburg: Zwijssen.
- Castelijns, J., Damen, H., Stevens, L.M., Werkhoven, W. van, & Jager, A. (1992). *Individuele hulp in de klas: bevorderen van taakgericht gedrag door responsiviteit*. Utrecht/Hoevelaken: Universiteit Utrecht/ CPS.

- Clark, C.M., & Peterson, P.L. (1986). Teachers' thought processes. In M.C. Wittrock (ed.). *Handbook of research on teaching*. Chicago: Rand McNally.
- Erp, M. van, & Tuyt, M. van (1981). *De basis van de nieuwe basisschool*. Den Haag: SVO/Ministerie van O&W.
- Erp, M. van, & Visch Eibergen, A. de (1983). *Op basis van de kleuterschool*. Amsterdam: SCO.
- Frankrijker, H. de, & Kler, H. de (1988). De OK in de nieuwe PABO: Integratie of annexatie? *Tijdschrift voor Lerarenopleiders*, 9, (5), 206-213.
- Gennip, H. van, & Ruys, A. (1992). *Aan de basis*. Nijmegen: ITS.
- Haan, L. de (1953). *Leidraad bij de methodiek en de didactiek in de kleuterschool*. Utrecht: Smits.
- Harskamp, E.G., Pijl, B., & Snippe, J. (1991). *Nascholing en spelstimulering in het primair onderwijs*. Groningen: RION.
- Harskamp, E., & T. Willemsen (1991). Programma's en speelleermaterialen voor voorbereidend rekenen in de basisschool. *Pedagogische Studietoeken*, 68, (9), 404-415.
- Harskamp, E., Edelenbos, P., Suhre, C., & Deinum, J. (1993). *Aanbod voor Nederlandse taal, Rekenen/wiskunde en Aardrijkskunde in het basisonderwijs. Technisch rapport ten behoeve van inspectie-onderzoek*. Groningen: RION.
- Inspectie van het Onderwijs (1994). *De toestand van het onderwijs aan jonge kinderen*. De Meern: Inspectie van het Onderwijs
- Jager, A. (1993). *Zorgverbreding in pedagogisch perspectief*. Dissertatie. Universiteit Utrecht, Faculteit Sociale Wetenschappen: ISOR.
- Janssen-Vos, F. (1987). *De brede basis*. Amsterdam: APS.
- Janssen-Vos, F. (1990). *Basisontwikkeling*. Assen: Van Gorcum.
- Janssen-Vos, F., B. Pompert & H. Vink (1990). *Kleuters in de Basisschool: Naar lezen, schrijven en rekenen*. Assen: Van Gorcum.
- Janssen-Vos, F. (1991). Van kleuteronderwijs naar basisonderwijs: wat zich in tien jaar afspeelde. *Het Jonge Kind* 18, 6.
- 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 (1985). *Ruimtelijke oriëntatie. Observatie- en hulpprogramma voor kleuters in de basisschool*. Arnhem: CITO.
- Kuyk, J.J. van (1987). Basisvorming in het kleuteronderwijs. In J.J. van Kuyk (red.). *Basisvorming in de basisschool*. Tilburg: Zwijzen.
- Kuyk, J.J. van (1992). Visie op onderwijs aan jonge kinderen en hulp aan kinderen met achterstanden. In B. van Oers & F. Janssen-Vos. *Visies op onderwijs aan jonge kinderen*. Assen: Van Gorcum.
- Laevers, F. (1992). Welbevinden en betrokkenheid. Richtsnoeren voor een ervaringsgerichte onderwijspraktijk. In B. van Oers & F. Janssen-Vos. *Visies op onderwijs aan jonge kinderen*. Assen: Van Gorcum.
- Lindt, M.W.J. (1987). De PABO: zorgverbreding en kleuteronderwijs, *ID, Tijdschrift voor lerarenopleiders*, 8, (4), 183-190.
- Lindt, M.W.J. (1988). De zorg van de PABO voor het onderwijs aan de kleuters op de basisschool. In *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs*, 4, (1), 2-15.
- Nijkamp, W.M. (1963). *Van beginschool naar kleuterschool*. Groningen: Wolters-Noordhoff.
- Nijkamp, W.M. (1966). *Ons kleuteronderwijs van gisteren naar morgen*. Groningen: Wolters-Noordhoff.
- Nijkamp, W.M. (1973). Honderd jaar kleuters. *De wereld van de kleuter*, nr.17. IJmuiden: Vermande & Zn.
- Oers, B. van & F. Janssen-Vos (red.) (1992). *Visies op onderwijs aan jonge kinderen*. Assen: Van Gorcum.
- Oers, B. van (1992). Onderwijsvisies en de vernieuwing van het onderwijs aan jonge kinderen: aanzet tot een discussiekader. In B. van Oers & F. Janssen-Vos. *Visies op onderwijs aan jonge kinderen*. Assen/Maastricht: Van Gorcum.
- Phillipi-Siewertz van Reesema, C. (1954). *Kleuterwereld - kleuterschool*. Den Haag: Wereldbibliotheek.
- Reezigt, G.J., Guldemon, H., & Ros, A.A. (1993). *Effectief onderwijs aan jonge kinderen*. Groningen: RION.
- Visitatiecommissie PABO (1993). *Bekwaamheid aan de basis*. Den Haag: HBO-Raad.
- Vroonhoven, K. van (1985). Van bewaarschool naar kleuterschool naar basisschool. In *De wereld van het jonge kind*, 6, 276-281.
- Wolbert, R.G.M., Schaap-Hummel, I., & Span, P. (1986). *Individualisering en differentiatie in de basisschool*. Den Haag: SVO.

Manuscript ontvangen 6-9-1994

Definitieve versie ontvangen 5-4-1995

Onderwijs aan jonge kinderen: historische ontwikkelingen

G.W. Meijnen

SCO-Kohnstamm Instituut, Universiteit van Amsterdam

ABSTRACT

In recent years the education of the (very) young child is very much in the public eye. The history of the education of the (very) young child shows that this type of education was not recognized as contributing to socialization, qualification and selection. In the past, two types of sources can be recognized: moral education and the upbringing of children whose parents both had to work. At this moment two societal factors are responsible for the growth of attention. First, there is the increasing participation of women with (very) young children on the labour market. Second, research findings show an overwhelming evidence of the backwardness in cognitive development of very young children from deprived families. Policy measures in order to prevent this kind of backwardness are carried out. Recently, the cooperation between primary schools (basisscholen) and infant schools (peuterspeelzalen) is stressed in order to cope with this problem.

INLEIDING

In de artikelen van Reezigt en Harskamp en Reezigt (dit nummer) wordt er op gewezen dat onderzoek naar aanbod en overdracht in wat tot 1985 het kleuteronderwijs was, schaars is. De eerste onderzoeken die door hen worden genoemd en die gebaseerd zijn op systematische dataverzameling, dateren van het begin van de tachtiger jaren (zie o.a. Van Erp en De Visch Eybergen, 1981). Een periode waarin de discussie over de wenselijkheid van de integratie van kleuter- en basisonderwijs nog in volle hevigheid woedde. Tot die tijd hielden universitaire vakgroepen dan wel onderwijsonderzoeksinstituten zich niet of nauwelijks bezig met het kleuteronderwijs. Uit een inventariserend onderzoek over de periode 1966-1974 blijkt bijvoorbeeld dat slechts 2,8% van al de projecten gericht is op dit onderwijstype (Verkenningscommissie Onderwijsresearch, 1976). Wanneer prioriteiten op de researchagenda een afspiegeling zijn van de publieke aandacht voor een maatschappelijk fenomeen (zie o.a. Harbers, 1986) dan zou een voorzichtig getrokken conclusie kunnen luiden dat het onderwijs aan jonge kinderen in het verleden niet is geassocieerd met de centrale functies van een onderwijssysteem: socialisatie, kwalificatie, selectie en allocatie.

In dit artikel zal in kort bestek een historische schets worden gegeven van de ontwikkeling van het onderwijs aan jonge kinderen in Nederland. Daarbij zal ook aandacht worden geschonken aan de maatschappelijke context. Uit deze schets is af te leiden dat de huidige discussies over inrichting van het onderwijs aan vier- tot zesjarigen het gevolg is van het besef dat onderwijs ook in deze fase al konsekwenties heeft voor de maatschappelijke functies van een onderwijssysteem. Afsluitend zal gewag worden gemaakt van recente beleidsvoornemens.

TOT 1800: INITIATIEVEN MET RELIGIEUZE ACHTERGROND

De vroegste gegevens over de schoolse opvoeding van het jonge kind in Nederland dateren van

Adres: Grote Bickerstraat 72, 1013 KS Amsterdam.

Dit is het laatste artikel behorende bij het themanummer.

de dertiende eeuw. In Breda werd in het jaar 1240 een instelling opgericht voor het geven van godsdienstonderwijs aan jonge kinderen. Het onderwijs werd gegeven door Rooms-Katholieke vrouwelijke religieuzen die de dominante levensbeschouwing van die tijd vertegenwoordigden. De opvoeding had tot doel kinderen op te voeden tot christelijk zedelijk leven. Daarbij werd weinig acht geslagen op de kinderlijke behoeften en ontwikkeling. Methoden die bruikbaar bleken te zijn voor ouderen zoals voor- en nazeggen van gebeden, psalmen en dergelijke werden ook gebruikt voor de instructie van het jonge kind (De Clercq, 1980).

Ten gevolge van de Reformatie in de eerste helft van de zestiende eeuw, waarbij het gezag van de Rooms-Katholieke kerk afbrokkelde, ontstonden er ook Protestantse schooltjes waar de kinderen kennismaakten met de beginselen van deze levensbeschouwing. Als de kinderen vijf jaar werden leerden ze in een aantal gevallen bovendien het alfabet en spellen. De huisvesting van de schooltjes was vaak slecht, de hygiënische omstandigheden lieten veel te wensen over, de kinderen werd geen enkele vorm van eigen initiatief toegestaan, de tucht was zeer streng en het aantal kinderen per (religieuze) leidster bedroeg vaak 50 of meer. Bovendien liet het intellectuele niveau van de leidsters veel te wensen over. De schooltjes werden niet bezocht door kinderen uit de hogere standen, zij werden opgevoed door de ouders en/of het huishoudelijk personeel.

In de zeventiende en achttiende eeuw was er in Nederland veel belangstelling voor het gedachtengoed van belangrijke pedagogen. Zo ontmoette Comenius (1592-1670) hier veel respons met als gevolg dat hij zich in Nederland vestigde. Ook is hij hier gestorven en begraven. De geschriften van Rousseau (1712-1778) vonden hier te lande eveneens veel aftrek. De praktijk van het onderwijs aan kleuters werd er echter nauwelijks door beïnvloed (De Clercq, 1980).

VAN 1800 TOT 1900: WORSTELING OM DE IDENTITEIT

De negentiende eeuw is staatkundig voor Nederland van beslissende betekenis geweest. Het werd een monarchie met een parlementaire democratie. Gegeven de dominantie van het Protestantisme ontstond er discussie over de vraag of openbaar onderwijs, gesubsidieerd door de overheid, een protestantse signatuur moest dragen. Tegen het einde van deze eeuw was deze 'schoolstrijd' grotendeels gestreden. De oplossing werd gevonden in de formule dat de overheid moest zorgen voor voldoende voorzieningen van openbaar onderwijs, dat qua levensbeschouwing neutraal diende te zijn, en dat bijzondere scholen, mits zij voldeden aan een aantal voorschriften, eveneens volledig door de overheid zouden worden betaald. Deze discussie evenals de daaruit voortvloeiende beleidsmaatregelen, hadden echter niet of nauwelijks betrekking op het onderwijs aan (zeer) jonge kinderen. Wet- en regelgeving met betrekking tot dit type onderwijs komt in Nederland in deze eeuw nauwelijks voor (Idenburg, 1964).

Sociaal-economisch gezien werd de negentiende eeuw gekarakteriseerd door de opkomende industrie. Fabrieksarbeid en zeer lage lonen veroorzaakten beide ouders evenals iets oudere kinderen tot lange werkdagen. Veel kleine kinderen werden daardoor aan hun lot overgelaten. De bourgeoisie trok zich het lot aan van de grote aantallen kinderen uit de arbeidersklasse die nauwelijks werden verzorgd en gevoed en geen onderwijs ontvingen. De vroegere 'klein-kinder' schooltjes verloren het karakter van instituten voor uitsluitend godsdienstonderwijs en werden naar hun aard meer en meer liefdadigheidsinstellingen, een tak van armenzorg. De bourgeoisie veronderstelde dat door het verbeteren van het onderwijs de armoede zou kunnen worden bestreden en de voedingsbodem voor revolutionaire tendenzen en toenemende criminaliteit binnen de arbeidersklasse zou doen verminderen (Dodde, 1983).

Op tal van plaatsen worden in aanvulling op de religieuze instellingen, schooltjes opgericht door charitatieve verenigingen of wel door particulieren met commerciële motieven. Omdat de bijdrage van de ouders afkomstig uit het proletariaat slechts zeer gering kon zijn, waren de omstandigheden in de commerciële schooltjes vaak erbarmelijk: nauwelijks daglicht, vrijwel geen sanitaire voorzieningen, geen verwarming, grote aantallen kinderen (tot 100) in kleine ruimtes en oude leidsters die niet meer in staat waren tot enige andere arbeid en daarom dit laaggewaardeerde en laag betaalde werk deden.

Weliswaar waren reeds vanaf 1806 alle 'klein-kinder' schooltjes onder toezicht geplaatst van inspecteurs en werd van de begeleidsters verwacht dat ze konden lezen en schrijven, maar veel effect had deze maatregel niet vanwege het ontbreken van sancties. Hoewel precieze cijfers ontbreken, bezocht naar schatting in het begin van de negentiende eeuw ongeveer 25% van de zeer jonge kinderen een 'klein-kinder' school. Kinderen uit de midden en hogere milieus behoorden daar nauwelijks toe, zij bleven meestal thuis en werden vaak mede opgevoed door het in dienst zijnde huishoudelijk personeel (Dodde, 1983). Vanaf 1850 keert de situatie enigszins ten goede. Onder auspiciën van de Maatschappij tot Nut van het Algemeen, een organisatie van de gegoede burgerij met als doel bij te dragen aan de verheffing van het volk, worden 'bewaarschooltjes' opgericht waar kinderen werden gevoed en gekleed en waar de ideeën van Pestalozzi tot uitvoering werden gebracht. Er werd aandacht besteed aan de oefening van zowel de verstandelijke als motorische ontwikkeling van de kinderen. Er werden bovendien opleidingen gestart voor de functie van bewaarschoolhouderes.

Vanaf 1830 stimuleert de centrale overheid het stichten van bewaarscholen voor kinderen tot zes jaar door gemeenten. Wanneer het particulier initiatief tekort schiet dienen de gemeenten voor het stichten van voldoende scholen te zorgen (Dodde, 1983). De gemeenten nemen deze taak niet erg serieus en pas aan het einde van de negentiende eeuw ontstaan er veel openbare, door de gemeenten bekostigde bewaarscholen. Vanaf 1840 komen er ook scholen voor kinderen uit de hogere milieus en ontstaan er 'burgerbewaarscholen' naast 'armenbewaarscholen'. Van de totale populatie kinderen in de kleuterperiode bezoekt ook in 1860 nog maar slechts ca 25% een bewaarschool, hetzelfde percentage als veertig jaar daarvoor (Dodde, 1983).

Zoals in een eerdere fase opvoedkundige ideeën in Nederland sterk werden ontleend aan buitenlandse pedagogen (Comenius, Rousseau), was dat in de tweede helft van de negentiende eeuw eveneens het geval en nu in de persoon van Fröbel. Zijn ideeën werden, in tegenstelling met die van zijn voorgangers, echter wel geadopteerd door de praktijk. Omstreeks 1900, wanneer ongeveer 50% van de kleuters een school bezoekt, wordt in hoofdzaak de methode van Fröbel toegepast al dan niet vermengd met ideeën van Herbart. De methodiek van Fröbel werd zelfs zo dominant dat men niet meer van bewaarscholen maar van Fröbelscholen spreekt. De naam voor het type werd identiek met de naam voor de meest gepraktiseerde methode.

Samenvattend: gedurende deze eeuw heeft het kleine-kinderschooltje haar karakter van instituut voor hoofdzakelijk godsdienstonderwijs verwisseld voor dat van liefdadigheidsinstelling op grond van sociale noodzaak om vervolgens het aanzien te krijgen van een echte onderwijsinstelling.

VAN 1900 TOT 1985: HET KLEUTERONDERWIJS KOMT TOT BLOEI

De invloed van buitenlandse pedagogen bleef bij lange na niet beperkt tot het gedachtengoed van Fröbel. In 1912 werd Maria Montessori uitgenodigd een lezing te houden in Nederland. Haar ideeën vonden ruime weerklank: in 1914 werd een eerste Montessori-klas in Den Haag geopend en in 1917 volgde de oprichting van de Nederlandse Montessori-vereniging. Daarna startte in 1918 een opleidingscursus voor Montessori-onderwijs.

Het enthousiasme waarmee haar ideeën in Nederland werden geadopteerd deden Maria Montessori zelfs besluiten zich een poos in Nederland te vestigen. De snelle groei van de Montessori-kleuterscholen werd vooral veroorzaakt door de adoptie van deze opvoedingsideologie door de hogere sociale milieus. Door de ruimere financiële bijdragen van de ouders staken deze scholen door zowel uitrusting als aanpak en professionaliteit van de leidsters, gunstig af bij Fröbelscholen. Het feit dat de Fröbel- en de Montessori-scholen zeer verschilden in materiële condities, professionele kwalificaties van de leidsters en sociale herkomst van de leerlingen, waren er de oorzaak van dat de discussie over welke methodiek inhoudelijk nu de beste was, zeer lang stand kon houden (Verbeeten, 1937).

Bovendien ontstonden er qua methodiek tal van mengvormen. Voor het uitbreken van de Tweede Wereldoorlog in 1940 was er dan ook sprake van een grote variatie in onderwijskundige

aanpak: (1) orthodoxe en vrijere Fröbelscholen, (2) orthodoxe en vrijere Montessori-scholen en (3) bewaarschooltjes zonder speciale methodiek vanwege het ontbreken van gediplomeerd personeel. Tijdens de economische crises in de dertiger jaren werden tal van openbare, door de gemeenten bekostigde bewaarscholen gesloten. Omdat het beroep van leidster nog steeds niet beschermd was, combineerden veel van hen het onderwijs aan jonge kinderen met andere activiteiten zoals het geven van naai- en handwerkles aan oudere meisjes.

Na 1945, het einde van de Tweede Wereldoorlog, werd de herinrichting van het onderwijs in Nederland met kracht ter hand genomen. Het onderwijs aan kinderen tot zes jaar kreeg echter niet de hoogste prioriteit. Pas in 1956 werd een wet op het kleuteronderwijs aangenomen waarin werd bepaald dat al de scholen, zowel openbare, die gesubsidieerd werden door de gemeente, als particuliere, mits zij voldeden aan bepaalde kwaliteitseisen, in de toekomst zouden worden bekostigd door de rijksoverheid. Daarmee werd de bekostiging identiek aan de regeling die reeds in 1920 was getroffen voor het overige, reguliere onderwijs. Het feit dat het onderwijs aan jonge kinderen aan het eind van de vorige eeuw en het begin van deze, geen onderdeel vormde van de schoolstrijd, geeft aan dat het maatschappelijk niet in hoog aanzien stond: te weinig in ieder geval om het voorwerp van politieke strijd te maken. In de genoemde wet wordt de minimum toelatingsleeftijd op vier jaar gesteld en verordonneerd dat de leidsters gediplomeerd dienen te zijn en dat er een 'speel-en werkplan' moet worden gemaakt. Intussen is de participatiegraad van de kleuters gegroeid tot ongeveer 80% in 1960.

In de jaren zeventig komt in Nederland de discussie op gang over de aansluiting van het kleuteronderwijs en het lager onderwijs. Door onder andere de invloed van de Engelse 'Nursery School Association' en de zich ook in Nederland verbreidende ideeën van Dewey en Decroly (Nijkamp, 1974), was in de kleuterschool het vrije spel en de vrije expressie steeds meer in het middelpunt van de belangstelling komen te staan. Het kleuteronderwijs werd daardoor meer en meer 'child-centered' met veel ruimte voor eigen initiatieven van het kind en met veel aandacht voor spelonderwijs opdoen van ervaringen. De kleuterschool als voorbereiding op het lager onderwijs verdween enigszins achter de horizon. De kloof tussen het kleuteronderwijs en het lager onderwijs werd qua inhoud en overdrachtsvormen aanzienlijk. Tal van pogingen werden ondernomen om de aansluiting beter te laten verlopen. Zo werd op vele lagere scholen het eerste leerjaar ingericht als een 'speelleerklas'.

Uiteindelijk is in 1985 gekozen voor een rigoureuze oplossing: de integratie van kleuter- en lager onderwijs tot één onderwijstype. In de Wet op het Basisonderwijs van dat jaar wordt de leerweg van vier tot twaalf jaar wettelijk vorm gegeven.

INTEGRATIE KLEUTER- EN BASISONDERWIJS: 1985 - HEDEN

De integratie van het kleuter- en lager onderwijs kan worden geïnterpreteerd als een uitvloeisel van een toenemende maatschappelijke bewustwording van de relevantie van de vroegkinderlijke socialisatie voor de latere levensloop. Ontwikkelingsmogelijkheden blijken voor een belangrijk deel al op zesjarige leeftijd uitgekristalliseerd te zijn. Het uit de onderwijsresearch bekende fenomeen dat prestaties op een vroeger tijdstip veruit de belangrijkste voorspellers zijn voor de actuele prestaties is daar een bewijs van. Een keten die op basis van beschikbaar onderzoeksmateriaal te reconstrueren is tot het tijdstip van de intrede in de voormalige lagere school. De voormalige lagere scholen transporteerden als het ware intellectuele aanvangsverschillen tussen kinderen (Slavenburg en Peters (red.), 1989).

Hoewel het verschijnsel niet dramatisch is qua effectgrootte, blijkt het genoemde transport zelfs gepaard te gaan met divergerende ontwikkelingslijnen: de kloof tussen de goed- en slecht presterenden wordt tijdens de lagere schoolperiode groter. Dit zogenaamde Matthéüs-effect (Matthéüs 25, vers 29) houdt in dat goede leerlingen meer profiteren van een uniform curriculumaanbod dan zwakke leerlingen. Het is dan ook niet verwonderlijk dat weinig scholen er in slagen aanvangsverschillen te nivelleren (zie o.a. Brandsma, 1993; Van der Werf, Nitert, Reezi, 1994). Het 'compenserend vermogen' van scholen blijkt niet groot te zijn. Gegeven deze empi-

rische bevindingen zijn er onderzoekers die pleiten voor een maximale inspanning met betrekking tot het voorkomen van verschillen prestatieontwikkeling bij jonge leerlingen (Slavin, e.a., 1994). Wanneer aanvangsverschillen ook nog sociaal-economische en etnische achtergronden hebben, stijgt het functioneren van het voormalige kleuteronderwijs op de politieke en beleidsagenda.

Van der Velden (1991) heeft op basis van een cohort leerlingen dat gevolgd is van het eerste leerjaar in het lager onderwijs tot aan de bereikte onderwijspositie op 18-jarige leeftijd, berekend dat van de verschillen tussen SES-groepen in gemiddeld bereikt onderwijsniveau, 40% is toe te schrijven aan de leeftijdsfase tot 6 jaar, 30% aan de fase tot de intrede in het voortgezet onderwijs en 30% aan de periode in het voortgezet onderwijs. Hoewel door het ontbreken van een dergelijk cohort met voldoende allochtone leerlingen identieke berekeningen voor deze categorie nog niet zijn te maken, blijkt uit tal van onderzoeksverslagen dat zich ook hier problemen voordoen in de linguïstische en cognitieve sfeer en dat reeds op jonge leeftijd. (zie o.a. Meijnen en Riemersma, 1992). Het artikel van Verhoeven en Narain (dit nummer) dat de ontwikkeling van tweetaligheid bij allochtone kleuters tot onderwerp heeft, sluit hierbij aan.

Maatschappelijk gezien blijkt de kleuterfase en dus het onderwijs aan vier- tot zesjarigen van meer belang te zijn dan algemeen werd beseft. De jarenlange afwezigheid van het besef dat de socialisatie-, kwalificatie en selectiefuncties van het onderwijs wel degelijk hun wortels hebben in het onderwijs aan jonge kinderen, creëerde voor het kleuteronderwijs de ruimte om doelstellingen te formuleren die niet direct met deze functies verbonden zijn.

Met de integratie werd beoogd de leerweg van 4 tot 12 binnen één systeem onder te brengen. In onderwijskundige termen: het creëren van een ononderbroken leerweg van 4 tot 12. Opvallend is dat de wetgever de daarmee corresponderende leerplicht niet volledig heeft doorgevoerd: deze geldt vanaf de vijfjarige leeftijd. Vanuit het perspectief van een efficiënte organisatie van overdracht en inhoud is dit niet een voor de hand liggende beslissing noch vanuit het oogpunt van de bestrijding van onderwijsachterstanden. Andere overwegingen zoals de verhouding tussen rechten en plichten van ouders versus de verantwoordelijkheid van de samenleving voor de opvoeding en opleiding van de jongste generatie, hebben tot deze beslissing geleid.

Uit het artikel van Reezigt (dit nummer) en uit het rapport van de Commissie Evaluatie Basisonderwijs (1994) blijkt dat er aan de uitwerking van het concept van de ononderbroken leerweg in de praktijk het een en ander schort. De CEB heeft tal van aanbevelingen geformuleerd om de realisering ervan toch dichterbij te brengen. Aanbevelingen die deductief voortvloeien uit de stelling dat zoveel mogelijk leerlingen het basisonderwijs dienen af te sluiten met een zo hoog mogelijke beheersing van kennis en vaardigheden met betrekking tot de kerndoelen. De argumentatie daarvoor wordt ontleend aan zowel de steeds hogere eisen die gegenereerd worden door maatschappelijke ontwikkelingen als aan het streven van het bestrijden van onderwijsachterstanden. In deze evaluatie maakt het onderwijs aan kleuters integraal onderdeel uit van het gekozen doelstellingsperspectief en is er geen plaats meer voor een, maatschappelijk gezien, contextloze invulling van het onderwijs.

Onderzoek waarin de effecten van instructie, differentiatie en materiaal in het onderwijs aan kleuters centraal staan, komt in de tweede helft van de tachtiger jaren op gang. Beproefde concepten uit research met een lange traditie in het 'reguliere' onderwijs. Het overzicht van onderzoeksresultaten door Reezigt (dit nummer) is dan ook gestructureerd volgens deze systematiek. Het CEB-perspectief is echter niet onomstreden. Harskamp en Reezigt (dit nummer) wijzen er op dat de laatste jaren verschillende, min of meer concurrerende visies zijn ontwikkeld over het gewenste onderwijsaanbod aan jonge kinderen en, in het verlengde daarvan, over gewenste opbrengsten. Stevens en Harskamp (dit nummer) doen verslag van een survey onder 40 kleuterleerkrachten naar hun opvattingen over doelstellingen en onderwijsaanbod. Zij komen tot de conclusie dat het overgrote deel van de leerkrachten (80%) een kindgerichte onderwijsvisie heeft. De visies zijn in de praktijk dus nogal scheef verdeeld. Het artikel van Van Kuyk en Verhoeven (dit nummer) is een voorbeeld van hoe de effecten van variaties in onderwijsaanbod uitwerken op de geletterdheid bij kleuters.

Op de variaties in visies wordt later ingegaan bij het bespreken van nieuwste ontwikkelingen.

Hier wordt volstaan met de constatering dat er kennelijk iets ontstaat wat in het lager onderwijs al lang gangbaar was: verschillende schoolconcepten. Het nog steeds groeiende aantal 'traditionele' vernieuwingsscholen en het nog in grote getale voorkomen van traditionele, klassikale scholen duiden erop dat er van een uniform schoolconcept in Nederland nog geen sprake is. De hoop mag worden uitgesproken dat er geen dubbele scheidslijnen gaan ontstaan: schoolconcepten voor de kleuterfase en, onafhankelijk daarvan, schoolconcepten voor 6 tot 12 jarigen. De legitimering voor variaties in schoolconcepten tussen scholen is om tal van redenen wellicht te verdedigen, die voor binnen scholen daarentegen niet.

EEN VOLGENDE INTEGRATIE?

Internationaal gezien loopt Nederland met de integratie van kleuter- en lager onderwijs voorop. In verreweg de meeste landen begint het 'reguliere' onderwijs op 5 of 6 jarige leeftijd. Onder 'Preschool education' wordt internationaal dan ook verstaan het net van voorzieningen voor de opvang van kinderen tot aan die leeftijd. Een net met veelal een aanzienlijk hogere dekking voor 2 tot 4 jarigen dan in Nederland het geval is. Daarmee is in Nederland de internationaal unieke situatie ontstaan dat het 'reguliere' onderwijs op zeer jonge leeftijd begint en voor bijna 100% dekkend is en dat educatieve voorzieningen voor de allerjongsten relatief schaars zijn.

Twee maatschappelijke ontwikkelingen versterken echter het streven naar institutionele bemoeienis met de opvoeding van (zeer) jonge kinderen, zowel waar het de inhoud als de vervroeging van de arrangementen betreft. Als eerste kan worden genoemd het beleid gericht op het vergroten van de mogelijkheden van kinderopvang in het kader van het verbeteren van de arbeidsparticipatie van vrouwen met jonge kinderen. Door maatregelen in de belasting- en subsidiesfeer is het aantal kinderen dat in kinderdagverblijven wordt opgevangen, drastisch toegenomen. Bijgevolg komt meer en meer de vraag omtrent de kwaliteit van de opvang aan de orde (zie o.a. Singer, 1993). Daartoe horen uiteraard ook doelstellings- en methodiek vraagstukken. De roep om versterking van de professionele kwalificaties van de (bege)leiding wordt in dit kader vaak gehoord.

Een tweede impuls voor de institutionele bemoeienis met de opvoeding van (zeer) jonge kinderen komt uit de sfeer van het reeds eerder genoemde Onderwijsvoorrangsbeleid. Ook op 4 jarige leeftijd zijn reeds verschillen in cognitieve en linguïstische ontwikkeling waar te nemen samenhangend met de sociaal-economische dan wel met etnische achtergrond (zie o.a. Meijnen en Riemersma, 1992). Als remedie worden stimuleringsactiviteiten ontworpen die respectievelijk 'home based' dan wel 'center based' zijn of een combinatie van beide vormen inhouden (zie o.a. Van Loggem en Bekkers (red.) 1994; Eldering en Choenni-Gobardhan, 1990; Rispens en Van der Meulen (red.) 1992).

Het Ministerie van VWS heeft de laatste jaren in dit kader zeer veel activiteiten ontplooid, in het bijzonder gericht op de thuissituatie van (zeer) jonge allochtone leerlingen. Programma's als Instapje, Opstapje, Opstap, Overstap, Klimrek, Spelvoorlichting aan Huis zijn daar voorbeelden van (Rispens en van der Meulen (red.), 1992). Het Ministerie van VWS subsidieert de programmaontwikkeling, de evaluatie en de voorlichting. Gemeenten dan wel welzijnsinstellingen kunnen de programma's vervolgens adopteren. Recentelijk heeft het Ministerie ook oog gekregen voor de ontwikkeling van centrumgerichte programma's. Binnen het scala aan voorzieningen dat reeds in de voorschoolse periode opereert, lijken de peuterspeelzalen, gelet op hun historie van doelgerichte pedagogische aanpak, naar het oordeel van het Ministerie, het meest in aanmerking te komen om een rol te vervullen in het bestrijden van ontwikkelingsachterstanden van met name allochtone leerlingen. Deze werksoort zou in het bijzonder bij het nieuw te ontwikkelen beleid moeten worden betrokken.

Ook op het Ministerie van Onderwijs, Cultuur en Wetenschappen is de aandacht voor het (zeer) jonge kind binnen het Onderwijsvoorrangsbeleid groeiende. In zijn reactie op het rapport 'Ceders in de Tuin' heeft de toenmalige staatssecretaris Wallage besloten extra middelen vrij te

maken ter bevordering van maatregelen die de toeleiding van kinderen uit achterstandssituaties naar de basisschool bevorderen (Ministerie van Onderwijs en Wetenschappen, 1993). In deze reactie wordt aangekondigd dat een en ander zal moeten worden afgestemd met het, toenmalige, Ministerie van WVC.

De toenmalige bewindslieden van WVC (D'Ancona) en O. en W. (Wallage) hebben in het voorjaar van 1993 de Commissie (Voor-)schoolse Educatie gevraagd advies uit te brengen over het toeleidingsvraagstuk. Het eindrapport is opgeleverd in januari 1994 (Leseman en Cordus, 1994). De kern van het advies omvat het inrichten van een aantal experimenten met samenwerkingsverbanden tussen peuterspeelzalen en basisscholen. Gepleit wordt voor het implementeren van zorgvuldig opgezette programma's voor 3 tot 6 jarigen.

Er is niet gekozen voor de uitwerking van een enkel programma. Een voorlopige inventarisatie leerde namelijk dat er in binnen- en buitenland tal van centrumgerichte programma's zijn ontworpen die soms zijn bestemd voor de kleuterleeftijd en soms voor nog jongere kinderen. De Commissie heeft vervolgens besloten alleen die programma's door te lichten die al een zekere wetenschappelijke en/of, in Nederland, maatschappelijke status hadden verworven. Dat zijn de volgende: (1) 'Recht is recht, krom is krom' van de Stichting Leerplan Ontwikkeling (SLO), (2) 'Basisontwikkeling' van het Algemeen Pedagogisch Studiecentrum (APS), (3) 'Ervaringsgericht werken met peuters en kleuters' van het Katholiek Pedagogisch Centrum (KPC), (4) 'Ontwikkelingsgericht programma voor 3 tot 6 jarigen' van het Centrum voor Educatieve Dienstverlening (CED), (5) het 'Van Kuyk/Slavin programma' van het Centraal Instituut voor Toetsontwikkeling (CITO) en (6) het 'High/Scope programma' van David Weikart uit de Verenigde Staten. Deze programma's zijn beoordeeld op hun wetenschappelijke verankering, op de relatie tussen programmadoelen, transferdoelen en de OVB-doelstelling, op de gepropageerde regulering van leer- en ontwikkelingsprocessen, op de directheid van de voorgestane interventie en bewaking van de programma-implementatie en, tenslotte, op de geschiktheid voor peutergroepen.

De conclusie van de Commissie luidt (p. 53): 'De voorgestelde programma's verschillen onderling qua concept, aanpak, opbouw en uitwerking nogal sterk, waardoor geen van de voorstellen op alle dimensies zonder meer als kwalitatief de beste of voor onderwijsvoorrandsgdoelen meest geschikte naar voren kwam'. Uiteraard verschilden de programma's wel in de mate van getal en niveau waarop aan bovengenoemde criteria werd voldaan. De Commissie besluit dan tot de volgende rangorde: (a) High/Scope en Van Kuyk/Slavin, (b) APS en KPC en (c) SLO en CED. Voorgesteld wordt te starten met een viertal experimenten uit de categorieën a en b. Per programma zouden dan vier koppels van peuterspeelzalen en basisscholen bereid moeten worden gevonden het programma gedurende ten minste drie jaar te adopteren. In totaal zouden dan 16 koppels aan de vier experimentele condities meedoen. Evaluatieonderzoek zou moeten aantonen welk programma het meest geschikt is voor de beoogde doelgroepen. De beide ministeries hebben vervolgens maatregelen getroffen die het starten van activiteiten in deze richting mogelijk maken.

DISCUSSIE

Uit een historische beschrijving van de institutionele zorg voor kleuters valt op te maken dat tot aan het begin van deze eeuw twee motieven een belangrijke rol speelden: religieuze en morele opvoeding en opvang voor het geval ouders daartoe niet in staat waren. Na de eeuwwisseling, wordt het gedachtengoed van met name buitenlandse pedagogen geïntroduceerd om het onderwijs aan kleuters vorm te geven. Onderwijs dat de gaven van 'hoofd, hart en hand' dient te stimuleren. De maatschappelijke relevantie wordt echter niet hoog ingeschat want deze onderwijsvorm wordt niet in de schoolstrijd betrokken en ook regelgeving met betrekking tot bijvoorbeeld de kwalificaties van de leidsters blijft lange tijd achterwege. Eerst in de tachtiger jaren van deze eeuw wordt kleuteronderwijs ook 'regulier' onderwijs.

Hoewel er nog steeds geluiden te horen zijn dat het ongedaan maken van de integratie wijs

beleid zou zijn, lijkt de keuze voor de basisschool voor 4 tot 12 jarigen een historische trend te bevestigen van een groeiende, maatschappelijke bemoeienis met opvoeding en opleiding van (zeer) jonge kinderen. Een trend die wordt bevestigd door basisscholen die deze weg al op zijn gegaan en al een peuterspeelzaal 'in huis hebben' hebben. Recent ingezet overheidsbeleid dat profilering van en concurrentie tussen basisscholen bevordert, zorgt voor een krachtige impuls met betrekking tot deze ontwikkeling.

LITERATUUR

- Brandsma, H.P. (1993). *Basisschoolkenmerken en de kwaliteit van het onderwijs*. Groningen: RION.
- Clercq, I. de (1980). *Begijne-, bewaar-, Frobel-, Kleuterschool*. Universiteit van Amsterdam.
- Commissie Evaluatie Basisonderwijs (1994). *Onderwijs aan jonge kinderen*. Den Haag: SDU.
- Dodde, N.L. (1983). *Het Nederlandse onderwijs verandert*. Muiderberg: Coutinho.
- Eldering, L., & Choeni-Gobardhan, S. (1990). *Interventieprogramma's voor jonge kinderen uit kansarme groepen*. Leiden: DSWO PRESS.
- Erp, M. van, & Visch Eybergen, A. de (1983). *Op basis van de kleuterschool*. Amsterdam: SCO.
- Harbers, H. (1986). *Sociale wetenschappen en hun speelruimte*. Groningen: Wolters-Noordhoff.
- Idenburg, Ph. J. (1964). *Schets van het Nederlands schoolwezen*. Groningen: Wolters.
- Leseman, P.P.M., & Cordus, J. (1994). *(Allochtone) kleuters meer aandacht*. Rijswijk: Ministerie van WVC.
- Loggem, D., van & Bekkers, B. (red.) (1994). *Het jonge kind tussen onderwijs en welzijn*. Amsterdam: Averroës Stichting.
- Meijnen, G.W., & Riemersma, F.S.J. (1992). *Schoolcarrières een klassenkwestie? De schoolloopbanen van leerlingen gezien in relatie tot de invloeden die maatschappelijke en binnenschoolse determinanten hierop hebben*. Amsterdam/Lisse: Swets en Zeitlinger.
- Ministerie van Onderwijs en Wetenschappen (1993). *Naar meer schoolsucces. Kabinetsreactie bij het advies 'Ceders in de tuin'*. Den Haag: Ministerie van O en W.
- Nijkamp, W.M. (1974). *100 jaar kleuters*. IJmuiden:
- Rispens, J., & B.F. van der Meulen (red.) (1992). *Gezinsgerichte stimulering van kinderen uit achterstands-situaties*. Amsterdam/Lisse: Swets en Zeitlinger.
- Singer, E. (1993). *Kinderopvang, goed of slecht?* Utrecht: Uitgeverij SWP.
- Slavenburg, J.H. & Peters, T.A. (red.) (1989). *Het project onderwijs en sociaal milieu: een eindbalans*. Rotterdam: SAD.
- Slavin, R.E., Karweit, N.L., Wasik, B.A., (1994). *Preventing early school failure*. Boston: Allyn and Bacon.
- Velden, R.K.W. van der (1991). *Sociale herkomst en schoolsucces*. Groningen: RION.
- Verbeeten, M. (1937). *Bijdrage tot de geschiedenis van het onderwijs van het jonge kind*. 's-Hertogenbosch: Verkenningcommissie Onderwijsresearch (1976). *Onderwijsonderzoek in Nederland*. Den Haag: Staatsuitgeverij.
- Werf, M.P.C. van der, Nitert, E.H.M. & Reezigt, G.J. (1994). *Effectieve en minder effectieve basisscholen voor allochtone leerlingen*. Groningen: GION.

Manuscript ontvangen 6-9-1994

Definitieve versie ontvangen 5-4-1993

Notities en Commentaren

REKENINHOUDELIJKE KRITIEK OP CEB-REPLIEK

A. Treffers

Freudenthal Instituut

Naar aanleiding van het themanummer over de evaluatie van het basisonderwijs, TOR 20,1 (1995), wil ik graag het volgende opmerken. De Commissie Evaluatie Basisonderwijs (CEB) is er vanuit gegaan dat de PPON-opgaven in principe door de leerlingen goed gemaakt moeten kunnen worden – de beheersingshypothese. Tevens is de CEB van mening dat de PPON-opgaven een goede afspiegeling vormen van de mogelijke items die uit de domeinbeschrijvingen kunnen worden afgeleid – de afspiegelingshypothese.

Op grond van deze twee aannames is de normering vastgesteld voor voldoende beheersing (gemiddeld minimaal 70 procent goed-score), matige beheersing (goed-score tussen 55 en 70 procent) en onvoldoende beheersing (minder dan 55 procent) van een (sub)domein.

Wijnstra (1995) betwist de juistheid van beide vooronderstellingen en de daaruit afgeleide normering:

‘Een dergelijke normering veronderstelt dat de gebruikte itemverzamelingen op de een of andere manier als representatief kunnen worden beschouwd voor de inhoud van het vigerende onderwijs en dat de normen passen bij de nagestreefde beheersing van deze stof. Deze veronderstelling is in elk geval niet te ontleen aan PPON, hoewel de commissie schrijft: De PPON-toetsen zijn zó geconstrueerd, dat leerlingen in principe de opgaven goed moeten kunnen maken. (Deelrapport 1, p. 55).

Deze zin is eenvoudig onjuist en in tegenspraak met de uitgangspunten bij de toetsconstructie (...).’ (p. 29)

Dat de leerlingen de PPON-toetsopgaven goed moeten kunnen maken, is irreëel. Vergelijking met de goedscores bij corresponderende internationale items toont aan dat de beheersingshypothese inderdaad iedere rationele en empirische grond mist. Dat ook de afspiegelingshypothese onjuist is, blijkt uit het simpele feit dat de PPON-opgaven voor het overgrote deel uit toepassingen bestaan, terwijl de gebruikte leerboeken in groep 8 anno 1987 voornamelijk een traditionele signatuur hadden en voor het overgrote deel kale rekenopgaven stelden.

Kortom, de basis van het CEB-onderzoek is ondeugdelijk.’

Ik wil een stap verder gaan met te stellen dat de PPON de pretentie van representativiteit ook nooit zou kunnen waarmaken en dat de CEB-aanname van afspiegeling niet alleen feitelijk maar ook principieel onjuist is.

Men dient namelijk bij de operationalisering van leerdoelen in dekkende opgaven onderscheid te maken tussen gesloten en open verzamelingen van corresponderende opgaven. Bij gesloten verzamelingen is het zinvol om over afspiegelingen in de vorm van een beperkt aantal items te spreken die representatief zijn voor de totale verzameling. Maar bij open verzamelingen is zulks niet mogelijk.

Het verschil tussen open en gesloten, in dit verband, zal ik illustreren met het complexe domein breuken.

Indien de doelstelling luidt dat leerlingen ‘kale’ breuken moeten kunnen vergelijken, is de verzameling opgaven om beheersing van dit leerdoel te meten, welomschreven van aard en gesloten van karakter. Of anders gezegd: men kan een serie van opgaven selecteren die ‘psychometrisch een afspiegeling vormen van de domeinen aan mogelijke items die men zich bij de domeinbeschrijvingen in kan denken’. (Van der Linden & Zwarts, 1995b, p. 36)

Maar bij toepassingsopgaven met breuken kan de verzameling dekkende items vrijwel onbeperkt worden uitgebreid en is het ondoenlijk om met een relatief beperkte set van enkele tientallen vraagstukken de totaliteit van mogelijke items te representeren. Dit blijkt bijvoorbeeld uit het feit dat ook elementaire opgaven gesteld kunnen worden die voor hooggeschoolde lezers van dit tijdschrift niet zo makkelijk zijn op te lossen.

Neem het leerdoel: 'via (gelijkwaardige) breuken deel-geheel relaties kunnen beschrijven en vergelijken'.

Een voorbeeld van een opgave die bij dit leerdoel past:

Op een feest danst op een gegeven moment $\frac{3}{4}$ deel van de vrouwen met $\frac{5}{6}$ deel van de mannen (er zijn dan alleen gemengde paren op de vloer).

Enkele vragen met een sterk variërende moeilijkheidsgraad:

- 1 Wat is het grootste deel dat danst: van de vrouwen of van de mannen?
- 2 Zijn er meer vrouwen of mannen op het feest?
- 3 Welk deel van de aanwezigen danst op dat moment?
- 4 Toon aan dat het gevraagde deel uit vraag 3 tussen de gegeven delen ($\frac{3}{4}$ en $\frac{5}{6}$) moet liggen en dat dit ook bij andere gegeven breuken zo moet zijn.

Vraag (1) is een opgave die direct naar het ordenen van breuken verwijst: $\frac{5}{6}$ is groter dan $\frac{3}{4}$, dus er zijn relatief meer dansende mannen op dat moment.

Vraag (2) betreft ook het vergelijken van breuken, maar de conclusie dat $\frac{5}{6}$ groter is dan $\frac{3}{4}$ voert tot de 'omgekeerde' slotsom dat er meer vrouwen dan mannen op het feest zijn.

Vraag (3) suggereert optellen van breuken (en vervolgens wellicht het halveren van de uitkomst), maar bij nadere beschouwing biedt deze basisoperatie geen oplossing. Sterker: het merendeel van hooggeschoolde volwassenen is niet in staat om dit probleem binnen een redelijke tijd van 5 à 10 minuten correct op te lossen, terwijl het achteraf gezien toch tamelijk elementair is ...

Vraag (4) is zo mogelijk nog complexer.

Moraal: toepassingen in een complex rekendomein zijn zo divers en dermate wisselend van moeilijkheidsgraad dat een tiental items geen adequate afspiegeling van alle mogelijke toepassingsopgaven kan geven – en dat is zeker het geval voor structureel rijke terreinen als die van (decimale) breuken, verhoudingen en procenten plus overlappende gebieden. Laat staan dat er een eenvoudige psychometrische procedure zou zijn om uit zo'n open verzameling een representatieve steekproef van te beheersen items te kiezen.

Of anders gezegd: het is arbitrair of men de opgaven 2, 3 en 4 als representatief beschouwt voor de leerstof die bij de genoemde doelstelling over 'deel-geheel relaties beschrijven en vergelijken' hoort – om deze doelstelling en deze opgaven over breuken maar als metafoor te gebruiken. Volstrekt onduidelijk is welke criteria de CEB bij de keuze heeft gebruikt en welke procedure van deskundigen-consultatie is gevolgd. (Dat wil zeggen: informeel is mij hierover wel wat bekend, maar het zou fair zijn als de CEB eens precies uit de doeken zou doen hoe een en ander bij rekenen-wiskunde precies in z'n werk is gegaan – dat zou heel onthullend kunnen zijn.) Maar één ding lijkt me wel duidelijk: een absolute niveauwaardering ontleen aan de goedscores van een serie opgaven die als partiële operationalisering van doelstellingen zijn op te vatten, is een uiterst hachelijke operatie die veel prudentie en deskundigheid vereist. Naar mijn mening is de CEB in dit opzicht niet voldoende omzichtig te werk gegaan.

Men kan zich overigens afvragen waardoor die openheid van dergelijke toepassingsterreinen wordt veroorzaakt.

Laten we nog even teruggrijpen op de breuken.

In de (toepassings-)realiteit verschijnen breuken als:

- deel-geheel relaties
- operatoren
- meetgetallen

- verhoudingen
- en verdelingen en delingen (Treffers, Streefland en De Moor, 1994).

Deze verschijningsvormen kunnen in vele contexten worden ingekleed en daardoor kunnen de verschillende operaties met breuken een sterk wisselende semantische structuur krijgen. Voeg daarbij de variërende moeilijkheidsgraad van rationale getallen plus de diverse verbindingen die met decimale getallen, verhoudingen en procenten gelegd kunnen worden (en die ook nadrukkelijk in de doelstellingen van het basisonderwijs zijn opgenomen), dan krijgt men een indruk van de enorme diversiteit aan mogelijke toepassingsopgaven waar breuken in vervat liggen. En daarmee ook van het problematische karakter van de CEB-aanname.

Kortom, de aannames waarop de CEB de 'domain-referenced measurement' baseert, plus de toegepaste normering, zijn niet alleen onjuist omdat de PPON-toets verkeerd is geïnterpreteerd en gebruikt, zoals Wijnstra terecht aangeeft, maar ook omdat de CEB-psychometrici toepassingsdomeinen van het rekenen te simpel identificeren, dat wil zeggen, te gesloten opvatten en menen met representatieve steekproeven van dekkende items te kunnen werken waar in werkelijkheid geen sprake van een goede afspiegeling kan zijn en ten slotte het arbitraire karakter van de keuzen onderschatten.

Vakdeskundigen, naast de leden van het PPON-team, hadden de CEB op dit probleem kunnen wijzen.

Van der Linden en Zwarts (1995a) schrijven:

'In fact, the definition of the qualifications was a long process in which such factors as familiarity with the curriculum, teaching practices, quality of the learning materials, previous evaluations, and extensive consulting of relevant parties played an important role.' (p. 16). Zoals gezegd, zijn bij mijn weten echter geen vakdeskundigen geraadpleegd. Om vele redenen is dat jammer. Eén van de redenen heb ik hier geschetst: een methodologische misser met verstrekkende gevolgen voor de normering van het (reken-)onderwijs.

Literatuur

- Treffers, A., Streefland, L. & Moor, E. de (1994). *Proeve van een Nationaal Programma* (Deel 3A Breuken). Tilburg: Zwijsen.
- Van der Linden, W.J. & Zwarts, M.A. (1995a). Robustness of judgements in evaluation research. *Tijdschrift voor Onderwijsresearch*, 20 (1), 13-27.
- Van der Linden, W.J. & Zwarts, M.A. (1995b). De opbrengsten van het basisonderwijs: een repliek. *Tijdschrift voor Onderwijsresearch*, 20(1), 34-41.
- Wijnstra, J.H. (1995). De opbrengsten van het basisonderwijs volgens de CEB: enkele kanttekeningen bij de gevolgde normeringsprocedures. *Tijdschrift voor Onderwijsresearch*, 20 (1), 28-33.

Notities en Commentaren

TREFFERS SCHIET MIS

W.J. van der Linden¹ en M.A. Zwarts²

¹Universiteit Twente

²Inspectie van het Onderwijs

Het eerste wat opvalt in de notitie van Treffers (1996) is de neiging tot het uitdelen van absolute kwalificaties. Het onderzoek van de CEB is "irreëel", "ondeugdelijk" en "principieel onjuist". Het tweede is de suggestie dat er van alles fout is gegaan, zonder evenwel man en paard te noemen. Neem bijvoorbeeld het volgende citaat over de criteria waarmee de CEB de opgavenverzamelingen heeft beoordeeld, waarover Treffers opmerkt: "informeel is mij hierover wel wat bekend, maar het zou fair (*sic!*) zijn als de CEB eens precies uit de doeken zou doen hoe een en ander bij rekenen-wiskunde precies in zijn werk is gegaan – dat zou heel onthullend (*sic!*) kunnen zijn" (Treffers, 1996, pag. 104). Zoiets lijkt verdraaid veel op de journalistiek van het roddelblad, waarin ook gewerkt wordt met het principe dat er net zoveel rook gemaakt moet worden tot de lezers gaan geloven dat er vuur is.

Dezelfde toon werd aangehouden in een interview met het *NRC Handelsblad* van 20 januari 1994 dat vrijwel direct na de publikatie van het CEB-rapport (Commissie Evaluatie Basisonderwijs, 1994) verscheen, in een later ingezonden stuk in dezelfde courant en in een voordracht die Treffers hield tijdens een conferentie aan het einde van de landelijke raadpleging van het onderwijsveld over het CEB-rapport. De CEB heeft op deze kwalificaties nooit willen reageren, omdat er geen zakelijke kritiek op haar werkwijze of resultaten werd gepresenteerd. Omdat in de voornoemde notitie voor het eerst een zakelijk aspect aan de orde komt, namelijk de vraag of een kleiner aantal bekende items een groter aantal onbekende items kan representeren, volgt er thans wel een reactie. Deze wordt verderop gegeven, maar we vermelden nu alvast de conclusie: opsomming van alle elementen is niet de enige manier om een onbekende verzameling vast te leggen. De wiskunde laat ons zien dat je zo'n verzameling ook kunt omschrijven. En bij deze beschrijving kan nu juist de moderne psychometrie behulpzaam zijn. Eerst evenwel het volgende.

Het onderliggende probleem in al de reacties van Treffers lijkt te zijn dat hij moeite heeft met de oordelen die de CEB over de opbrengsten van het rekenonderwijs heeft geveld. In het eerder genoemde interview in het *NRC Handelsblad* werd hier onmiddellijk stelling tegen genomen. Het Nederlandse rekenonderwijs zit volgens Treffers goed in elkaar en de CEB, die voor de opbrengsten van een aantal onderdelen de beoordeling "matig" uitreikte, zit er naast. Nu heeft iedereen recht op een eigen oordeel en we laten dit aspect van Treffers' acties dan graag ook in zijn waarde. Maar de vraag welke evidentie hij voor zijn oordelen aanvoert, is natuurlijk een andere zaak. Waar de CEB vooral aansluiting heeft gezocht bij een evaluatietraditie waarin onafhankelijke, grootschalige peilingsonderzoeken en evaluatiestudies met representatieve steekproeftrekking het materiaal leveren, adstrueert Treffers vanuit een andere traditie. Zowel in het interview als in de later ingezonden stukken en de conferentiebijdrage baseerde Treffers zijn oordeel op casuïstiek en het aanvoeren van "internationale bijval" voor de vernieuwing van de didactiek van het rekenen in Nederland. De casuïstiek blijkt uit de veelvuldige voorbeelden van losse opgaven met "percentages leerlingen die deze internationaal of in Nederland goed maken". Daaruit moest dan blijken dat de CEB verkeerd had geëvalueerd. Over de omvang en de aard van de steekproeftrekking waarop deze cijfers gebaseerd waren, werd geen mededeling gedaan. Evenmin over de kwantificatie van de onzekerheid door middel van standaardschattingsfouten. Illustratief is dat toen één van ons tijdens de eerder genoemde conferentie bij de zoveelste bespreking van een losse opgave de vraag stelde bij welke steek-

proef deze afgenomen was, als antwoord kreeg te horen dat dit gebeurd was bij enkele scholen in de buurt van Utrecht, maar dat deze zeker representatief waren voor heel Nederland.

De internationale bijval zou moeten blijken uit het feit dat Treffers voortdurend aanvoert dat het Nederlandse rekenonderwijs "tot in China en de Verenigde Staten beroemd is om zijn didactische aanpak". Maar om welke Chinezen en Amerikanen gaat het eigenlijk? Hebben die allemaal het rekenonderwijs gezien, zoals dat in de Nederlandse taal wordt gegeven, door leraren met een opleiding aan een Nederlandse Pedagogische Academie, aan de hand van verschillende leerboeken gedrukt in het Nederlands en met een leerplan dat 6-8 schooljaren omvat? Of kennen ze alleen een beschrijving van de didactische aanpak die in sommige leerboeken gematerialiseerd is? In het Engels? Door degenen die deze aanpak zelf ontwikkeld hebben? Chinezen en Amerikanen hebben gelukkig nog een vriendelijke cultuur en een sterk ontwikkelde neiging om anderen te complimenteren met hun werk. Maar hoe zouden ze iets kunnen weten van de *feitelijke* opbrengsten van het Nederlandse basisonderwijs, zoals dat onder de hier geldende omstandigheden op ruim 8000 scholen verzorgd wordt?

Het beeld bestaat bij ons dat Treffers een rekendidacticus is die met overgave werkt aan de vernieuwing van de leerstof voor het rekenen in het basisonderwijs, hier graag scholen en instellingen bij betreft en de resultaten met enthousiasme uitdraagt. Inmiddels is er in Nederland reeds een aanzienlijke traditie van vernieuwing van het rekenonderwijs opgebouwd, maar de missie is nog lang niet geklaard want niet alle leraren werken al met de nieuwe aanpak. Voor de energie bestaat bij ons bewondering en we realiseren ons dat een noodzakelijke voorwaarde voor de vernieuwing van een belangrijke zaak als het rekenonderwijs ligt in de aanwezigheid van vakbekwame didactici met grote gedrevenheid. De inbreng van hen bij de beoordeling van itemverzamelingen in evaluatie- of peilingsonderzoek is van eminent belang. In dit opzicht zijn we blij dat het Cito bij het samenstellen van het PPON-materiaal uitgebreid vakdeskundigen heeft geraadpleegd. We hebben dankbaar kennis genomen van hun reacties op het materiaal, waaronder de reactie van Treffers, die we verderop nog uitgebreid zullen citeren. Maar één van de fundamentele principes van evaluatie-onderzoek is het principe van de onafhankelijkheid, dat impliceert dat je een ontwikkelaar niet de opbrengsten van het eigen ontwikkelingswerk moet laten beoordelen. Dan zijn de uitkomsten namelijk voorspelbaar. Enerzijds mag er niet teveel kritiek op het lopende onderwijs gegeven worden, want er is reeds aanzienlijk geïnvesteerd in nieuwe materialen en aanbevelingen. Anderzijds moet er nog wel ruimte voor toekomstig werk blijven. Deze beperkingen leiden tot curieuze zinnen, zoals de volgende die het *NRC Handelsblad* uit de mond van Treffers optekent: "Het Nederlandse rekenonderwijs scoort, zeker als je het internationaal bekijkt, helemaal niet slecht, maar het kan nog aanmerkelijk beter". Wat is dat voor onderwijs dat tegelijkertijd "helemaal niet slecht" is en "nog aanmerkelijk beter kan"?

In de rapporten van de CEB wordt niet gesteld dat het rekenonderwijs slecht en niet meer te redden is, maar wordt een genuanceerder beeld gegeven. Als dit beeld per se op één noemer gebracht moet worden, dan kan gesteld worden dat de opbrengsten, zeker voor de hogere vaardigheden als verhoudingen en breuken, matig zijn en verbeterd kunnen worden, o.a. door een herziening van het aanbod, onderwijs op maat in de eerste leerjaren en een meer op opbrengst gerichte cultuur. Het door de CEB geschetste beeld van de opbrengsten is vrijwel identiek aan dat van de vakdidacticus rekenen uit het PPON-project in een onafhankelijk maar vrijwel gelijktijdig gepubliceerd themanummer van het blad *School* (zie voor deze parallel onze reactie op de notitie van Wijnstra in Van der Linden & Zwarts, 1995b). En komt het inderdaad niet verrassend dicht bij de tweede helft van de hierboven geciteerde woorden van Treffers, namelijk dat het rekenonderwijs op onderdelen "nog aanmerkelijk beter kan"?

Goed, dit keer probeert Treffers het dus via een kritiek op de gebruikte itemverzamelingen. Dit is tenminste een zakelijk punt waarop we commentaar kunnen geven. Allereerst moeten we echter een misverstand uit de weg ruimen dat wellicht bij de lezers zou kunnen zijn ontstaan. Zinsneden in Treffers' notitie als "de PPON-toets" en "een relatief beperkt setje van enkele tientallen vraagstukken" doen vermoeden dat de evaluatie van de CEB van het rekenonderwijs op de empirische gegevens uit één PPON-toets teruggaat. Dat is niet zo. In het PPON-project

zijn gegevens over de opbrengsten van het rekenonderwijs gepeild met behulp van een verzameling van 473 items voor het einde en 264 items voor het midden van het basisonderwijs. Daarnaast is ons oordeel bepaald door de resultaten uit de OVB-evaluatie en uit overig empirisch onderzoek. We beperken ons hier evenwel tot het door Treffers gekozen object van de rekenitems in PPON. Zijn centrale stelling ligt vervat in het volgende citaat: "Men dient namelijk bij de operationalisering van leerdoelen in dekkende opgaven onderscheid te maken tussen gesloten en open verzamelingen van corresponderende opgaven. Bij gesloten verzamelingen is het zinvol om over afspiegeling in de vorm van een beperkt aantal items te spreken die representatief zijn voor de totale verzameling. Maar bij open verzamelingen is zulks niet mogelijk" (Treffers, 1996, pag. 103). Voor het overige bestaat zijn notitie, in herkenbare stijl, uit een uitgebreide bespreking van een reeks voorbeelden van rekenopgaven met breuken. In onze reactie gaan we niet op deze voorbeelden in. Dat opgaven in moeilijkheid kunnen verschillen, was ons al bekend. Bovendien behoort ongeveer de helft van de opgaven tot het domein van het voortgezet onderwijs en dat staat hier niet ter discussie.

Waar het ons om gaat, is de centrale vraag of men aan mag nemen dat de verzamelingen opgaven in het PPON-project representatief zijn voor de in principe open verzameling opgaven die men zich bij de leerstof voor het rekenen in het basisonderwijs in kan denken. De stelling van onze opponent is dat dit niet kan. Bij de beantwoording van deze vraag moet men echter de in het PPON-project gebruikte methodologie in ogenschouw nemen. De items in het PPON-project werden gecalculeerd op een aantal schalen met behulp van een responsmodel uit de moderne psychometrie. Om een goede modelpassing te verkrijgen, bleken voor het einde van het basisonderwijs in totaal maar liefst 27 verschillende schalen noodzakelijk. Dit empirische gegeven mag dus opgevat worden als een aanwijzing dat 27 verschillende dimensies nodig zijn om de verzameling van 473 items te representeren. Als dit aantal voldoende is voor de beschrijving van de opbrengsten van het Nederlandse onderwijs, is het enige wat er nog kan variëren de verdeling van de items op deze schalen. Als we ons, in navolging van Treffers, beperken tot de moeilijkheidsparameter in het model, valt de vraag naar de representativiteit van de opgaven dus uiteen in de volgende twee deelvragen:

1. Is het aantal van 27 PPON-schalen voldoende om de opbrengsten van de volledige leerstof van het rekenen in het basisonderwijs te beschrijven? Of zijn er aanwijzingen dat belangrijke onderwerpen over het hoofd zijn gezien?
2. Is de verdeling van de moeilijkheid van de items op ieder van deze schalen representatief voor de verzameling opgaven die men zich zou kunnen indenken bij deze leerstof?

Onze eerste reactie op Treffers is dus dat er niet één vraag naar de representativiteit van de items gesteld moet worden, maar dat het er *twee* zijn. We zullen nu betogen dat er voortreffelijke indicaties zijn dat het antwoord op de eerste vraag positief is en het antwoord op de tweede vraag er niet toe blijkt te doen.

Om een antwoord op de eerste vraag te krijgen, is de beoordeling van de itemverzameling door inhoudsdeskundigen van het grootste belang. De PPON-verzameling is door diverse inhoudsdeskundigen beoordeeld, zowel op initiatief van het PPON-project als naderhand in een klein beoordelingsexperiment door de CEB. De bevindingen waren in het algemeen zeer positief. Wellicht is het voldoende als we hier alleen de bevindingen van Treffers (1988) citeren:

"Voorts kan worden vastgesteld dat de kwaliteit van de PPON-opgave in het algemeen gesproken ruimschoots toereikend is om het onderwijsniveau passend te peilen - de Cito-samenstellers Bokhove c.s. verdienen daarvoor lof - al zullen bepaalde problemen in deze peiling de toets der kritiek niet kunnen doorstaan, zoals zal blijken." (pag. 182);

"De onderhavige PPON-toetsen voor rekenen zijn een bruikbaar instrumentarium om het onderwijsniveau te meten; de opgaven zijn in het algemeen gesproken van goed gehalte en representeren in behoorlijke mate de concrete leerdoelen op de gebieden van basisvaardigheden, cijferen, verhoudingen, procenten, breuken, kommagetallen en meten." (pag. 187).

De in het eerste citaat genoemde "bepaalde problemen" blijken vervolgens alleen te bestaan uit het feit dat Treffers van oordeel is dat de zakrekenmachine ingepast had moeten worden en dat de opgaven met toepassingen nogal afhankelijk zijn van de beheersing van de Nederlandse taal. We citeren zijn oordeel op dit laatste punt, omdat dit laat zien dat er desniettemin geen consequenties voor de tweede representativiteitsvraag zijn:

"Iedere schaal van dit PPON-onderzoek bevat toepassingen, vaak voor meer dan de helft. Dit is in overeenstemming met de doelstellingen van het rekenen: iedere vaardigheid moet tenslotte ook toegepast kunnen worden. Het gevolg is echter wel dat de collectie opgaven nogal 'talig' is. Sommigen zullen wellicht hieraan de relatief teleurstellende (*sic!*) resultaten toeschrijven. Dit lijkt mij echter niet terecht. Soms staan tekstopgaven laag op de ladder en soms hoog." (pag. 186)

Als je deze teksten van Treffers ter gelegenheid van het verschijnen van de PPON-resultaten leest, vraag je je vertwijfeld af waarom er nu niets meer van de PPON-opgaven zou deugen.

Om de tweede vraag te kunnen beantwoorden, heeft de CEB de methode van de robuustheidsanalyse gevolgd. Via computersimulatie werd de feitelijke verdeling van de moeilijkheid van de items op de PPON-schalen op diverse manieren systematisch verstoord om het eventuele effect op de beoordelingen door de CEB na te gaan. De resultaten van deze uitgebreide studie zijn gepresenteerd in Van der Linden en Zwarts (1995a). De lezer wordt opgeroepen de uitkomsten van deze studie te (her)lezen. In het algemeen blijken, zelfs bij eenzijdige verstoring van de verdeling van itemmoeilijkheden, op een enkel onderwerp na, de parameters waarop de CEB zijn oordelen heeft gebaseerd, verrassend stabiel te zijn. Of voor iedere schaal in de PPON-peiling de verdeling van de moeilijkheid van de opgaven volledig representatief is voor de verdeling van alle mogelijke opgaven op deze schaal die de leerstof goed representeren, is een vraag waar we nooit definitief antwoord op zullen krijgen. Belangrijker is evenwel de vraag in welke mate de conclusies van de CEB van deze onzekerheid afhangen. Onze studie laat zien dat deze onzekerheid er voor de CEB-oordelen niet zo veel toe doet.

Tenslotte nog een leeswijzer voor Treffers. In het CEB-rapport over de opbrengsten van het basisonderwijs staat één zin die eerst door Wijnstra (1965) en thans door Treffers uit zijn verband is gelicht. De zin luidt: "De PPON-toetsen zijn zo geconstrueerd, dat leerlingen *in principe* de opgaven goed moeten kunnen maken" (cursivering van ons). Deze zin is natuurlijk geen voorschrift dat alle leerlingen in het basisonderwijs op de beide itemverzamelingen uit PPON een perfecte score zouden moeten hebben. Deze interpretatie is inconsistent met de overige tekst in het rapport en de door de CEB gekozen normeringen, die toelaten dat om het predikaat voldoende te krijgen een substantieel deel van de leerlingen zelfs meer dan 40-50% van de opgaven fout mag hebben. Wat er met deze zin wel gezegd werd, kan aangegeven worden met de eerste van de twee bovenstaande representativiteitsvragen: de leerlingen moeten in principe de opgaven kunnen maken omdat de PPON-schalen zich alleen richten op de leerstof van het basisonderwijs en er dus geen opgaven zijn die thuishoren op schalen die dat niet doen. Het is daarom naïef om te menen dat volgens de CEB alle leerlingen in het basisonderwijs zo knap zouden moeten zijn dat ze iedere opgave in de PPON-verzamelingen onveranderlijk goed maken. De relatie tussen de vaardigheden van de leerlingen en de opgaven in de verzamelingen is probabilistisch van aard. Dat is nu juist de kern van de methodologie van de itemresponstheorie, waar PPON mee gewerkt heeft en waar wij onze eigen berekeningen op hebben gebaseerd. Het referentiekader van de lezers is niet altijd dat van de schrijvers. Dit is onderhands al eens aan Wijnstra uitgelegd. Desondanks maakte hij van deze ene zin de hoeksteen van zijn notitie. We hadden dan ook liever gezien dat Treffers zich op de CEB-rapporten zou richten in plaats van blindelings Wijnstra te citeren.

LITERATUUR

- Commissie Evaluatie Basisonderwijs (1994). *Inhoud en opbrengsten van het basisonderwijs* (Deelrapport 1). De Meern: Inspectie van het Onderwijs.
- Van der Linden, W.J. & Zwarts, M.A. (1995a). Robustness of judgments in evaluation research. *Tijdschrift voor Onderwijsresearch*, 20, 13-27.
- Van der Linden, W.J. & Zwarts, M.A. (1995b). De opbrengsten van het basisonderwijs: een repliek. *Tijdschrift voor Onderwijsresearch*, 20, 34-41.
- Treffers, A. (1988). Over de merkbare invloed van onderwijsmethoden op leerprestaties. In J.M. Wijnstra (red). *Balans van het rekenonderwijs in de basisschool* (pag. 181-189). Arnhem: Cito.
- Treffers, A. (1996). Rekeninhoudelijke kritiek op CEB-repliek. *Tijdschrift voor Onderwijsresearch*, 21, 103-105.
- Wijnstra, J.M. (1995). De opbrengsten van het basisonderwijs volgens de CEB: Enkele kanttekeningen bij de gevolgde normeringsprocedure. *Tijdschrift voor Onderwijsresearch*, 20, 28-33.

Notities en Commentaren

MAAR RAAKT WEL ZIJN EIGEN VOET

W.J. van der Linden¹ en M.A. Zwarts²

¹Universiteit Twente en ²Inspectie van het Onderwijs

Het blijft moeilijk om in de betogen van Treffers een consistente lijn te ontdekken. Wat moeten we denken van een auteur die een paar jaar geleden de representativiteit van de PPON itempool nog de hemel in prees (Treffers, 1988; zie ook onze vorige reactie), nu uit twee merkwaardige stellingen het tegendeel concludeert, maar vervolgens wel in een adem doorgaat en verheugd meedeelt dat de resultaten uit PPON-2 overtuigend aantonen dat de rekenmethoden die zich op het werk van het Freudenthal Instituut baseren het beter doen (Treffers, 1996b)? Wat voor waarde heeft deze conclusie als het PPON-materiaal niet representatief en dus niet inhoudsvalide is? In zijn eerste notitie schoot Treffers mis, maar thans lijkt het er verdraaid veel op dat hij in zijn eigen voet heeft geschoten. Gelukkig maar dat de CEB bij zijn oordeelsvorming over de kwaliteit van het rekenonderwijs wel aandachtig kennis heeft genomen van de commentaren van Treffers, maar er niet blind op heeft gevestigd.

Halve waarheden en hele suggesties

In de tweede notitie van Treffers staan nogal wat halve waarheden en hele suggesties. We laten deze puntsgewijs de revue passeren en nemen daarbij de volgorde in acht die Treffers in deze notitie heeft aangehouden.

1. In stelling 1 wordt betoogd dat als de CEB-normen op het beste rekenland ter wereld - volgens Treffers was dat toch Nederland? - toegepast worden, er geen voldoende uit de bus zal komen. Dit is een experiment dat de CEB nooit voor haar rekening zou willen nemen. In evaluatie-onderzoek is het absoluut onjuist om eisen aan de opbrengsten van een vak te stellen zonder rekening te houden met de condities waarover de onderwijsgegenden zelf geen controle hebben. De CEB heeft dit reeds gesteld in zijn plan voor de evaluatie van het basisonderwijs en als voorbeelden van zulke condities genoemd het landelijk onderwijsbeleid en de daarop gebaseerde uitgaven, de materiële voorzieningen (o.a. leer- en hulpmiddelen en de mogelijkheden van onderwijs-op-maat), de kenmerken van de schoolbevolking (o.a. etniciteit, gezinssituatie en schoolverzuim) en de kenmerken van de leraarpopulatie (o.a. opleidingsniveau en arbeidssatisfactie) (Commissie Evaluatie Basisonderwijs, 1992, pag. 13). Voor de opbrengsten van een individueel vak komt daar nog bij het beslag dat de andere schoolvakken op de beschikbare tijd leggen. De normen die de CEB heeft aangelegd, zijn gebaseerd op een zo compleet mogelijke empirische beschrijving van de toestand van het basisonderwijs. In andere landen gelden andere condities en is de beschikbare onderwijstijd anders verdeeld. Als de werkwijze van de CEB elders herhaald zou moeten worden, zouden er dus ongetwijfeld andere normen uit rollen. Iedere serieuze evaluator zal hetzelfde principe volgen. Als de *Consumentenbond* stereoversterkers beoordeelt, houdt ze ook rekening met de prijs van de aanbevolen apparaten.

2. In stelling 2 wordt beweerd dat het mogelijk is om toetsen te ontwikkelen die psychometrisch identiek zijn maar verschillende scoreverdelingen opleveren. Hierover kunnen we kort zijn. Kennisname van twee bladzijden in een leerboek over psychometrie is voldoende om te weten dat verschillende test- en itemparameters verschillende scoreverdelingen impliceren. Deze stelling is klinkklare onzin.

3. De bewijsvoering voor beide stellingen moet het twee keer doen met een beroep op anonieme autoriteiten die Treffers zouden bijvallen. Zo adstrueert men niet in een wetenschappelijke discussie. Wie zijn al die vakdeskundigen die het met stelling 1 eens zijn? Onder wiskundigen wordt bepaald verschillend gedacht over de kwaliteit van het rekenonderwijs. Direct

Adres: Universiteit Twente, Faculteit Toegepaste Onderwijskunde, Postbus 217, 7500 AE Enschede.

na het vaker gememoreerde interview met Treffers in het *NRC Handelsblad* kon men in dezelfde courant een ingezonden stuk lezen van de wiskundigen Barendregt (KUN) en Ruttkay (VU), die zich zeer bezorgd uitlieten over het Nederlandse rekenonderwijs en kritiek hadden op de aanpak die het Freudenthal Instituut voorstaat. Opnieuw in dezelfde courant werd op 6 september j.l. een artikel afgedrukt van de hand van H.K. Kuiken, hoofdredacteur van het *Journal of Engineering Mathematics*, waarin deze betoogde dat hij op grond van zijn ervaringen met het universitaire toelatingsexamen in China - het land dus waar volgens Treffers de Nederlandse aanpak wereldberoemd is - tot de conclusie komt dat vrijwel geen Nederlandse kandidaat het onderdeel met goed gevolg zou doorstaan. Ook wijst hij erop dat in de wiskunde-olympiade Nederland de laatste jaren onveranderlijk tussen de dertigste en de veertigste plaats figureert. En wie zijn die Cito-vakdeskundigen die de merkwaardige bewering in stelling 2 voor hun rekening durven nemen? Ze bestaan niet of zouden op staande voet ontslagen moeten worden.

4. De waarheid van de eerste stelling meent Treffers verder aan te kunnen tonen door een "vergelijking van overeenkomstige items in verschillende nationale onderzoeken" en door "de resultaten van het tweede en het derde internationale onderzoek voor rekenen-wiskunde van de IEA te analyseren". Welke nationale onderzoeken heeft Treffers precies geraadpleegd? Buiten Australië, de V.S. en Nederland zijn er amper landen waar ooit een nationale peiling van de opbrengsten van het basisonderwijs op basis van een representatieve steekproeftrekking is uitgevoerd. Of gaat het hier om buitenlandse herhalingen van het in onze vorige reactie op Treffers gememoreerde onderzoek op "enkele scholen in de buurt van Utrecht die zeker representatief zijn voor heel Nederland"? Bovendien weten alle toetsconstructeurs dat kleine verschillen in de bewoording of de vraagvorm van een item tot even onvoorspelbare als drastische verschuivingen in de moeilijkheid ervan kunnen leiden. Zie bijvoorbeeld de opmerking terzake van Wijnsstra (1995). Scoreverdelingen op één verzameling items voorspellen uit verdelingen op een andere verzameling kan niet op grond van didactische analyse van de inhoud van opgaven, maar alleen met behulp van steekproeftechnieken en statistische analyses die bekend staan onder de naam testequivalering. Voor een recent overzicht van degelijke methoden verwijzen we onze opponent naar Kolen en Brennan (1995). Wat de IEA-toetsen betreft, deze zijn veel te beperkt om uitspraken te kunnen doen op het niveau van de kerndoeldomeinen voor rekenen die de CEB in zijn oordelen heeft aangehouden. Trouwens, het TIMSS-project, waarbinnen de derde IEA-peiling van het rekenonderwijs wordt uitgevoerd, heeft zijn toetsen nog niet zo lang geleden afgenomen en zit nu midden in de voorbereidingen van zijn data-analyses. Resultaten van internationale vergelijkingen zijn op dit moment nog niet beschikbaar.

5. Het gebrek aan representativiteit van de itempool uit PPON meent Treffers verder aan te kunnen tonen met "elementaire vakinhoudelijke middelen". Het wondermiddel blijkt een taxonomie voor de leerstof over breuken te zijn die ontstaat door vier verschillende indelingscriteria met elkaar te combineren. Maar wat kun je nu in vredesnaam met een taxonomie bewijzen? Met het toevoegen van indelingscriteria aan taxonomieën kun je blijven doorgaan en iedere didacticus heeft waarschijnlijk zijn eigen voorkeur voor een taxonomie. Voor de vraag of de verzameling PPON-opgaven een representatieve steekproef uit het leerstofdomein vormt, zijn deze indelingen statistisch gezien van geen enkel belang. Het enige wat ze in feite mogelijk maken, is de trekking van een gestratificeerde in plaats van een enkelvoudig aselechte steekproef items. Ironische genoeg leert de steekproeftheorie ons dat in het eerste geval het aantal items in de steekproef juist kleiner kan zijn om de aantal-goedscore op het domein met een gelijke precisie te schatten (zie bijvoorbeeld Bhattacharyya & Johnson, 1977, hoofdstuk 16). Het belangrijkste is evenwel het volgende. Wat voor het onderwerp telt, zijn niet de theoretische discussies over de juiste taxonomie, maar het *empirische antwoord op de vraag met hoeveel dimensies de leerstof en de itemverzameling beschreven moet worden*. De in PPON gebruikte methodologie van een op een IRT-model gebaseerde schalingsmethode is een effectief antwoord op deze vraag. Als het aantal dimensies bekend is, is het enige waar de verdeling van de aantal-goedscores nog van afhangt de verdeling van de itemparameters over deze schalen. Door middel van robuustheidsanalyse heeft de CEB laten zien dat haar oordelen amper gevoelig waren voor verschuivingen in deze parameterverdelingen (Van der Linden & Zwarts, 1995a). Dit is op-

nieuw door ons betoogd in onze reactie op Wijnstra (Van der Linden & Zwarts, 1995b) en tenslotte nog een keer in onze vorige reactie op Treffers (Van der Linden & Zwarts, 1996). We vragen ons zo langzamerhand oprecht af of Treffers onze stukken wel leest.

6. In de volgende alinea meent Treffers ons te moeten wijzen op de moeilijkheidsgraad van de opgaven in de PPON-pool. De suggestie dat deze ons niet bekend zouden zijn, is wel heel pedant. Zoals al het overige materiaal, beschikte de CEB ook over deze informatie. Datafiles met de itemparameters werden ons door het Cito welwillend ter beschikking gesteld en zijn uitvoerig gebruikt in ons rekenwerk. Het meest krasse is nog wel dat Treffers weet dat één van ons ten tijde van de rekenpeiling als Cito-medewerker op PPON werkte. Nee hoor, de CEB heeft deze informatie heus niet vergeten toen zij haar beoordelingen opstelde.

7. In zijn analyse van de vraag of er sinds 1987 veel in het rekenonderwijs is gewijzigd, heeft onze opponent alleen oog voor het didactische karakter van de gebruikte leermethoden—en dan nog slechts voor de vraag of ze zich baseren op de ideeën van het Freudenthal Instituut. Kennelijk verwacht hij aardverschuivingen in de opbrengsten van het onderwijs ten gevolge van een verandering van leerboek. Zo gaat het er in het onderwijs niet aan toe. Er zijn factoren die veel krachtiger zijn. De betekenis van leerboeken zinkt zelfs volledig in het niet als, zoals in het PPON-onderzoek is gebleken, een aanzienlijk aantal scholen amper toekomt aan de behandeling van sommige onderwerpen. Voor andere voorbeelden van effectiviteitsbepalende factoren verwijzen we naar de uitkomsten van een meta-analyse in Fraser (1989). De CEB stond voor de taak om het basisonderwijs in al zijn aspecten te evalueren. Onze stelling blijft dat er in de betreffende 5 jaar, met uitzondering van de toename van het gebruik van realistische methoden, in het geheel van het rekenonderwijs niet veel is veranderd. Belangrijk is dat de lezer opmerkt dat het kwantitatieve overzicht dat Treffers schijnbaar ter ondersteuning van zijn bewering geeft, slechts slaat op de verschillen in opbrengsten tussen traditionele en realistische rekenmethoden en dus *niet* op de vergelijkingen van de opbrengsten tussen 1987-1992. (De tweede peiling vond plaats in 1992 en niet in 1994 zoals Treffers lijkt te willen suggereren.) Deze verschillen waren al in 1987 bekend en volgen het vaste patroon uit het leerboekenonderzoek dat methoden het beter doen op de onderdelen die hun auteurs belangrijk vinden en slechter op de onderdelen die ze minder belangrijk vinden. In tegenstelling tot Treffers heeft de CEB, direct na het beschikbaar komen van de laatste gegevens, de consequenties van het verschil in opbrengsten tussen 1987-1992 voor haar beoordelingen van het rekenonderwijs wel geanalyseerd (Zwarts, 1994). De uitkomst was dat op het niveau van de 27 individuele rekenschalen er 2 positiever (gehele getallen en hoofdrekenen met vermenigvuldigen en delen) en 2 negatiever (cijferend vermenigvuldigen en berekenen van het aantal maateenheden) beoordeeld zouden zijn geworden. Worden de oordelen evenwel geaggregeerd op het niveau van de kerndoel domeinen, en dat is het rapportageniveau dat de CEB in zijn eindrapport aangehouden heeft, dan vallen deze verschuivingen tegen elkaar weg en blijft het oordeel onveranderd.

8. Tenslotte dan de kwestie van de vier dansopgaven die Treffers in zijn vorige notitie formuleerde. De lezer moet ze nog maar eens nalezen en voor zichzelf de vraag beantwoorden of de beheersing van de laatste twee wel tot het basisonderwijs gerekend mag worden. Wij blijven apert van mening dat dit niet het geval mag zijn. Deze opgaven verschillen veel te veel in niveau van de PPON-opgaven, die Treffers gemiddeld al te moeilijk voor het basisonderwijs vond. Opgaven die een algemeen bewijs vragen ("Toon aan dat ...") mogen niet voorgelegd worden aan leerlingen die nog enkel geen idee over methoden van bewijsvoering hebben. Dat onder de vele manieren waarop een opgave opgelost kan worden er één is die eenvoudig blijkt te zijn, betekent nog niet dat de opgave voor de leerlingen van de basisschool eenvoudig is. We gaan dan nog voorbij aan het feit dat de redactie van de gemeenschappelijk stam van deze vier opgaven niet specifiek genoeg is voor leerlingen van de basisschool en dat reeksen opgaven waarvan het antwoord op de volgende vraag afhangt van succes op een voorafgaande vraag meettechnisch gezien uit de boze zijn. Dergelijke zaken tellen mee als de moeilijkheid van opgaven empirisch wordt vastgesteld. We zien hier opnieuw een fraaie demonstratie van het feit dat als didactici lang met een onderdeel van de leerstof bezig zijn, ze de neiging krijgen om de moeilijkheid ervan te gaan onderschatten.

SLOTWOORD

Wat ons betreft zijn we aan het einde gekomen van onze discussie met Treffers. Het is moeilijk discussiëren met een opponent die slechts oog heeft voor het didactische aspect van het rekenonderwijs, die zijn theorieën baseert op informele analyses van losse opgaven, zich voor evidentie beroept op meningen van gelijkgezinden en weinig interesse toont voor de statistische methodologie van peilingsonderzoek. Bovendien reageert hij in zijn tweede notitie nergens op de eerder door ons aangevoerde argumenten, zodat een echte discussie niet ontstaat. Opnieuw lijkt het er verdraaid veel op alsof hij in het wildeweg schiet in de veronderstelling dat rook de beste manier is om vuur te suggereren. Dat hij zijn eigen voeten raakt en vervolgens in zijn betoog over de betekenis van het realistische rekenonderwijs eigenlijk geen been meer heeft om op te staan, komt niet voor onze rekening. Voor een eerlijk duel zijn we altijd in, maar van losvliegende kogels blijven we graag verschoond.

LITERATUUR

- Bhattacharyya, G.K., & Johnson, R.A. (1977). *Statistical concepts and methods*. New York City, NY: Wiley.
- Commissie Evaluatie Basisonderwijs (1992). *Plan voor de evaluatie van het basisonderwijs*. De Meern: Inspectie van het Onderwijs.
- Fraser, B.J. (1989) Research syntheses on school and instructional effectiveness. *International Journal of Educational Research*, 13, 707-720.
- Kolen, M.J., & Brennan, R.L. (1995). *Test equating: Methods and practices*. New York City, NY: Springer-Verlag.
- Van der Linden, W.J. & Zwarts, M.A. (1995a). Robustness of judgments in evaluation research. *Tijdschrift voor Onderwijsresearch*, 20, 13-27.
- Van der Linden, W.J. & Zwarts, M.A. (1995b). De opbrengsten van het basisonderwijs: een repliek. *Tijdschrift voor Onderwijsresearch*, 20, 34-41.
- Treffers, A. (1988). Over de merkbare invloed van onderwijsmethoden op leerprestaties. In J.M. Wijnstra (red). *Balans van het rekenonderwijs in de basisschool*. Arnhem: Cito.
- Treffers, A. (1996a). Rekeninhoudelijke kritiek op CEB-repliek. *Tijdschrift voor Onderwijsresearch*, 21,
- Treffers, A. (1996b). Rekeninhoudelijke kritiek op CEB-repliek (2).
- Wijnstra, J.M. (1995). De opbrengsten van het basisonderwijs volgens de CEB: Enkele kanttekeningen bij de gevolgde normeringsprocedure. *Tijdschrift voor Onderwijsresearch*, 20, 28-33.
- Zwarts, M.A. (1994). *Beoordeling gegevens nieuwe rekenpeiling* (intern memorandum). De Meern: Inspectie van het Onderwijs.

NOOT VAN DE REDACTIE

In dit commentaar reageren de auteurs op een niet gepubliceerde dupliek van Treffers. Nadat zowel de dupliek van Treffers als dit weerwoord daarop van Van der Linden en Zwarts door de redactie waren geaccepteerd, trok de heer Treffers zijn dupliek terug. De redactie betreurt de gang van zaken. Op verzoek van de heren Van der Linden en Zwarts wordt hun tweede commentaar toch gepubliceerd. Geïnteresseerde lezers kunnen de tekst van de dupliek van Treffers bij de auteur opvragen. Zijn adres is: Prof.dr. A. Treffers, Freudenthal Instituut, Tiberdreef 4, 3561 GG Utrecht.

Onderwijsexpansie of meritocratisering?

Een verklaring voor de afgenomen samenhang tussen het ouderlijk sociaal milieu en het hoogst bereikte opleidingsniveau van kinderen

M.H.J. Wolbers en P.M. de Graaf
Vakgroep Sociologie, Katholieke Universiteit Nijmegen

ABSTRACT

In the Netherlands, effects of family background on highest educational level attained by children have decreased during the 20th century. Previous analyses show that family background effects on given transitions in educational careers have been stable. By using complete data on educational careers of cohorts born between 1928 and 1967 we investigated whether these findings are conflicting. Mare proposed a model of educational careers that distinguishes between the results of educational expansion and the results of changes in the principles by which schooling is allocated. This model shows that stable family background effects on transitions will go together with decreasing family background effects on highest educational attainment, because the effect of social origins declines over the educational career. Our results corroborate Mare's hypotheses. A cohort comparison showed no trend in the effects of social origins on the odds of making transitions. This shows that the declining effect of family background on highest educational attainment in the Netherlands is a result of educational expansion only. No meritocratization can be observed, when one focuses on individual transitions.

1. INLEIDING EN PROBLEEMSTELLING

In Nederland heeft veel onderzoek plaatsgevonden naar de relatie tussen de maatschappelijke positie van ouders en die van kinderen. Er zijn twee redenen te veronderstellen dat deze samenhang in de twintigste eeuw is verminderd. Enerzijds maakt men tegenwoordig gebruik van efficiënte selectieprocessen ('achievement' in plaats van 'ascription'). Anderzijds heeft de overheid een activistisch beleid gevoerd om ongelijke kansen tussen verschillende sociale milieus te verkleinen. In een moderne samenleving zou daarom iemands sociale herkomst minder bepalend zijn geworden voor zijn opleiding en beroep, terwijl tegelijkertijd de relatie tussen iemands bereikte opleidings- en beroepsniveau sterker is geworden. Het huidige empirisch onderzoek levert echter geen eenduidig beeld over de veronderstelde afname van de invloed van het sociaal milieu op onderwijssucces. Er bestaan onderzoeken waaruit inderdaad naar voren is gekomen dat in deze eeuw de samenhang tussen het ouderlijk sociaal milieu en het hoogst bereikte opleidingsniveau van kinderen is afgenomen (Ganzeboom, 1984; Ganzeboom & De Graaf, 1989; De Graaf & Ganzeboom, 1993). Andere onderzoekers daarentegen hebben bij een aantal cohorten van leerlingen, die de overgang van het lager naar het voortgezet onderwijs maakten, geen afzwakking gevonden van het effect van de sociale achtergrond (bijvoorbeeld Dronkers (1983), Vrooman & Dronkers (1986), Faasse, Bakker, Dronkers & Schijf (1987) en Bakker & Schouten (1991)), of slechts in een geringe mate (Bakker & Cremers, 1994).

De vraag is of deze onderzoeksuitslagen tegenstrijdig zijn, of dat ze berusten op een verschillende weergave van de samenhang tussen sociale herkomst en onderwijssucces. Het laatste is meer voor de hand liggend, want er bestaat een tweeslachtigheid in het meten van onderwijsongelijkheid (Dronkers & De Graaf, 1995, p. 47). Enerzijds kan men de invloed van het ouderlijk sociaal milieu op het hoogst behaalde diploma vaststellen (bereikt niveau), anderzijds is het mogelijk de invloed van het sociaal milieu op belangrijke overgangen gedurende de onderwijsloopbaan te onderzoeken (overgangskansen). In de bovengenoemde studies is dit onderscheid duidelijk te herkennen. De onderzoekers die besluiten dat er een afnemende ongelijkheid is, gebruikten het hoogst bereikte opleidingsniveau als maatstaf voor onderwijsongelijkheid. Daarentegen bestudeerden de onderzoekers die concluderen dat er geen sprake van verandering is de overgang van het lager naar het voortgezet onderwijs. Mare (1979, 1980, 1981) verhelderde deze kwestie door te laten zien dat beide benaderingen er niet in slagen een duidelijk onderscheid te maken tussen twee verschillende ontwikkelingen: 1) veranderingen in de selectie- en allocatieprocessen op basis waarvan iemands opleidingsniveau wordt bepaald en 2) de algemene groei in onderwijsdeelname. Mare stelde een wiskundig model voor waarin onderwijsloopbanen worden gezien als een opeenvolging van keuzes of transities. Dit is analoog aan Boudon (1974). Bij elke transitie maakt een leerling de keuze te stoppen met het volgen van onderwijs of door te gaan met een verdere (dat wil zeggen hogere) opleiding. De kansverhouding om een transitie al dan niet te 'overleven' wordt beïnvloed door (onder meer) de sociale achtergrond. Gebruik makend van dit model liet Mare zien dat wanneer *a)* er sprake is van onderwijsexpansie en wanneer *b)* de milieu-effecten over de onderwijsloopbaan heen afnemen en als daarnaast *c)* per transitie in de tijd stabiele milieu-effecten aanwezig zijn, dat dan *d)* de samenhang tussen de sociale achtergrond en het hoogst bereikte opleidingsniveau van de respondent moet zijn afgenomen. In bovengenoemde studies gaat het dan om een schijnbare tegenspraak. Volgens het model van Mare is het heel goed verenigbaar dat bij de transitie van het lager naar het voortgezet onderwijs *geen* verandering wordt gevonden in de invloed van het sociaal milieu in het onderwijs, terwijl de samenhang tussen sociale herkomst en het hoogst bereikte opleidingsniveau *wel* is afgenomen. De combinatie van een algemene onderwijsgroei met een afnemende milieu-invloed gedurende de schoolloopbaan zorgt voor een afnemend verband tussen sociale herkomst en het uiteindelijk bereikte onderwijsniveau. De oorzaak hiervoor is dat de latere transities in de loopbaan steeds belangrijker worden voor het bereikte eindniveau, omdat steeds meer leerlingen daaraan toe komen, en deze latere transities minder milieu-gevoelig zijn. Het enige dat we dan nog wel aannemelijk moeten maken, is dat aan de condities *a)* en *b)* is voldaan. Dit gaan we nu doen.

Dat er onderwijsexpansie heeft plaatsgevonden (conditie *a)*) mag bekend worden verondersteld. De onderwijsparticipatie heeft in Nederland enorme proporties aangenomen (Idenburg, 1964; Van der Ploeg, 1993; CBS, 1994).

Conditie *b)* zegt dat sociale selectie het sterkst een rol speelt aan het begin van de onderwijs-carrière en afzwakt in de gevorderde fasen van de onderwijsloopbaan. Er zijn twee redenen aan te geven waarom dit zo is (Shavit & Blossfeld, 1993). Het eerste argument luidt dat jongere kinderen sterker afhankelijk zijn van de voorkeuren en materiële omstandigheden van hun ouders dan oudere kinderen. Volgens Müller (1990) zijn scholieren en studenten, naarmate ze ouder worden, in toenemende mate in staat zelf te beslissen wat ze willen en rekenen zij minder op ouderlijke hulpbronnen. Ten tweede zullen kinderen met weinig hulpbronnen de eerste selectiebarrières in een schoolloopbaan slechts met zeer veel moeite overleven, en zullen de overgebleven leerlingen de verdere horden makkelijker kunnen nemen. Zeer weinig arbeiderskinderen komen daarom in de hogere niveaus van het onderwijssysteem terecht, terwijl kinderen uit hogere strata van de samenleving daarentegen gemakkelijker het voortgezet en hoger onderwijs binnenstromen. Het gevolg hiervan is dat bij transities op een later tijdstip in de onderwijsloopbaan de sociale achtergrond steeds minder gecorreleerd is met talent of aanleg en andere leerlingkenmerken – zoals motivatie en doorzettingsvermogen – die succes in het onderwijs bewerkstelligen. Daarom zal volgens Mare (1981) het indirecte effect van sociale herkomst, dat wordt geïntermedieerd door aanleg en motivatie, steeds geringer worden of zelfs helemaal niet meer bestaan, en zal de totale invloed van sociale herkomst klein zijn.

In Nederland is recentelijk een aanzet gegeven tot het bestuderen van onderwijstransities op de manier zoals die door Mare is voorgesteld (Dronkers, 1993; De Graaf & Ganzeboom, 1993). Dronkers onderzoekt twee cohorten geboren in 1953 en 1965 en laat zien dat de milieu-afhankelijkheid van een aantal transities (tot ongeveer 17-jarige leeftijd) niet is veranderd. De Graaf & Ganzeboom menen echter wel aan te tonen dat de invloed van het sociaal milieu op de eerste transities in het voortgezet onderwijs is afgenomen.

Beide studies zijn niet zonder problemen. Dronkers bestudeert slechts transities in de eerste fase van het voortgezet onderwijs en het tijdvak waarvoor hij gegevens heeft bestrijkt niet meer dan twaalf jaar. Zo'n beperkte analyse vermindert de kans op het vinden van significante verschillen. Ook raakt men door slechts een gedeelte van de onderwijsloopbaan te bestuderen het zicht kwijt op de totale invloed van de sociale achtergrond, waardoor een onvolledig beeld ontstaat over het belang van het sociaal milieu in het onderwijs.

De Graaf en Ganzeboom bekijken wel een lange historische periode en complete onderwijsloopbanen, maar omdat zij in hun gegevens slechts de beschikking hadden over het hoogst bereikte onderwijsniveau zijn de transities die zij bestuderen achteraf geconstrueerd. Deze aanpak is twijfelachtig omdat het Nederlandse onderwijsstelsel vele transities kent die niet hiërarchisch zijn. Zo kunnen individuen die in het VWO terecht zijn gekomen daar direct vanuit het basisonderwijs voor hebben gekozen, maar ook de omweg hebben bewandeld via het HAVO. Daarom is het mogelijk dat het ene individu de overgang naar het VWO al op zijn twaalfde heeft gemaakt en het andere op zijn zeventiende. De keuze voor het HBO kan via de route langs het HAVO of VWO zijn gegaan, maar ook via het MAVO en het MBO.

In dit artikel presenteren we analyses op basis van informatie over complete onderwijsloopbanen van 1617 Nederlanders geboren tussen 1928 en 1967. De gegevens zijn retrospectief verzameld en geven een representatief beeld van de milieu-afhankelijkheid van onderwijstransities in de periode tussen ongeveer 1940 en 1990. Dit maakt het mogelijk de invloed van het sociaal milieu in het onderwijs te bestuderen aan de hand van het door Mare voorgestelde loopbaanmodel. We gaan na of de afgenomen samenhang tussen het ouderlijk sociaal milieu en het hoogst bereikte opleidingsniveau van kinderen alleen bewerkstelligd wordt door de toegenomen onderwijsparticipatie of mede het resultaat is van veranderende selectie- en allocatieprocessen. De te beantwoorden onderzoeksvraag luidt daarom:

In hoeverre hangen in Nederland transities gedurende de onderwijsloopbaan af van iemands sociale achtergrond en hoe is deze invloed veranderd in de periode 1940-1990?

De opbouw van het artikel is als volgt. In paragraaf 2 wordt de onderzoeksopzet besproken. Vervolgens worden in paragraaf 3 de analyseresultaten gepresenteerd die een antwoord moeten geven op de door ons gestelde onderzoeksvraag. Paragraaf 4 ten slotte omvat de belangrijkste conclusies van dit onderzoek.

2. ONDERZOEKSOPZET

2.1 Onderwijsloopbaanmodellen

Bij het bestuderen van onderwijstransities gaan we uit van de meest gangbare overgangen die het Nederlandse onderwijsstelsel (figuur 1) kent¹. Iedere van de 14 pijlen in deze figuur stelt een overgang weer en voor elke transitie zullen we de invloed van sociale herkomst schatten. Daarnaast onderzoeken we binnen ieder opleidingsniveau de invloed van het sociaal milieu op de kans een diploma te behalen². In totaal analyseren we daarmee 23 transities, die in bijlage 1 staan beschreven. Voor elk van die 23 transities komen alleen diegenen in aanmerking die de noodzakelijke eerdere transities met succes hebben doorlopen. Daarom noemen we deze transities *conditionele* overgangen.

De structuur van het Amerikaanse onderwijsstelsel, waarvoor Mare zijn model heeft ontwikkeld, is vrij eenvoudig en kan adequaat worden gemodelleerd als een afvalrace. In tegenstelling

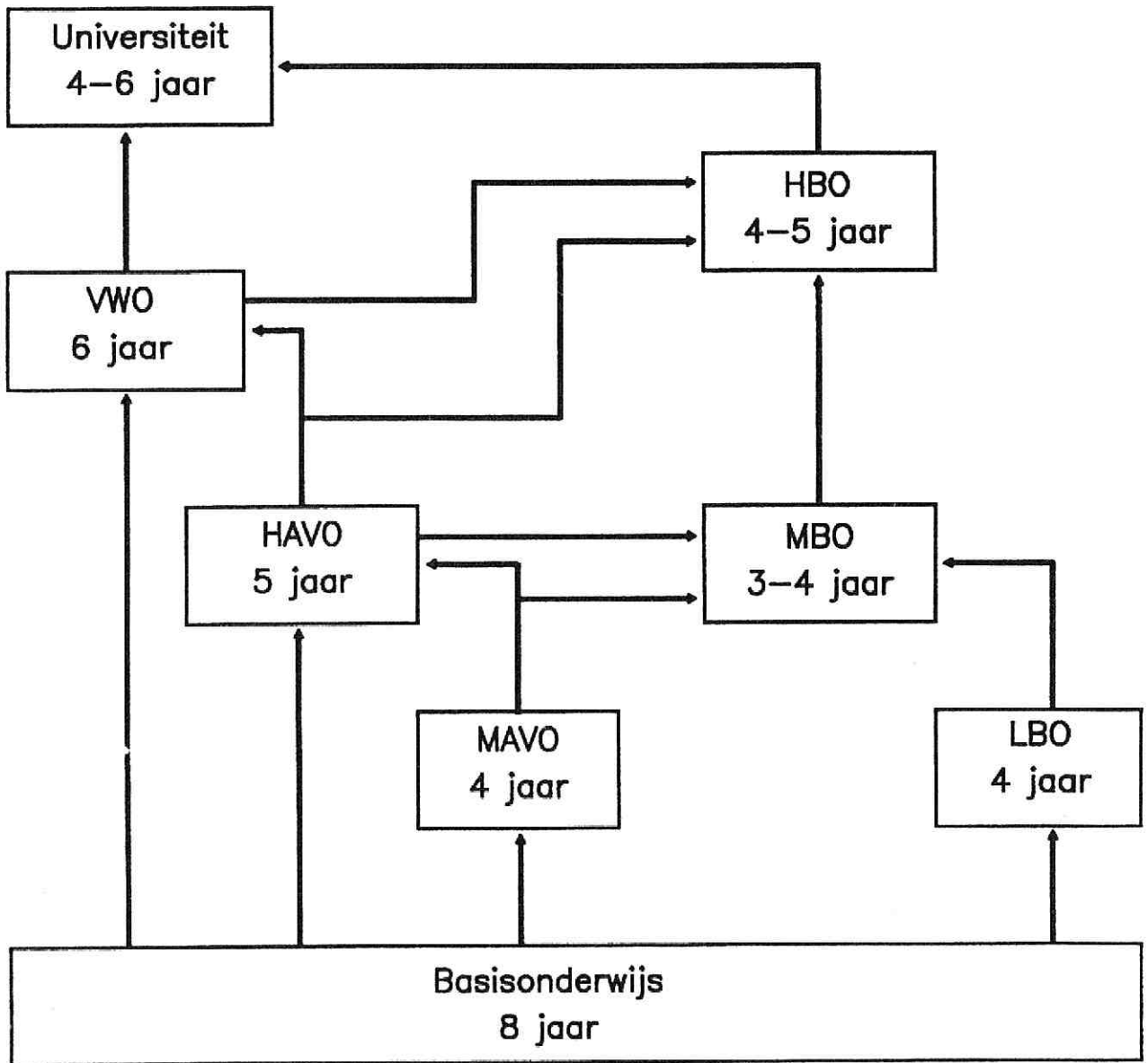


Fig. 1. Het Nederlandse onderwijsstelsel.

(Bronnen: Van der Ploeg (1993), figuur 2.1, p. 7; CBS (1993), p. 445).

tot dit systeem heeft het Nederlandse onderwijsstelsel de eigenschap dat er vaak sprake is van meer dan twee keuzemogelijkheden. In Nederland is bijvoorbeeld de overgang van het basis- naar het voortgezet onderwijs niet eenvoudig als succes of het ontbreken daaraan te beschouwen. Leerlingen hebben de keuze tussen vijf alternatieven: stoppen met het volgen van onderwijs (voor zover het bereiken van de leerplichtige leeftijd dit toelaat), LBO, MAVO, HAVO of VWO, die natuurlijk wel gerangordend zijn. De oplossing die we voor dit probleem hebben gekozen, is om uit één transitie met vijf alternatieven vier conditionele overgangen te maken, namelijk van LO naar LBO+, van LO naar MAVO+, van LO naar HAVO+ en van LO naar VWO. De transitie van LO naar LBO+ houdt in de mogelijke overgang naar een niveau voortgezet onderwijs (LBO of hoger), onder de conditie dat een individu niet is gestopt met het volgen van onderwijs na de basisschool. Van de personen die deze transitie met succes hebben genomen, is vervolgens bepaald of zij gekozen hebben voor LBO of voor een hoger schoolniveau (MAVO, HAVO of VWO). Gegeven dat men voor minimaal MAVO heeft gekozen, wordt de kans bepaald dat deze keuze HAVO of VWO is. Van deze overgebleven selectie mensen

wordt ten slotte nagegaan wie naar het VWO is gegaan. Soortgelijke procedures hebben we gevolgd bij de overgang van MAVO naar MBO of HAVO, bij de overgang van HAVO naar MBO, VWO of HBO, en die van VWO naar HBO of WO.

In eerste instantie analyseren we de invloed van sociale herkomst op elk van de 23 onderwijstransities afzonderlijk. Er zijn echter een drietal problemen verbonden aan zo'n analyse. Allereerst is voor verschillende transities het aantal respondenten dat er voor in aanmerking komt gering. Het statistische onderscheidingsvermogen voor deze transities is beperkt en veel effecten zullen daarom niet significant zijn. Wanneer men afzonderlijke transities bestudeert, verkrijgt men ook geen beeld van de totale invloed van het sociaal milieu. Ten slotte kan niet nagegaan worden of de invloed van sociale herkomst tijdens de schoolloopbaan afneemt.

Aan deze drie tekortkomingen is tegemoet te komen door de afzonderlijke overgangen te stapelen zodat een simultane schatting van alle transities mogelijk wordt (zie Hout, 1989; Raftery & Hout, 1993). Elk individu komt door deze procedure net zo vaak in de analyse voor als hij of zij in aanmerking is gekomen voor een overgang. Deze werkwijze levert een in statistische zin meer krachtige toetsing op, geeft een duidelijk overzicht van het effect van het sociaal milieu in het onderwijs en kan de verandering van het milieu-effect gedurende de onderwijsloopbaan uitdrukken.

We analyseren zulke gestapelde loopbaanmodellen op twee manieren. Om te beginnen hebben we een datamatrix gemaakt waarin alle 23 onderwijstransities gecombineerd zijn. Dit noemen we het totaalmodel. In dit totaalmodel wordt een compleet beeld gegeven van de invloed van sociale herkomst op de voorkomende stromen binnen alle niveaus in het Nederlandse onderwijsstelsel. Daarnaast hebben we een beperkter model opgesteld waarin we ons concentreren op overgangen in de bovenstroom van het voortgezet onderwijs en in het hoger onderwijs. Dit noemen we het bovenstroommodel. Over het voordeel van het bovenstroommodel komen we te spreken nadat we de bevindingen van het totaalmodel hebben behandeld.

2.2 Gegevens

De gebruikte gegevens in dit artikel zijn gebaseerd op de Familie-enquête Nederlandse Bevolking 1992/'93 (Ultee & Ganzeboom, 1993). Doel van deze enquête was het verzamelen van informatie over de samenhangen tussen sociale kenmerken van familieleden, in het bijzonder van echtgenoten of samenwonende partners, maar ook van ouders en kinderen.

Tussen mei 1992 en mei 1993 werden de respondenten geïnterviewd. Hun namen en adressen waren aselekt getrokken uit de bevolkingsregisters van een aantal Nederlandse gemeenten. Deze gemeenten vormden een steekproef uit alle Nederlandse gemeenten die was gestratificeerd naar regio en urbanisatiegraad. In totaal zijn 1000 primaire respondenten ondervraagd en 800 partners. De hier vermelde analyses beperken zich tot personen die geboren zijn tussen 1928 en 1967. Omgerekend naar leeftijd betreft het individuen van 25 tot en met 64 jaar. Respondenten jonger dan 25 jaar zijn buiten de analyses gelaten, omdat ze wellicht hun onderwijsloopbaan nog niet hebben afgesloten. De loopbanen van mensen van 65 jaar en ouder daarentegen zijn niet geanalyseerd vanwege te verwachten onbetrouwbare antwoorden als gevolg van het feit dat een retrospectief onderzoeksdesign een sterke nadruk legt op het geheugen van de respondenten. Het resultaat van deze selectie is een steekproef van 1617 respondenten.

2.3 Operationalisaties

Wanneer we onderwijstransities analyseren, hebben we te maken met een dichotome afhankelijke variabele, die de waarde 0 of 1 heeft. Een 1 duidt op een succesvolle overgang, terwijl een 0 geen succes betekent.

Voor de meting van het ouderlijk sociaal milieu gebruiken we drie indicatoren, namelijk het opleidingsniveau van beide ouders en het beroep van de vader. Het opleidingsniveau van zowel vader als moeder is gemeten door middel van de hoogst bereikte opleiding en wordt als interval variabele gebruikt. Het is als volgt geoperationaliseerd: 1. minder dan LO; 2. LO; 3. LBO; 4. MAVO; 5. MBO; 6. HAVO; 7. VWO; 8. HBO; 9. WO; 10. meer dan WO. Het beroepsniveau van de vader van de respondent is geïndiceerd aan de hand van de voor Nederland ontwikkelde

beroepsprestigeschaal van Sixma & Ultee (1983). Deze schaal geeft het sociaal aanzien weer dat beroepen in de Nederlandse samenleving genieten en heeft een bereik van 13 tot en met 87.

Om verschillen in onderwijskansen tussen mannen en vrouwen te beschrijven, nemen we de variabele sekse mee in de analyses. Bij deze dichotome variabele treedt de categorie man als referentiegroep op.

Veranderingen in het effect van het sociaal milieu over de tijd heen analyseren we door vier geboortecohorten te onderscheiden: 0. 1928-1937; 1. 1938-1947; 2. 1948-1957; 3. 1958-1967.

Het voordeel van de gevolgde procedure van stapelen is dat we in staat zijn eigenschappen van de overgang in de analyse mee te nemen³. Daarom nemen we bij de simultane schatting van onderwijstransities naast de bovengenoemde predictoren nog twee variabelen in de modellen op. Allereerst maken we het onderscheid tussen onderwijstransities die een overgang van het ene naar het andere opleidingsniveau weergeven en overgangen die het behalen van een diploma in een opleiding betreffen, gegeven dat men die opleiding is gaan volgen. Daarnaast nemen we de leeftijd waarop een transitie plaatsvindt, op in ons modellen⁴. Dit doen we om veranderingen in de invloed van het sociaal milieu gedurende de onderwijsloopbaan vast te stellen. In de analyse is de leeftijd verminderd met twaalf jaar, zodat de leeftijd de afwijking in jaren aangeeft ten opzichte van het tijdstip waarop door de bank genomen de overgang van het basis- naar het voortgezet onderwijs plaatsvindt.

2.4 Analysemethode

Om onderwijstransities adequaat te analyseren, maken we gebruik van logistische regressie-analyse. In een logistisch model wordt de logit over een variabele lineair afhankelijk gesteld van een of meer interval variabelen (vergelijk DeMaris, 1992). Onder een logit – of logodds – wordt de natuurlijke logaritme verstaan uit een kansverhouding (= odds). De kans voor individu i op succes noteren we als p_{1i} en de kans op geen succes als p_{0i} . De som van deze kansen is 1. Als we de predictor X als een interval variabele beschouwen, dan is het logistische model als volgt:

$$\ln \frac{p_{1i}}{p_{0i}} = \beta + \beta_{yx} X_i \quad (1)$$

waarin:

$$p_{0i} = 1 - p_{1i} \quad (2)$$

In model (1) wordt een logit voor individu i , de natuurlijke logaritme over de verhouding van de kans dat individu i succes heeft versus de kans dat individu i geen succes kent, lineair afhankelijk gesteld van zijn score op variabele X . De interpretatie van de parameters rechts komt overeen met die in een lineair regressie model. Het intercept (β) is de waarde van de logit bij $X=0$ en β_{yx} is de effectparameter die de verandering van de logit aangeeft per eenheid toename van X . In dit artikel geven we niet de geschatte b -coëfficiënten weer, maar de antilog eruit ($\exp(b)$). Deze waarden geven de voorspelde (multiplicatieve) verandering aan in de odds per eenheid toename van X . De schattingstechniek die gebruikt wordt om de verschillende parameters te berekenen is die van de grootste aannemelijkheid (maximum likelihood). De relatieve sterkte van de onafhankelijke variabelen in het model wordt berekend door het verschil tussen de hoogste en de laagste categorie van een predictor als exponent van de geschatte toename van de voorspelde kansverhouding te nemen.

3. RESULTATEN

In deze paragraaf gaan we alleen uitgebreid in op de analyseresultaten van de simultane schatting van onderwijsloopbanen. De analyse van afzonderlijke onderwijstransities bespreken we niet, want we zijn vooral geïnteresseerd in het totaalbeeld. Echter de resultaten van zo'n afzonderlijke analyse zijn wel gepresenteerd in bijlage 2. In deze bijlage zijn naast de hoofdeffecten van sociaal milieu, sekse en cohort, de significante interactie-effecten tussen deze variabelen weergegeven. De resultaten van het uitgebreide onderwijsloopbaanmodel met een simultane schatting van alle 23 onderwijstransities staan weergegeven in tabel 1.

Uit model 1 van tabel 1 blijkt dat de drie indicatoren voor het sociaal milieu een significante invloed uitoefenen op de waarschijnlijkheid een transitie te overleven. De verhouding van de kans op succes versus de kans op geen succes neemt toe, naarmate het beroepsprestige van de vader en het opleidingsniveau van beide ouders hoger zijn. In relatieve termen geformuleerd is de invloed van vaders opleiding op de odds het grootst. De odds op succes is voor kinderen van de hoogst opgeleide vaders (categorie 10) bijna twee keer zo groot als voor kinderen van de laagst opgeleide vaders (categorie 1), want $1.077^9 = 1.950$. Het overeenkomstige maximale effect van moeders opleiding bedraagt $1.051^9 = 1.565$ en dat van vaders beroep $1.003^{74} = 1.248$.

Tabel 1. Resultaten logistische regressie-analyse met simultane schatting van alle voorkomende onderwijstransities (weergegeven is $\exp(b)$).

model	1.	2.	3.	4.	5.	6.
beroep vader	1.003*	1.003*	1.006**	1.002	1.004*	1.005
opleiding vader	1.077**	1.095**	1.129**	1.114**	1.094**	1.145**
opleiding moeder	1.051*	1.057**	1.147**	1.143**	1.073*	1.267**
sekse	.894*	.853**	.912	.693**	.978	.888
cohort	1.111**	1.122**	1.120**	1.175*	1.122**	1.144
leeftijd transitie		.927**	1.037*	.926**	.927**	1.037*
diploma		2.832**	2.848**	2.843**	2.833**	2.860**
<i>interacties met leeftijd transitie</i>						
beroep vader			.999*			.999*
opleiding vader			.993			.993
opleiding moeder			.985**			.984**
sekse			.984			.986
<i>interacties met cohort</i>						
beroep vader				1.001		1.001
opleiding vader				.990		.991
opleiding moeder				.963		.964
sekse				1.119*		1.136*
<i>interacties met sekse</i>						
beroep vader					.999	.998
opleiding vader					1.001	1.000
opleiding moeder					.971	.953
N	8166	8166	8166	8166	8166	8166
aantal succesvolle transities	5319	5319	5319	5319	5319	5319
model Chi ²	158.099**	582.413**	652.307**	595.008**	583.852**	667.869**
df	5	7	11	11	10	18

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

Bron: Familie-enquête Nederlandse Bevolking 1992/'93

Het significante effect van sekse houdt in dat meisjes door de bank genomen minder succes hebben om transities te overleven dan jongens. De invloed van cohort ten slotte laat zien dat jongere respondenten verder zijn gekomen in hun onderwijsloopbaan dan oudere. Dit is uiteraard niets anders dan de toegenomen onderwijsdeelneming. Vervolgens breiden we het model uit met de leeftijd waarop een bepaalde overgang plaatsvindt en een dummy die het contrast weer-geeft tussen een transitie van het ene opleidingsniveau naar het andere en het behalen van een diploma, gegeven dat een keuze al is gemaakt (model 2). We zien dat de fit aanzienlijk toeneemt. Hoe hoger de leeftijd waarop een overgang plaatsvindt, des te kleiner de odds. Dit laat zien dat de afvalpercentages in het hoger onderwijs groter zijn dan in het voortgezet onderwijs. Daarnaast geldt dat de kansverhouding succes versus geen succes groter is als de onderwijstransitie het behalen van een diploma betreft dan wanneer er sprake is van een overgang van het ene opleidingsniveau naar het andere.

In model 3 zijn interacties van sociaal milieu en sekse met de leeftijd waarop een transitie plaatsvindt, toegevoegd. Alle drie effecten van sociale herkomst blijken gedurende de schoolloopbaan af te nemen. Bij vaders beroep en moeders opleiding is deze afname significant. De leeftijd waarop beslissingen zijn genomen door de respondenten in dit bestand varieert globaal tussen 12 en 25. De invloed van moeders opleiding op de odds een transitie te overleven is het grootst op 12-jarige leeftijd en bedraagt 1.147. Deze invloed op de odds daalt met ruim 11 procent per jaar, want het effect van moeders opleiding is op ongeveer 21-jarige leeftijd uitgedoofd ($1.147 * .985^9 = 1.001$). De invloed van vaders beroepsprestige vertoont een zelfde patroon over de schoolloopbaan. De invloed van vaders beroepsprestige op de kansverhouding is op 12-jarige leeftijd 1.006, terwijl dit effect afneemt met iets meer dan 16 procent per jaar. Het is dus ook verdwenen aan het eind van het voortgezet onderwijs. De interactie van sekse met de leeftijd waarop een transitie plaatsvindt daarentegen laat zien dat de verschillen tussen jongens

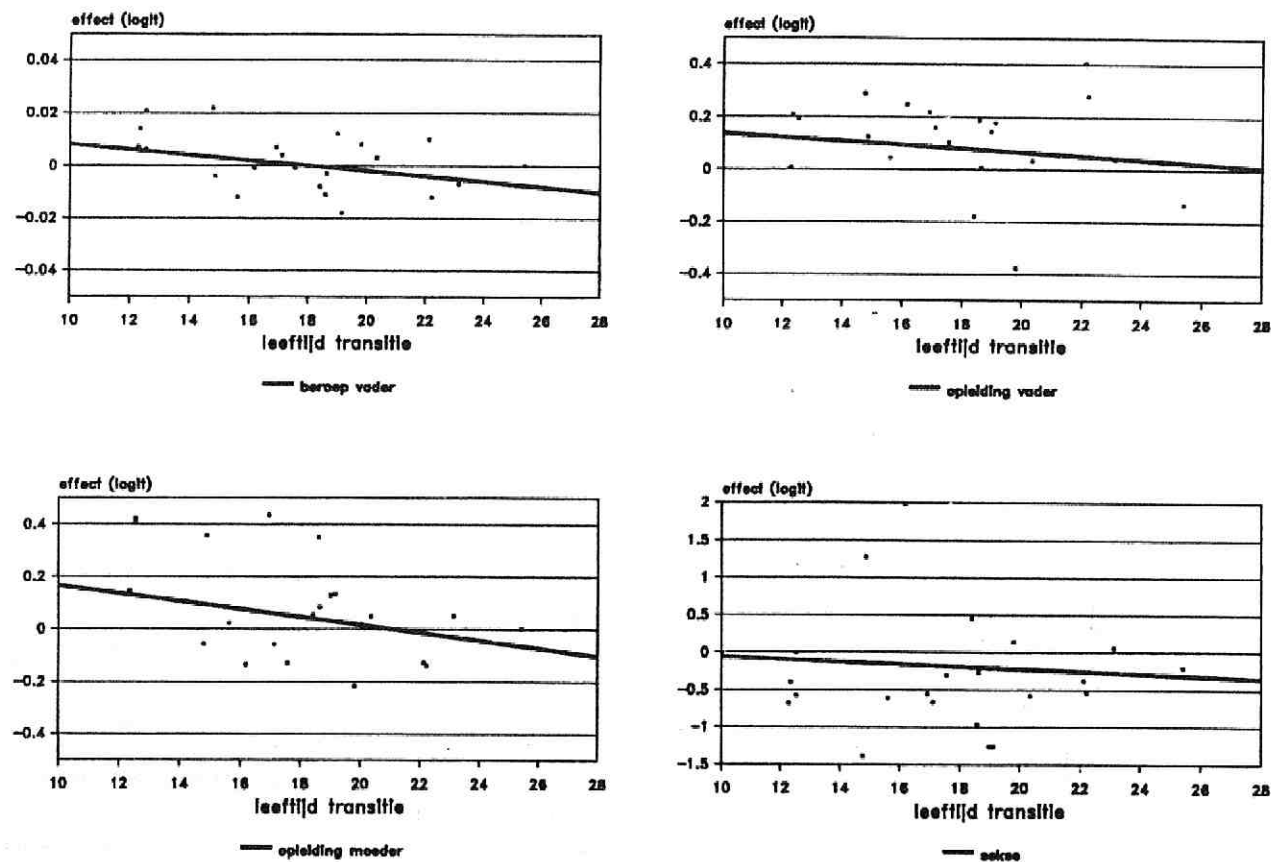


Fig. 2. Veranderingen in de invloed van sociale herkomst en sekse gedurende de onderwijsloopbaan.

en meisjes bij latere transitie groter zijn dan bij eerdere. Hoewel dit interactie-effect niet significant is, lijkt het erop dat de onderwijsachterstand van meisjes gedurende de onderwijsloopbaan groter wordt. In figuur 2 zijn de veranderingen in de effecten van sociale herkomst en sekse gedurende de schoolloopbaan nog eens grafisch weergegeven. De effecten in deze figuur betreffen geen effecten op de odds, maar effecten op de logodds. Het strooidiagram is gebaseerd op de schattingen uit bijlage 2, terwijl de trendlijnen berusten op de lineaire interactie-termen uit tabel 1.

Wanneer we in plaats van interacties met de leeftijd waarop een transitie plaatsvindt interacties met cohort opnemen, verkrijgen we model 4. Uit dit model blijkt dat er geen historische verschuivingen in de invloed van sociale herkomst op onderwijstransities hebben plaatsgevonden. Alleen de interactie van sekse met cohort is significant. De ratio van de odds voor meisjes om succes dan wel geen succes te hebben en de overeenkomstige odds voor jongens is in het oudste cohort .693 en in het jongste .971 ($.693 * 1.119^3$). Met andere woorden: de onderwijsachterstand van meisjes ten opzichte van jongens is zo goed als ingelopen. Uit model 5 blijkt dat er geen enkele significante interactie tussen sociale herkomst en sekse aanwezig is. Model 6 ten slotte vormt het eindmodel waarin alle interacties tegelijkertijd zijn opgenomen. De gevonden resultaten wijken niet af van de eerdere bevindingen. Er is een afname in de invloed van het sociaal milieu gedurende de onderwijsloopbaan, maar er zijn geen historische ontwikkelingen in de herkomsteffecten per transitie geweest.

Het zojuist besproken onderwijsloopbaanmodel heeft als nadeel dat het geen natuurlijke loopbaan betreft. Weliswaar geeft het een compleet overzicht van de invloed van sociale herkomst voor alle mogelijke onderwijstransities, maar het model heeft niet het karakter van een afvalrace, waaraan een bepaald aantal mensen begint en gaandeweg steeds meer individuen afvallen. Het beeld van de onderwijsloopbaan als afvalrace komt van Boudon (1974). Boudon stelde dat een onderwijsloopbaan bestaat uit een aantal transitie – hij noemt het letterlijk breekpunten – waar mensen steeds opnieuw de beslissing moeten nemen onderwijs te blijven volgen of (voorgoed) af te haken. Volgens Boudon valt iedere beslissing uit in het nadeel van leerlingen uit de lagere milieus. Daarnaast werkt het cumulatief; iedere beslissing meer die genomen moet worden is extra nadelig voor kinderen met een lage sociale achtergrond. Voor Nederland heeft Willemse (1987) Boudons simulatiemodel empirisch onderbouwd door voor een zestal breekpunten het effect van sociale herkomst te berekenen. Op een meer geavanceerde wijze bestuderen wij vier onderwijstransities in de bovenstroom van het voortgezet onderwijs en in het hoger onderwijs. De eerste overgang betreft de keuze na het basisonderwijs voor de bovenstroom van het voortgezet onderwijs (HAVO of VWO). Van de leerlingen die deze overgang maakten, bestuderen we vervolgens of ze de gekozen opleiding in de bovenstroom van het voortgezet onderwijs met succes hebben afgesloten. Daarna bekijken we van deze groep leerlingen met een HAVO- of VWO-diploma wie is doorgegaan met een opleiding in het hoger onderwijs (HBO of VWO) en wie zijn onderwijscarrière heeft gestaakt. De vierde transitie ten slotte geeft weer of iemand zijn diploma in het hoger onderwijs al dan niet heeft behaald. In figuur 3 zijn de vier transitie weergegeven inclusief de aantallen doorstromers en afvallers.

De resultaten van dit bovenstroommodel staan weergegeven in tabel 2. De concentratie op een duidelijk afgebakende afvalrace blijkt zinvol te zijn, want de resultaten van deze analyse versterken de bevindingen uit tabel 1. De specificatie van de afzonderlijke modellen uit tabel 2 is overeenkomstig aan die uit tabel 1. Stuk voor stuk leveren de modellen grotere effecten en een betere fit. We bespreken hier slechts de belangrijkste verschillen tussen de modellen uit het totaalmodel en het bovenstroommodel. Ten eerste is het effect van de leeftijd waarop een transitie plaatsvindt, omgedraaid (model 2). In het totaalmodel (zie tabel 1) werd een odds kleiner dan 1 gevonden, terwijl het model voor de bovenstroom van het voortgezet onderwijs en voor het hoger onderwijs een odds groter dan 1 laat zien tussen de leeftijd waarop een overgang plaatsvindt en schoolsucces. Deze positieve samenhang heeft te maken met het meer natuurlijke

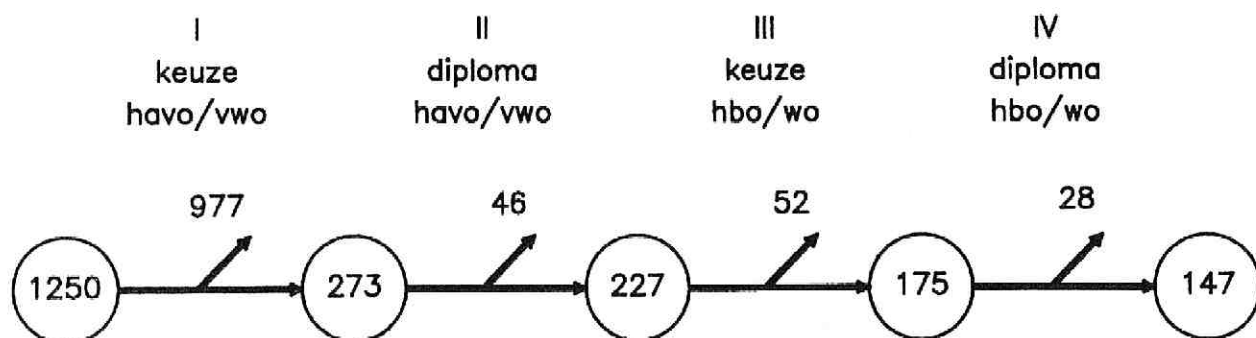


Fig. 3. Onderwijstransities in de bovenstroom van het voortgezet onderwijs en in het hoger onderwijs.

ke karakter van het bovenstroommodel. Eenmaal op weg, behaalt men vaak de eindstreep. Het diploma in het algemeen voortgezet onderwijs bezit namelijk niet zo veel waarde op de arbeidsmarkt. Ten tweede neemt in het bovenstroommodel naast het effect van vaders beroep en moeders opleiding ook de invloed van vaders opleiding significant af gedurende de onderwijsloop-

Tabel 2. Resultaten logistische regressie-analyse met simultane schatting van onderwijstransities in de bovenstroom van het voortgezet onderwijs en in het hoger onderwijs (weergegeven is $\exp(b)$).

model	1.	2.	3.	4.	5.	6.
beroep vader	1.018**	1.014**	1.020**	1.021**	1.017**	1.031**
opleiding vader	1.208**	1.171**	1.298**	1.181*	1.120*	1.261*
opleiding moeder	1.260**	1.197**	1.297**	1.325*	1.225**	1.444**
seks	.764*	.856	.806	.622	.885	.656
cohort	1.481**	1.439**	1.442**	1.892**	1.441**	1.940**
leeftijd transitie		1.259**	1.900**	1.258**	1.260**	1.906**
diploma		3.288**	3.041**	3.250**	3.270**	2.984**
<i>interacties met leeftijd transitie</i>						
beroep vader			.997*			.997*
opleiding vader			.965**			.964**
opleiding moeder			.972*			.973*
seks			1.029			1.031
<i>interacties met cohort</i>						
beroep vader				.997		.996
opleiding vader				.995		.991
opleiding moeder				.951		.960
seks				1.170		1.169
<i>interacties met seks</i>						
beroep vader					.995	.994
opleiding vader					1.085	1.088
opleiding moeder					.957	.946
N	1767	1767	1767	1767	1767	1767
aantal succesvolle transitie	769	769	769	769	769	769
model Chi ²	454.262**	787.677**	853.469**	793.131**	789.997**	860.468**
df	5	7	11	11	10	18

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

Bron: Familie-enquête Nederlandse Bevolking 1992/'93

baan (model 3). Terwijl dit effect bij de overgang van het basis- naar het voortgezet onderwijs nog beduidend is (1.298), op negentienjarige leeftijd – zeven jaar later – is het zo goed als verdwenen ($1.298 * .965^7 = 1.012$). Ten derde blijkt er in de bovenstroom geen significante invloed van sekse te zijn op onderwijssucces. Ook bestaan er tussen de onderscheiden geboortecohorten niet langer verschillen in de ratio van de odds voor jongens op succes versus geen succes en de overeenkomstige odds voor meisjes, terwijl we in het totaalmodel vonden dat meisjes vroeger minder succesvol waren op school.

In het dagelijkse spraakgebruik zijn we vaak niet geïnteresseerd in de invloed van sociale herkomst op de odds een bepaalde overgang te maken. We praten liever over kansen, omdat we hier gemakkelijker iets bij kunnen voorstellen. Daarom laten we tot slot van onze analyse zien wat de modellen betekenen in termen van kansen. We nemen als voorbeeld een van de belangrijkste overgangen in het Nederlandse onderwijsstelsel, namelijk die van het basisonderwijs naar het HAVO of VWO. Dit is de eerste overgang in ons bovenstroommodel. Voor drie milieus van herkomst⁵ bepalen we de historische ontwikkeling in de waarschijnlijkheid van een overgang naar het HAVO of hoger. De ontwikkelingen in de voorspelde kansen zijn weergegeven in figuur 4⁶. Ze worden geplaatst in het perspectief van een natuurlijke groeicurve. Het logistische kansmodel is gebaseerd op deze groeicurve. De figuur bestrijkt een langer tijdvak dan de geboortecohorten die wij in dit artikel bestuderen. Hierdoor worden de eigenschappen van de natuurlijke groeicurve duidelijk. Een van deze eigenschappen is dat de sterkte van de toename in de kans afhangt van de grootte van de kans. Wanneer de kans zeer lage of zeer hoge waarden aanneemt, stijgt deze maar weinig. Daarentegen is de toename het sterkst als de kans 50 procent is.

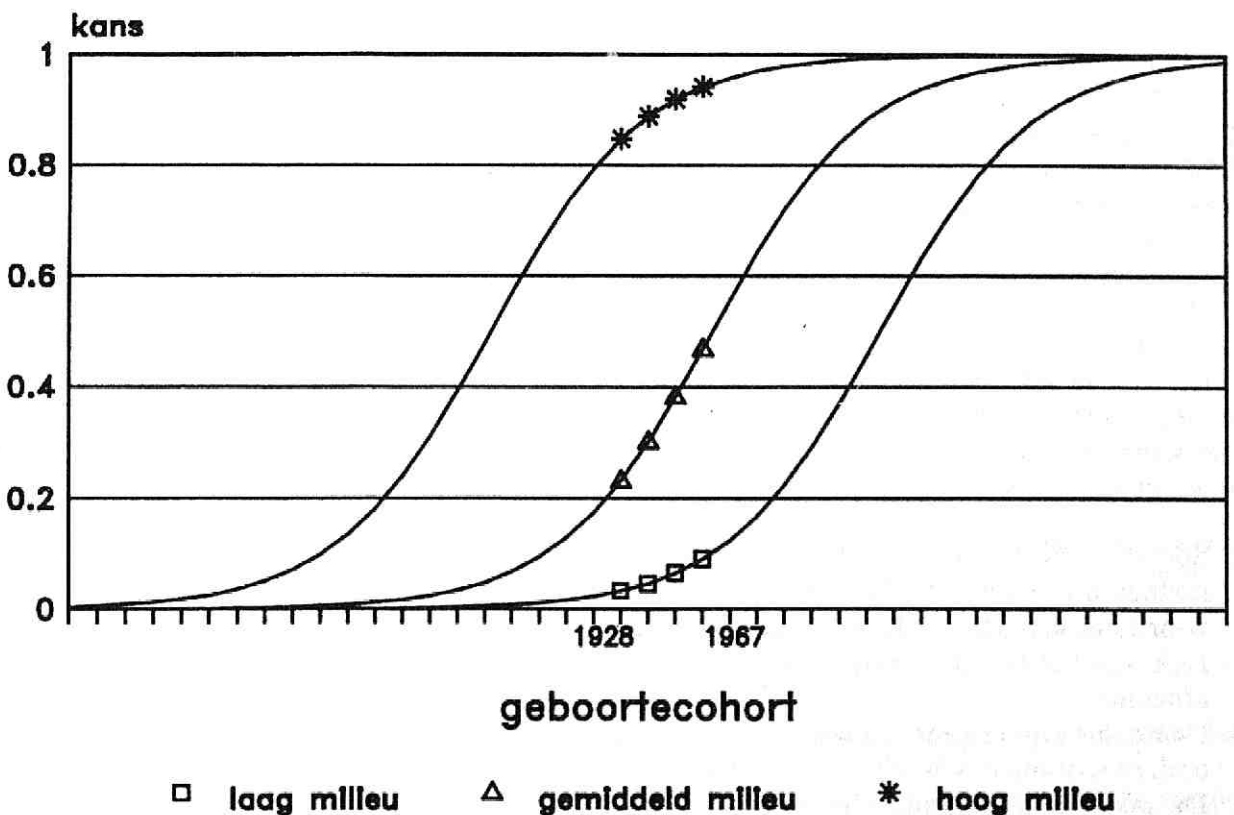


Fig. 4. De ontwikkeling in de voorspelde kans op succes ten aanzien van de overgang van het basisonderwijs naar het HAVO of VWO (N=1139).

In figuur 4 valt direct op dat er binnen ieder cohort grote verschillen bestaan tussen de drie onderscheiden milieus wat betreft hun kans naar het HAVO of VWO te gaan. In het oudste cohort (1928-37) bijvoorbeeld is de kans voor kinderen uit een laag milieu 3 procent, uit een gemiddeld milieu 23 procent en uit een hoog milieu 84 procent. Verder laat de figuur zien dat het percentage leerlingen dat kiest voor het HAVO of VWO is toegenomen. In veertig jaar tijd is dit percentage voor jongens en meisjes afkomstig uit een laag milieu van 3 naar 9 procent gestegen, voor kinderen van gemiddelde komaf is het toegenomen van 23 naar 47 procent en voor leerlingen met een hoge sociale achtergrond is de voorspelde kans op succes gegroeid van 84 naar 94 procent. De drie curves hebben precies dezelfde vorm; ze kunnen op elkaar worden gelegd, maar bevinden zich voor een gegeven cohort in een verschillende fase.

Hoewel het er op het eerste gezicht op lijkt dat de kinderen uit de middengroep er de laatste decennia het meest op vooruit zijn gegaan, en dat er dus iets is veranderd in de invloed van sociale herkomst op een succesvolle transitie, is deze uitspraak bedriegelijk. Dit komt voort uit een van de eigenschappen van de natuurlijke groeicurve, namelijk dat de groei het sterkst is in het middentraject rond de 50 procent. Omdat de kans dat een leerling uit de middengroep de overgang naar het HAVO of VWO maakt zich juist beweegt in dit middentraject, is voor hem de toename in de overgangswaarschijnlijkheid het grootst. De odds ratio houdt echter rekening met de bodem- en plafondeffecten voor respectievelijk het laag en hoog milieu en laat zien dat er geen ontwikkelingen in de *relatieve* kansen hebben plaatsgevonden. Inhoudelijk betekent dit dat de groei in onderwijsparticipatie verschillend verloopt tussen de sociale milieus, maar dat de selectie- en allocatieprocessen in het onderwijs *niet* zijn veranderd.

4. CONCLUSIES EN DISCUSSIE

Uitgangspunt van dit artikel wordt gevormd door de tegenstrijdige bevindingen omtrent veranderingen in de relatie tussen het ouderlijk sociaal milieu en onderwijssucces van kinderen in Nederland. Dat het hier een schijnbare tegenspraak betreft, is in het begin van de jaren tachtig aangetoond door Mare. Volgens Mare zal door de toegenomen onderwijsdeelname sinds de Tweede Wereldoorlog de samenhang tussen het ouderlijk sociaal milieu en het hoogst bereikte opleidingsniveau van kinderen zijn afgenomen, wanneer er per transitie niets is veranderd. De onderwijsloopbanen voor jongere geboortecohorten zijn gemiddeld genomen een stuk langer geworden en daarom kunnen ouders de schoolloopbanen van hun kinderen steeds minder beïnvloeden. Bij deze structurele verklaring hoeft er geen sprake te zijn van historische verschuivingen in de selectieregels. Puur de onderwijsexpansie is verantwoordelijk voor de afgenomen samenhang tussen sociale herkomst en het hoogst bereikte diploma.

Daarbovenop is het mogelijk dat er een proces naar meer meritocratie heeft plaatsgevonden en dat ook de selectieregels zelf zijn veranderd. Om duidelijkheid te krijgen in de invloed van beide processen, hebben we schoolloopbaanmodellen ontwikkeld waarin de processen van elkaar kunnen worden onderscheiden. De analyseresultaten van deze loopbaanmodellen kunnen als volgt worden samengevat:

- a. Sociale herkomst – in het bijzonder het ouderlijk opleidingsniveau – oefent een positieve invloed uit op onderwijssucces bij allerlei transities in Nederlandse schoolloopbanen, maar voor lang niet alle onderwijstransities in even sterke mate. De grootste effecten bevinden zich aan het begin van de onderwijsloopbaan, terwijl de invloed gedurende de loopbaan afneemt.
- b. Daarnaast zijn er geen aanwijzingen gevonden dat de invloed van de sociale achtergrond op onderwijstransities is afgenomen in de periode 1940-1990.
- c. De sekse-ongelijkheid in het onderwijs is sterk afgenomen in de afgelopen decennia. Jongens hadden over het algemeen meer succes op school dan meisjes en hun voorsprong leek in het verleden nog toe te nemen gedurende de onderwijsloopbaan. Tegenwoordig vertonen schoolloopbanen van jongens en meisjes geen verschillen meer.

Deze bevindingen komen met uitzondering van punt *b*. overeen met de onderzoeksuitslagen van De Graaf & Ganzeboom (1993). Dit punt is echter wel zeer cruciaal. In tegenstelling tot onze uitkomsten vonden De Graaf en Ganzeboom wel een afname in de invloed van het sociaal milieu op onderwijstransities over de tijd heen. Hoewel in die studie kunstmatige transities zijn bestudeerd, hadden zij wel veel meer data (11244 respondenten) ter beschikking over een veel langere periode (geboortecohorten van 1891 tot 1960). Misschien vinden wij geen trend, omdat deze of vóór de door ons bestudeerde periode heeft plaatsgevonden, of omdat het statistische onderscheidingsvermogen te gering is.

Hoe dan ook, de belangrijkste consequentie van onze resultaten is dat niet zozeer de selectie- en allocatieprocessen in het onderwijs zijn veranderd, maar dat de afgenomen samenhang tussen sociale herkomst en het hoogst bereikte onderwijsniveau wordt verklaard door de toegenomen onderwijsdeelname. Als gevolg van het feit dat iedereen in Nederland in *absolute* zin een stukje meer opleiding heeft genoten sinds de Tweede Wereldoorlog, lijkt het alsof de onderwijsongelijkheid is afgenomen. Dit is niet zo, want de kansverhouding op al dan geen succes bij onderwijstransities gedurende de schoolloopbaan is tussen sociale milieus niet verminderd.

NOTEN

1. In de onderzochte periode hebben verschillende veranderingen in het onderwijssysteem plaatsgevonden. In het geval dat een respondent een opleiding opgaf die niet meer (onder dezelfde naam) bestaat, is de overeenkomstige hedendaagse opleiding gecodeerd. Veel voorkomende combinaties zijn (M)ULO/3 jaar HBS-MAVO, MMS-HAVO en GYMNASIUM/ATHENEUM/HBS-VWO.
2. Behalve de in de afbeelding weergegeven onderwijsniveaus en overgangen hebben we de transitie van wetenschappelijk naar post-wetenschappelijk onderwijs (promotieplaats, specialisaties) en succes daarbinnen bestudeerd.
3. Daarnaast lijkt het meenemen van eigenschappen van de overgang noodzakelijk om zo een mogelijke afhankelijkheid tussen de waarnemingen (autocorrelatie) te voorkomen. In de vakliteratuur wordt hier echter nooit op gewezen.
4. Wanneer de onderwijstransitie het behalen van een diploma betrof, is als leeftijd genomen het midden tussen de leeftijd waarop men de opleiding is begonnen en die waarop zij is afgesloten.
5. De drie milieus zijn als volgt: laag milieu (laagste beroepsprestige vader, bijvoorbeeld vuilnisman; opleidingsniveau vader en moeder LO), gemiddeld milieu (gemiddelde beroepsprestige vader, bijvoorbeeld boekhouder; opleidingsniveau vader en moeder MBO) en hoog milieu (hoogste beroepsprestige vader, bijvoorbeeld rechter; opleidingsniveau vader en moeder WO).
6. Om te bepalen hoe de kans op succes afhankelijk is van *X*, moeten we het logitmodel uit (1) omzetten in een kansmodel. Af te leiden valt dat de kans dat individu *i* succes heeft:

$$p_{1i} = \frac{1}{e^{-(\beta + \beta_{yx}X)} + 1} \quad (3)$$

LITERATUUR

- Bakker, B., & Cremers, P. (1994). Gelijke kansen in het onderwijs? Een vergelijking van vier cohorten leerlingen in hun overgang naar het voortgezet onderwijs. *Tijdschrift voor Onderwijsresearch*, 19, 191-203.
- Bakker, B., & Schouten, S. (1991). Trends in onderwijskansen. *Sociale Wetenschappen*, 34, 1-21.
- Boudon, R. (1974). *Education, opportunity, and social inequality*. New York: Wiley.
- Centraal Bureau voor de Statistiek (1993). *Statistisch Jaarboek 1993*. Den Haag: SDU.
- Centraal Bureau voor de Statistiek (1994). *Vijfennegentig jaren statistiek in tijdreeksen 1899-1994*. Den Haag: SDU.
- DeMaris, A. (1992). *Logit modeling. Practical applications*. (Sage University Papers series on Quantitative Applications in the Social Sciences, 07-086). Newbury Park, California: Sage.
- Dronkers, J. (1983). Have inequalities in educational opportunity changed in the Netherlands? A review of empirical evidence. *Netherlands' Journal of Sociology*, 19, 133-150.

- Dronkers, J. (1993). Educational reform in the Netherlands. Did it change the impact of parental occupation and education? *Sociology of Education*, 66, 262-277.
- Dronkers, J., & Graaf, P. de (1995). Ouders en het onderwijs van hun kinderen. In J. Dronkers & W. Ultee (red.), *Verschuivende ongelijkheid in Nederland*. Assen: Van Gorcum, 46-66.
- Faasse, J., Bakker, B., Dronkers, J., & Schijf, H. (1987). The impact of educational reform. Empirical evidence from two Dutch generations. *Comparative Education*, 23, 261-278.
- Ganzeboom, H. (1984). Causal models for intergenerational transmission of social inequality in the Netherlands in 1958 en 1977. In B. Bakker, J. Dronkers & H. Ganzeboom (eds.), *Social stratification and mobility in the Netherlands*. Amsterdam: SISWO, 109-122.
- Ganzeboom, H., & Graaf, P. de (1989). Verandering van onderwijskansen in Nederland tussen 1900 en 1980. In I. Gadourek & J. Peschar (red.), *De open samenleving*. Bockaflevering Mens en Maatschappij, Deventer: Van Loghum Slaterus, 58-78.
- Graaf, P. de, & Ganzeboom, H. (1993). Family background and educational attainment in the Netherlands for the 1891-1960 birth cohorts. In Y. Shavit & H-P Blossfeld (eds.), *Persistent inequality. Changing educational attainment in thirteen countries*. Boulder/Summertown: Westview Press, 75-100.
- Hout, M. (1989). *Following in father's footsteps. Social mobility in Ireland*. Cambridge, MA: Harvard University Press.
- Idenburg, Ph. (1964). *Schets van het Nederlandse schoolwezen*. Groningen: Wolters.
- Mare, R. (1979). Social background composition and educational growth. *Demography*, 16, 55-71.
- Mare, R. (1980). Social background and school continuation decisions. *Journal of the American Statistical Association*, 75, 295-305.
- Mare, R. (1981). Change and stability in educational stratification. *American Sociological Review*, 46, 72-87.
- Müller, W. (1990). *Does education matter? Evidence from cross-national comparisons*. University of Mannheim, manuscript.
- Ploeg, S. van der (1993). *The expansion of secondary and tertiary education in the Netherlands*. Nijmegen: ITS.
- Raftery, A., & Hout, M. (1993). Maximally maintained inequality. Expansion, reform, and opportunity in Irish education, 1921-75. *Sociology of Education*, 66, 41-62.
- Shavit, Y., & Blossfeld, H-P. (eds.) (1993). *Persistent inequality. Changing educational attainment in thirteen countries*. Boulder/Summertown: Westview Press.
- Sixma, H., & Ultee, W. (1983). Een beroepsprestigeschaal voor Nederland in de jaren tachtig. *Mens en Maatschappij*, 58, 360-382.
- Ultee, W., & Ganzeboom, H. (1993). *Familie-enquête Nederlandse Bevolking 1992/'93* (databestand). Nijmegen: Vakgroep Sociologie Katholieke Universiteit.
- Vrooman, J., & Dronkers, J. (1986). Changing educational attainment processes. Some evidence from the Netherlands. *Sociology of Education*, 59, 69-78.
- Willems, P. (1987). Overleving in de bovenstroom van het voortgezet onderwijs. Veranderde milieu-effecten voor jongens en meisjes. *Tijdschrift voor Onderwijsresearch*, 12, 329-343.

Manuscript ontvangen: 30-6-1995

Definitief aanvaard: 28-9-1995

Bijlage 1. Overzicht van de bestudeerde onderwijstransities.

no.	transitie	verklaring
1.	lo - lbo+	transitie van LO naar LBO, MAVO, HAVO of VWO
2.	lo - mavo+	transitie van LO naar MAVO, HAVO of VWO - gegeven minimaal LBO-niveau
3.	lo - havo+	transitie van LO naar HAVO of VWO - gegeven minimaal MAVO-niveau
4.	lo - vwo	transitie van LO naar VWO - gegeven minimaal HAVO-niveau
5.	lbo - mbo	transitie van LBO naar MBO
6.	mavo - mbo+	transitie van MAVO naar MBO of HAVO
7.	mavo - havo	transitie van MAVO naar HAVO - gegeven minimaal MBO-niveau
8.	mbo - hbo	transitie van MBO naar HBO
9.	havo - mbo+	transitie van HAVO naar MBO, VWO of HBO
10.	havo - vwo+	transitie van HAVO naar VWO of HBO - gegeven minimaal MBO-niveau
11.	havo - hbo	transitie van HAVO naar HBO - gegeven minimaal VWO-niveau
12.	vwo - hbo+	transitie van VWO naar HBO of WO
13.	vwo - wo	transitie van VWO naar WO - gegeven minimaal HBO-niveau
14.	hbo - wo	transitie van HBO naar WO
15.	wo - wo+	transitie van WO naar WO+
16.	diploma lbo	kans op succes binnen LBO
17.	diploma mavo	kans op succes binnen MAVO
18.	diploma mbo	kans op succes binnen MBO
19.	diploma havo	kans op succes binnen HAVO
20.	diploma vwo	kans op succes binnen VWO
21.	diploma hbo	kans op succes binnen HBO
22.	diploma wo	kans op succes binnen WO
23.	diploma wo+	kans op succes binnen WO+

Bijlage 2. Resultaten logistische regressie-analyse van afzonderlijke onderwijstransities op sociale achtergrond + significante interacties (weergegeven is $\exp(b)$).

	a.	b.	c.	d.	e.	f.	g.	h.	i.	j.	k.	l.	m.	n.	o.	p.
1.	1.006	2.866**	1.500	.992	1.794**								1207	1142	101.117**	5
2.	1.021**	1.213**	1.524**	.563*	.869				1.490**				1142	695	225.923**	6
3.	1.014**	1.232**	1.160*	.672*	1.260**								695	291	122.983**	5
4.	1.007	1.007	1.139	.508**	.870								291	174	15.952**	5
5.	1.007	1.245*	1.545**	.574**	1.556**								477	175	74.861**	5
6.	1.004	1.171	.944	.516*	2.000**								320	211	41.409**	5
7.	.999	1.107	.879	.739	1.745**								211	63	9.962	5
8.	1.003	1.034	1.050	.561*	.942								372	81	6.441	5
9.	.989	1.205	1.421	.380	1.084								162	144	10.219	5
10.	.997	1.008	1.088	.766	.742								144	102	2.785	5
11.	.992	.835	1.057	1.575	.185*								102	88	12.800*	5
12.	1.012	1.155	1.137	.283**	2.048**								187	164	23.439**	5
13.	.982	1.195	1.142	.283**	1.248								164	109	21.013**	5
14.	1.010	1.496*	.881	.686	2.932*	.831*							259	32	8.160	6
15.	1.001	.875	1.001	.818	.846								103	14	1.862	5
16.	1.022	1.335**	.945	.248*	1.371	.986*						1.791*	650	503	24.836**	7
17.	.996	1.132	1.429	3.566**	2.607**			.828*	.635*				476	354	24.646**	7
18.	1.008	.686**	.804**	1.155	.405**	1.148*							495	379	29.378**	6
19.	.999	1.281*	.874	7.271**	1.013						.736*		215	159	10.653	6
20.	.988	1.044	1.024	.543	1.290								206	163	6.102	5
21.	.988	1.325*	.870	.583	1.332				1.892*				290	219	20.628**	7
22.	.993	1.039	1.053	1.062	.714								172	110	5.025	5
23.	1.465	4.110	.004	.000	20.398								18	16	9.785	5

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

Legenda: 1. lo - lbo+; 2. lo - mavo+; 3. lo - havo+; 4. lo - vwo; 5. lbo - mbo; 6. mavo - mbo+; 7. mavo - havo; 8. mbo - hbo; 9. havo - mbo+; 10. havo - vwo+; 11. havo - hbo; 12. vwo - hbo+; 13. vwo - wo+; 14. hbo - wo; 15. wo - wo+; 16. diploma lbo; 17. diploma mbo; 18. diploma mavo; 19. diploma havo; 20. diploma vwo; 21. diploma hbo; 22. diploma wo; 23. diploma wo+; a. beroep vader; b. opleiding vader; c. opleiding moeder; d. sekse; e. cohort; f. beroep vader*cohort; g. opleiding vader*cohort; h. opleiding moeder*cohort; i. sekse*cohort; j. beroep vader*sekse; k. opleiding vader*sekse; l. opleiding moeder*sekse; m. totaal aantal transitities; n. aantal succesvolle transitities; o. model χ^2 ; p. aantal vrijheidsgraden

Bron: Familie-enquête Nederlandse Bevolking 1992/93

De factorstructuur van schrijfvaardigheid: een toepassing van multilevel analyse¹

Hans Kuhlemeier* en Huub van den Bergh**
Instituut voor Toetsontwikkeling (Cito)

ABSTRACT

In the Dutch Educational Assessment Program, the students' language proficiency is measured in grade 9 (at age 15). The writing performance of the students is measured by means of several performance-based writing tasks. Each task is rated on a variety of aspects such as content, style, organisation, punctuation, spelling and grammar. As a consequence, national performance levels are reported in a rather detailed way. This article focuses on the efficiency of the measurement and the reporting of writing performance. Multilevel factor analysis (MLFA) was used to determine dimensionality and reliability. It appears that the between and within-school factor structures are not identical. At the school level, a distinction between two factors is warranted: composition versus technical conventions. At the student level, the composition factor remains intact. However, the technical conventions factor has to be divided into three subskill factors: punctuation, spelling and a factor that is interpreted as the technical quality at the sentence level. Furthermore, it is found that the factor structure is not independent of the writing task. It is also shown that, in general, functional writing proficiency is measured relatively unreliably at the school, student, and score level. It is concluded that the potential usefulness of national assessment for the purpose of school effectiveness studies is limited. Recommendations are made concerning the instrumentation of writing performance, the sample and administration design, the statistical analysis and the reporting of national assessment data.

INLEIDING

Schrijfvaardigheid vormt een essentieel onderdeel van de recente taalpeilingen van het onderwijsniveau in het Nederlandse taalgebied (o.a. Zwarts, 1990; Sytstra, 1993) en daarbuiten (o.a. Conry & Jerolski, 1980; Neville, 1988). In deze taalpeilingen wordt de taalvaardigheid van de leerlingen op een gedetailleerde wijze beschreven. Voor het onderdeel schrijfvaardigheid wordt niet volstaan met één schrijftaak. In de praktijk van het schrijfonderwijs moeten leerlingen immers verschillende teksttypen leren schrijven. De leerlingen krijgen dan ook verschillende schrijfp opdrachten voorgelegd, zoals opstellen, brieven en formulieren. Een andere reden om meer schrijfp opdrachten in een peiling op te nemen is dat de schrijfp prestaties tot op zekere hoogte afhankelijk zijn van de schrijfp opdracht die men de leerling voorlegt. Deze zogeheten taakeffecten verklaren veelal een flink deel van de variantie in schrijfp prestaties (Breland, Camp, Jones, Morris & Rock, 1987; Van den Bergh, 1988; Kuhlemeier & Van den Bergh, 1991; Schoonen, 1991; Linn & Burton, 1994).

In de taalpeilingen worden de schrijfp prestaties van de leerlingen door getrainde beoordelaars op een groot aantal aspecten beoordeeld. Geschreven teksten moeten immers niet alleen voldoen aan eisen op het terrein van inhoud, stijl en organisatie. Ook de meer technische aspecten, zoals interpunctie, spelling en grammatica, vormen belangrijke kwaliteitscriteria.

1. Met dank aan G.J. Mellenbergh en J.M. Wijnstra voor hun commentaar op een eerdere versie van dit artikel en aan H. Goldstein voor zijn adviezen bij de data-analyse.

*Adres: Instituut voor Toetsontwikkeling (Cito), Bedrijfsgroep Peilingsonderzoek, Nieuwe Oeverstraat 65, 6801 MG Arnhem.

**Adres: Universiteit Utrecht, Centre for Language and Communication, Trans 10, 3512 JK Utrecht.

Het grote aantal opdrachten en beoordelingsaspecten maakt een uitgebreide en genuanceerde beschrijving van het niveau van het schrijfonderwijs mogelijk. Naast voordelen van psychometrische aard (vgl. Beaton, 1988; Wijnstra, 1988) zijn er echter ook nadelen aan verbonden. Een daarvan is de bewerkelijkheid van de verzameling, verwerking, analyse en rapportage van de prestaties voor de vele opdrachten en beoordelingsaspecten. Een ander nadeel betreft de (on)toegankelijkheid van de vaak omvangrijke en gedetailleerde rapportage, met name voor niet-onderzoekers (Gipps & Goldstein, 1983). Veelal is het aantal rapportage-eenheden even groot als het aantal schrijfoopdrachten maal het aantal beoordelingsaspecten¹. Of dit wel zo efficiënt is, valt niet zonder meer te zeggen. Daarvoor is nader onderzoek naar de dimensionaliteit van de meting van schrijfvaardigheid noodzakelijk.

De vraag naar de dimensionaliteit van de meting heeft ook een pendant op schoolniveau. Een neven doel van peilingsonderzoek is het ondersteunen van scholen bij het evalueren van hun onderwijs. Zo ontvingen scholen in het kader van een recente peiling in het voortgezet onderwijs (Kuhlemeier, Van den Bergh, Notté, Wagenaar, Verstralen & Cappers, 1994) naast de normale leerlingrapportage ook een zogeheten schoolrapportage. Dit verschaft de scholen een gedetailleerd beeld van de effectiviteit van hun onderwijs in vergelijking met een landelijke referentiegroep van scholen die in vergelijkbare omstandigheden verkeren (qua opleidingstype, instroom, e.d.). De vraag naar de dimensionaliteit van de meting op schoolniveau heeft implicaties voor deze schoolrapportage. Zouden aan de vele vakonderdelen i.c. rapportage-eenheden slechts enkele latente factoren ten grondslag liggen, dan kan het schoolrapport aanmerkelijk worden vereenvoudigd.

Multilevel factoranalyse

De dimensionaliteit van schrijfvaardigheid is tot op heden vrijwel uitsluitend geanalyseerd met behulp van technieken van unilevel analyse, zoals exploratieve factoranalyse of toetsende factoranalyse via covariantie-structuuranalyse (Brelan e.a., 1987; Van Gelderen, 1987; Van den Bergh, 1988; Kuhlemeier & Van den Bergh, 1991; Sytstra & Zwarts, 1991; Zwarts & Rijlaarsdam, 1991). Recentelijk is een nieuwe techniek beschikbaar gekomen waarmee de dimensionaliteit van de data op meer analyseniveaus simultaan bestudeerd kan worden: multilevel factoranalyse (Goldstein & McDonald, 1988; Raudenbush, Rowan & Kang, 1991; Longford & Muthén, 1992; Muthén, 1992). In onze toepassing ervan analyseren wij de factorstructuur van schrijfvaardigheid simultaan op het tussen-scholenniveau en het binnen-scholenniveau. Voor het gebruik van multilevel factoranalyse in ons onderzoek zijn twee hoofdargumenten aan te voeren.

Het eerste argument is statistisch van aard. In een unilevel factoranalyse wordt doorgaans verondersteld dat de data zijn verzameld als een aselechte steekproef uit de populatie en dat de waarnemingen statistisch gezien onafhankelijk zijn. In peilingsonderzoek hebben we echter te maken met een getrapte of clustersteekproef. Eerst worden bijvoorbeeld scholen getrokken, en pas daarna leerlingen binnen scholen. Omdat leerlingen van één school ten gevolge van selectie- en onderwijsfactoren meer gemeenschappelijk hebben dan leerlingen van verschillende scholen wordt de aanname van onafhankelijke waarnemingen geweld aangedaan; de precisie van een clustersteekproef is daarmee lager dan die van een enkelvoudige steekproef van dezelfde omvang (Kish, 1965). Onder constanthouding van onder meer het aantal leerlingen per cluster wordt deze overschatting van de precisie ernstiger naarmate de intraklassecorrelatie toeneemt (Tate & King, 1994). In het geval van een tweetrapssteekproef met leerlingen binnen scholen is de intraklassecorrelatie gedefinieerd als de ratio van de variantie op het schoolniveau tot die van de totale variantie (Kish, 1965). In de praktijk van peilingsonderzoek doen zich vaak aanzienlijke intraklassecorrelaties voor (De Gloppe, 1988; Kuhlemeier & Van den Bergh, 1989; Rymenans, Leroy & Daems, 1991; Tate & King, 1994). Unilevel analyse geeft dan te optimistische schattingen van de nauwkeurigheid van de populatie-indices en leidt te snel tot significantie van de geanalyseerde relaties tussen bijvoorbeeld leerprestaties en schoolkenmerken (Goldstein, 1987).

Het tweede argument is van inhoudelijke aard. Normaliter wordt verondersteld dat de factor-

structuur op alle niveaus dezelfde is, maar dit is zeker geen wet van Meden en Perzen. Soms blijkt een factor wel aantoonbaar op leerlingniveau, maar niet op het hogere niveau van de school (b.v. Longford, 1990; Balke, 1991). Er zijn dan wel systematische verschillen tussen leerlingen (binnen scholen), maar het gemiddelde niveau varieert niet van school tot school. Ook de samenhang tussen de factoren kan van niveau tot niveau verschillen (voor voorbeelden, zie: Hox, 1993; Melse & Kuhlemeier, 1993). De interpretatie van tussen- en binnen-groepen factorstructuren in onderzoek naar schooleffectiviteit is een nog onontgonnen gebied. Het ligt echter voor de hand de binnen-scholen-structuur te interpreteren als de (cognitieve of mentale) structuur bij leerlingen, analoog aan de gangbare interpretatie van de uitkomsten van een unilevel factoranalyse. De tussen-scholen-structuur verwijst daarentegen naar een samengesteld effect van verschillen tussen scholen in de selectie van leerlingen, het leerstofaanbod en de effectiviteit van het gegeven onderwijs (vgl. Longford & Muthén, 1992). Vooral deze tussen-scholenstructuur is van groot belang voor het bestuderen van de effectiviteit van de instructie en de school.

Vraagstelling

In dit artikel wordt de dimensionaliteit i.c. efficiëntie van de meting van schrijfvaardigheid in peilingsonderzoek aan de orde gesteld. Om bovengenoemde redenen maken we daarbij gebruik van multilevel factoranalyse. De data zijn ontleend aan de proefpeiling Nederlands (Kuhlemeier & Van den Bergh, 1989). In dat onderzoek kregen derdeklassers uit het voortgezet onderwijs verschillende schrijftaken voorgelegd (waarvan een leerling er hooguit twee maakte). Van elk van deze taken werden de leerlingprodukten door getrainde jury's op twaalf aspecten beoordeeld, zoals inhoud, stijl, organisatie, interpunctie, spelling en grammatica. Overeenkomstig het afname-ontwerp zijn slechts twee schrijftaken - een excuusbrief en een sollicitatiebrief - aan dezelfde substeekproef van leerlingen voorgelegd. Om deze reden maken we in deze studie alleen gebruik van deze twee schrijftaken. Beide zijn op (dezelfde) tien aspecten beoordeeld. In dit artikel stellen we de vraag naar de empirische onderscheidbaarheid van de beide opdrachten en de tien aspecten op school- en leerlingniveau. Deze vraag naar de dimensionaliteit van schrijfvaardigheid valt uiteen in vijf deelvragen:

- 1 Hoeveel verschillende factoren moeten er op schoolniveau en op leerlingniveau worden onderscheiden om een adequate beschrijving te geven van de schrijfvaardigheid in het voortgezet onderwijs?
- 2 Hoe groot zijn de verschillen tussen scholen in het niveau van schrijfvaardigheid?
- 3 Hoe groot zijn de verschillen tussen leerlingen binnen scholen?
- 4 Wat is de samenhang tussen de factoren op school- en leerlingniveau?
- 5 Hoe betrouwbaar is schrijfvaardigheid gemeten op het niveau van de school en de leerling?

METHODE VAN ONDERZOEK

Proefpersonen

De data zijn verzameld in het kader van de proefpeiling Nederlands (Kuhlemeier & Van den Bergh, 1989). In deze peiling is een landelijk representatieve beschrijving gegeven van de inhoud en het niveau van het vak Nederlands in het derde leerjaar van het voortgezet onderwijs. De afname vond plaats aan het einde van het schooljaar 1986-1987. Voor dit artikel maken we uitsluitend gebruik van de scores van 1451 derdeklassers op twee schrijfoopdrachten: een excuusbrief en een sollicitatiebrief. De leerlingen zijn afkomstig van 184 opleidingen voor VWO (19%), HAVO (19%), MAVO (24%), LTO (21%) of LHNO (18%); de verdeling naar sekse is 48% jongens en 52% meisjes; de gemiddelde leeftijd is 15 jaar en bijna 6 maanden.

Instrumenten

De excuusbrief en de sollicitatiebrief zijn functionele schrijftaken. De eerste opdracht, de excuusbrief, behelst het schrijven van een brief aan de directeur van de school. De leerling moet

zich verontschuldigen voor het feit dat hij of zij een afspraak wegens ziekte niet kan nakomen. De sollicitatiebrief schrijft de leerling aan de personeelschef van een groot warenhuis, op grond van een advertentie waarin om een vakantiehulp gevraagd wordt. Beide brieven zijn zogeheten bepaalde opdrachten, het type schrijfoopdracht dat in peilingsonderzoek het meest gebruikt wordt (Schoonen, 1991). In tegenstelling tot het traditionele opstel is de communicatieve context, zoals het schrijfdoel en het lezerspubliek, voorgeschreven. Net als bij het opstel heeft de leerling grote vrijheid in het bepalen van onder meer inhoud, stijl, organisatie en de meer technische aspecten van de tekst. De twee schrijfoopdrachten hebben veel gemeenschappelijk. Het zijn voorbeelden van korte, functionele brieven. We mogen deze tot op zekere hoogte beschouwen als twee equivalente metingen van functionele schrijfvaardigheid, of als steekproeven uit een bij benadering gelijk inhoudelijk domein.

De brieven zijn beoordeeld op tien verschillende aspecten i.c. tekstkenmerken. A priori kunnen deze tien aspecten in twee clusters worden verdeeld: een stelvaardigheidscomponent (compositie) en een schrijftechnische component (schrijftechniek). Voor de stelvaardigheidscomponent van zowel de excuus- als de sollicitatiebrief zijn scores op vijf aspecten voorhanden:

1. Globale kwaliteit. Hierbij gaven de beoordelaars een globaal oordeel over de kwaliteit van het schrijfprodukt na één keer lezen².
2. Inhoud algemeen. Bij dit aspect gaven de beoordelaars een algemeen oordeel over de inhoudelijke kwaliteit van de brief.
3. Stijl. Beoordeeld is of het taalgebruik passend en begrijpelijk is en de communicatie niet verstoort; daarbij is tevens gelet op de treffendheid of vaagheid van de woordkeus.
4. Organisatie. Het oordeel over de organisatie weerspiegelt de wijze waarop de leerling de inhoud gepresenteerd heeft. Warrige en onoverzichtelijke schrijfproducten scoren op dit aspect laag.
5. Inhoud specifiek. Hier telden de beoordelaars met behulp van een checklist het aantal relevante inhoudselementen.

Met betrekking tot de schrijftechnische component zijn gegevens over vijf soorten fouten beschikbaar:

6. Interpunctie. Het aantal interpunctiefouten per honderd zinnen.
7. Spelling. Het aantal spelfouten per honderd woorden.
8. Werkwoordspelling. Het aantal fouten in de werkwoordspelling per honderd zinnen.
9. Grammatica en idioom. Het aantal grammaticale en idiomatische fouten per honderd zinnen.
10. Welgevormde zinnen. Het aantal grammaticaal niet-correcte zinnen per honderd zinnen.

Iedere brief is op elk aspect beoordeeld door een aselechte steekproef van twee à drie uit negen leraren. Voorafgaand aan de beoordeling ontvingen de leraren een intensieve training. Het aantal beoordelaars per aspect bedraagt drie voor de oordelen over de globale kwaliteit, inhoud, stijl en organisatie; voor de overige zes aspecten zijn dat er twee. Ter ondersteuning van de beoordeling hadden de leraren de beschikking over van te voren geschaalde ankerteksten van lage, middelmatige en hoge kwaliteit. Uiteraard waren er verschillende ankerteksten voor verschillende aspecten. De jurybetrouwbaarheid varieert van .66 voor de telling van het aantal werkwoordspelfouten in de excuusbrief tot .99 voor het oordeel over de inhoud van de sollicitatiebrief, met een gemiddelde van .83 (vgl. Kuhlemeier & Van den Bergh, 1989). Voor aanvullende beschrijvende gegevens zij verwezen naar de Bijlage.

Structuur dataset

De in totaal bijna 70000 scores van de beoordelaars zijn per schrijfoopdracht per aspect gesommeerd tot één score per leerling. In totaal beschikken we over 28635 van deze leerlingsscores, afkomstig van 1451 derdeklassers van 184 scholen voor voortgezet onderwijs. Het aantal leerlingen per school varieert van vijf tot negen, met een gemiddelde van 7.91. Het aantal scores

per leerling bedraagt maximaal twintig: tien voor de excuusbrief en tien voor de sollicitatiebrief. Het werkelijke aantal scores per leerling varieert van vier tot twintig, met een gemiddelde van 19.73. Er is dus sprake van een vrijwel volledige dataset.

Statistische analyse

Overeenkomstig de structuur van de data onderscheiden we drie niveaus: school (niveau 3), leerling (niveau 2) en score (niveau 1). De variantie van een factor wordt opgesplitst in drie delen:

- een tussen-scholencomponent (schoolniveau);
- een tussen-leerlingen-binnen-scholencomponent (leerlingniveau);
- een tussen-scores-binnen-leerlingcomponent (scoreniveau);

In dit variantiecomponenten model worden de varianties en covarianties van de factoren gemoeddeeld op school- en leerlingniveau. De residuele variantie, die niet aan deze factoren kan worden toegeschreven, schatten we op het eerste niveau (vgl. Raudenbush, Rowan & Kang, 1991). Het eerste niveau is het meetgedeelte van het model; het beschrijft de relatie tussen de scores en de factoren uit het structurele gedeelte van het model (het binnen- en tussen-scholen-niveau).

De scores zijn per opdracht per aspect gestandaardiseerd, met een gemiddelde van nul en een variantie van één. Door deze standaardisatie worden de gemiddelden en de (totale) varianties per opdracht per aspect aan elkaar gelijk gemaakt. Voor de verdeling van de varianties en covarianties over de verschillende niveaus geldt dat uiteraard niet.

In de analyse onderscheiden we drie modellen die alleen van elkaar verschillen in de structuur op het tweede en derde niveau; de structuur op het eerste niveau is telkens gelijk. Op dit 'tussen-scores-binnen-leerlingniveau' schatten we voor elke combinatie van opdracht maal beoordelingsaspect een afzonderlijke residuele variantie. De mate waarin de scores de factor(en) indiceren, mag dus van opdracht tot opdracht en van aspect tot aspect verschillen. De residuele variantie representeert dat deel van de totale variantie dat niet aan deze factoren kan worden toegeschreven, en dat, gegeven het model, als 'ruis' te beschouwen is. De factoren op het binnen- en tussen-scholenniveau zijn hiermee gecorrigeerd voor onbetrouwbaarheid (vgl. Raudenbush, Rowan & Kang, 1991).

Basismodel (model I)

In eerste instantie is de structuur van de data beschreven met behulp van een tien-aspectenmodel (model I). In dit basismodel is voor elk van de tien beoordelingsaspecten een aparte latente factor (zowel op school- als op leerlingniveau) opgenomen. Elke aspectfactor heeft twee indicatoren: de scores op de beide opdrachten. Dit lijkt acceptabel, aangezien de twee schrijfopdrachten een grote gelijkenis vertonen. Stel Y_{hijk} verwijst naar de score op aspect h ($h = 1, 2, \dots, 10$), van opdracht i ($i = 1, 2$), van leerling j ($j = 1, 2, \dots, N_k$), in school k ($k = 1, 2, \dots, N$), en X_{hijk} is een dummy-variabele die 'aan' staat (= 1) als het een score voor aspect h van opdracht i betreft. In alle andere gevallen staat de dummy-variabele 'uit' (= 0). We kunnen het model nu schrijven als:

$$Y_{hijk} = \sum_h (\beta_{h0jk} * X_{hijk}) + e_{hijk} \quad (1)$$

In deze formule is β_{h0jk} het regressiegewicht voor aspect h van leerling j in school k . β_{h0jk} kan geïnterpreteerd worden als het gemiddelde over beide opdrachten van deze leerling voor aspect h . De tweede term, e_{hijk} , is de afwijking van aspect h van opdracht i van het gemiddelde voor dit aspect van leerling j in school k . We nemen aan dat e_{hijk} voor elk aspect normaal verdeeld is met gemiddelde nul en variantie s_{ehi}^2 . De matrix met residuele varianties op scoreniveau is dus een diagonale matrix met $(10*2=)$ twintig elementen.

Op het tweede niveau schrijven we voor alle aspecten het gemiddelde van leerling j als afwijking van het schoolgemiddelde:

$$\beta_{h0jk} = g_{h00k} + u_{h0jk} \quad (2a)$$

Hierbij is g_{h00k} het gemiddelde voor aspect h van school k . De random term, u_{h0jk} , geeft de afwijking van leerling j van school k van dit gemiddelde aan. We nemen aan dat deze residuele score ongecorrleerd is met (e_{hijk}) uit Vergelijking (1) en dat u_{h0jk} normaal verdeeld is met gemiddelde nul en variantie s_{uhj}^2 . De covariantiematrix van de residuele scores op leerlingniveau is een 10×10 symmetrische matrix met 55 elementen waarvan 10 varianties en 45 covarianties.

Op het derde niveau kan het gemiddelde van school k ten slotte geschreven worden als afwijking van het populatiegemiddelde:

$$g_{h00k} = \pi_{h000} + v_{h00k} \quad (2b)$$

De fixed parameter π_{h000} is een schatting van het populatiegemiddelde. Echter, daar de scores (per aspect per opdracht) gestandaardiseerd zijn, zijn deze (tien) parameters betekenisloos. We nemen aan dat de residuele scores (v_{h00k}) ongecorrleerd zijn met die op leerling- en scoreniveau. Voorts nemen we aan dat deze residuen normaal verdeeld zijn met gemiddelde nul en variantie s_{vhk}^2 . De tussen-scholen covariantiematrix van het basismodel is, net zoals de covariantiematrix op leerlingniveau, een 10×10 symmetrische matrix met 55 elementen.

Vereenvoudigd model (model II)

Inspectie van de covariantiematrices zoals geschat in het basismodel verschaft een eerste indruk van de dimensionaliteit van de data. Vervolgens is getracht het model zonder verlies aan informatie te vereenvoudigen. Daarbij hanteren we twee criteria. Om een factor als zelfstandige entiteit in het model te handhaven moet de variantie ervan significant groter zijn dan nul ($p < .05$); is dit niet het geval, dan wordt de desbetreffende factor op het desbetreffende niveau uit het model verwijderd. Ten tweede moet de correlatie met alle andere factoren significant kleiner zijn dan één (of groter dan min één); perfect correlerende factoren zijn samengevoegd tot één nieuwe factor (waarbij verschillen in varianties zijn meegemodelleerd)³. De toepassing van beide criteria resulteerde in een sterk vereenvoudigd model (model II).

Taakeffectenmodel (model III)

Op het niveau van de school en de leerling is in model I en II nog geen onderscheid gemaakt tussen de beide schrijfoopdrachten. In een verkennende analyse bleken er verschillen in varianties tussen de beide opdrachten te bestaan. De invloed van de opdracht schatten we zowel op school- als leerlingniveau. Daartoe breiden we Vergelijking (2a) uit met een dummy die het verschil tussen beide opdrachten representeert. Als we met $O2_{02jk}$ een dummy-variabele aanduiden die 'aan' staat als het de sollicitatiebrief betreft, dan krijgen we

$$\beta_{h0jk} = g_{h00k} + u_{h0jk} + u_{00jk} * O2_{02jk} \quad (3a)$$

en,

$$g_{h00k} = \pi_{h000} + v_{h00k} + v_{000k} * O2_{02jk} \quad (3b)$$

Vergelijking 1, 3a (of 2a) en 3b (of 2b) kunnen ook in één vergelijking weergegeven worden. Substitutie van vergelijking 3b in 3a, en het resultaat daarvan in 1 geeft:

$$Y_{hijk} = \sum_h (\pi_{h000} + u_{h0jk} + v_{h00k}) + O2_{02jk} (u_{00jk} + v_{000k}) + e_{hijk} \quad (4)$$

De taakdummy ($O2_{02jk}$) heeft de waarde nul voor de excuusbrief en één voor de sollicitatiebrief. In het taakeffectenmodel (model III) representeren de varianties van de schrijffactoren de

varianties van de excuusbrief. De covariantiematrix voor een model met tien schrijffactoren en één taakfactor op leerlingniveau kan als volgt worden weergegeven:

$$\begin{array}{cccccc}
 s_{u10j}^2 & & & & & \\
 s_{u10j, u20j} & s_{u20j}^2 & & & & \\
 \cdot & & & & & \\
 \cdot & & & & & \\
 s_{u10j, u(10)0j} & \dots & \dots & s_{u(10)0j}^2 & & \\
 s_{u10j, u00j} & \dots & \dots & s_{u(10)0j, u00j} & s_{u00j}^2 &
 \end{array} \quad (5)$$

In vergelijking (5) is s_{u10j}^2 de variantie van de eerste schrijffactor op leerlingniveau, $s_{u10j, u20j}$ de covariantie tussen de eerste en tweede schrijffactor, s_{u20j}^2 de variantie van de tweede schrijffactor, $s_{u(10)0j}^2$ de variantie voor het tiende aspect, s_{u00j}^2 de variantie van de taakfactor, en $s_{u10j, u00j}$ de covariantie van deze taakfactor met de eerste schrijffactor. Merk op dat de covariantiematrix op schoolniveau analoog is aan die op leerlingniveau, zij het dat op beide niveaus een verschillend aantal factoren onderscheiden kan worden.

De variantie van een factor voor de sollicitatiebrief is uit te rekenen als de som van de variantie van een schrijffactor, twee maal de covariantie van deze factor met de desbetreffende taakfactor, en de variantie van de taakfactor; bijvoorbeeld: de variantie voor de eerste schrijffactor van de sollicitatiebrief is te berekenen als $(s_{u10j}^2 + 2*s_{u10j, u00j} + s_{u00j}^2)$.

In model III representeren de covarianties tussen de schrijffactoren de covarianties voor de excuusbrief. De correlaties tussen de schrijffactoren worden berekend door de gemeenschappelijke covariantie te delen door de wortel uit het produkt van de desbetreffende varianties. Deze correlaties geven inzicht in de samenhang van de factoren op de onderscheiden niveaus.

De verhouding tussen de varianties op school- en leerlingniveau geeft inzicht in de verdeling van de systematische variantie over beide niveaus. Omdat de scores per opdracht per aspect gestandaardiseerd zijn, is de totale variantie van een factor in beginsel gelijk aan één³. De intraklassecorrelatie, gedefinieerd als de ratio van de tussen-scholenvariantie ten opzichte van de totale variantie (Kish, 1965), is dan gelijk aan de tussen-scholenvariantie. Omdat de totale variantie een deel 'ruis' bevat, is de hoogte van de intraklassecorrelatie een onderschatting van de 'ware' intraklassecorrelatie (Muthén, 1992). Deze 'error free' intraklassecorrelatie berekenen we door de tussen-scholenvariantie te delen door de totale systematische variantie (de som van de tussen-scholenvariantie en de tussen-leerlingenvariantie).

Betrouwbaarheid

Op schoolniveau wordt de betrouwbaarheid van het gemiddelde $(\rho_{f, school})$ voor factor f als volgt berekend (Raudenbush, Rowan & Kang, 1991):

$$\rho_{f, school} = S_{vkf}^2 / (S_{vkf}^2 + S_{ujf}^2 / N_j + S_{ehi}^2 / N_i) \quad (6)$$

waarbij

- S_{vkf}^2 : de tussen-scholenvariantie voor factor f ;
- N_j : het (gemiddeld) aantal leerlingen per school (circa 8);
- S_{ujf}^2 : de tussen-leerlingenvariantie voor deze factor;
- N_i : het (gemiddeld) aantal indicatoren;
- S_{ehi}^2 : de residuele variantie (berekend als het gemiddelde van de residuele varianties van de gegeven indicatoren h).

De betrouwbaarheid van de factorgemiddelden op leerlingniveau $(\rho_{f, leerling})$ wordt berekend als:

$$\rho_{f, \text{leerling}} = S^2_{\text{ujf}} / (S^2_{\text{ujf}} + S^2_{\text{ehi}} / N_i) \quad (7)$$

In het eindrapport van de proefpeiling Nederlands zijn de schrijffprestaties voor elke combinatie van opdracht en aspect afzonderlijk beschreven (zie Kuhlemeier & Van den Bergh, 1989). De betrouwbaarheid van deze scores wordt op analoge wijze berekend door de systematische factorvariantie ($S^2_{\text{vkf}} + S^2_{\text{ujf}}$) te delen door de som van de systematische factorvariantie ($S^2_{\text{vkf}} + S^2_{\text{ujf}}$) en de residuele variantie van de desbetreffende scores (S^2_{ehi}).

De multilevel factoranalyse is uitgevoerd met het programma ML3 (Prosser, Rasbash & Goldstein, 1991). Gebruik is gemaakt van de IGLS-schattingmethode (Iterative Generalized Least Squares). De significantietoetsingen van de verschillen tussen varianties en covarianties (correlaties) zijn uitgevoerd met behulp van een multiple vergelijkingsprocedure (zie Goldstein, 1987, p. 29). Als de nulhypothese waar is, levert deze procedure een toetsingsgrootte op die bij benadering χ^2 -verdeeld is. In de statistische toetsingen is het 5%-significantiecriterium gehanteerd.

De relatieve passing van het taakeffectenmodel (model III) ten opzichte van het analoge model zonder zo'n factor (model II) is geëvalueerd aan de hand van het verschil in de passingsgrootte $-2 \text{ Log Likelihood}$ tussen beide modellen. Dit verschil is asymptotisch χ^2 -verdeeld met het bijbehorende verschil in het aantal vrijheidsgraden (Bentler & Bonet, 1980). Op grond van deze toetsing verwerpen we model II ten faveure van model III ($p < .001$). Met dit laatste model beginnen we de rapportage van de bevindingen⁴.

RESULTATEN

Het aantal en de aard van de factoren

Hoeveel verschillende factoren moeten er op school- en leerlingniveau worden onderscheiden om een adequate beschrijving te geven van de schrijfvaardigheid in het voortgezet onderwijs (eerste onderzoeksvraag)? Tabel 1 geeft de patroonmatrix van model III dat de data met zo weinig mogelijk factoren zo goed mogelijk beschrijft. Horizontaal zijn voor één leerling de twintig toetsscores weergegeven en verticaal de factoren waarmee deze scores beschreven worden.

Voor een adequate beschrijving van de samenhang tussen de twintig toetsscores zijn in totaal acht factoren nodig (zie Tabel 1). Twee factoren, FAC1 en TAAK, treffen we zowel op schoolniveau als op leerlingniveau aan, zodat het aantal verschillende factoren zes bedraagt.

Eerst beschrijven we de structuur op schoolniveau. Wanneer we afzien van de taakfactor, blijkt de oorspronkelijke 10×10 -matrix (model I) gereduceerd te kunnen worden tot een 2×2 -matrix. Er zijn slechts twee schrijfvaardigheidsfactoren aantoonbaar. De eerste factor (FAC1) wordt geïndiceerd door de scores voor globale kwaliteit, inhoud, stijl, organisatie en inhoudselementen. Deze eerste factor interpreteren we als de algemeen-inhoudelijke kwaliteit van de schrijfprodukten (gedefinieerd in termen van kenmerken van de schrijfprodukten) of als de vaardigheid in het stellen (gedefinieerd in termen van vaardigheden). De tweede factor (FAC2) heeft de vijf foutentellingen als indicatoren; deze factor benoemen we als de technische kwaliteit van het schrijven (schrijftechniek).

Op leerlingniveau kunnen we de samenhang tussen de toetsscores met vier schrijfvaardigheidsfactoren en één taakfactor adequaat beschrijven. Het gepostuleerde tien-aspectenmodel, met voor elk aspect een aparte factor, vindt derhalve geen ondersteuning in de data. Net als op schoolniveau doet de eerste factor (FAC1) een beroep op algemeen-inhoudelijke kwaliteit van het schrijven ofwel de stelvaardigheid. Anders dan op schoolniveau blijken er aan de tien schrijftechnische scores niet één maar drie schrijffactoren ten grondslag te liggen: interpunctie (FAC3), spelling (FAC4) en een factor die geïndiceerd wordt door de scores voor werkwoordspelling en de beide grammaticale maten (FAC5). Deze laatste factor interpreteren we als de technische kwaliteit op zinsniveau, aangezien alle drie de indicatoren op dit niveau betrekking hebben.

Tabel 1. Patroonmatrix op school- en leerlingniveau (model III; EB: excuusbrief; SB: sollicitatiebrief).

	Tussen-scholen			Tussen-leerlingen				
	FAC1	FAC2	TAAK	FAC1	FAC3	FAC4	FAC5	TAAK
Stelvaardigheid								
1.1 Globale kwaliteit: EB	1			1				
1.2 Globale kwaliteit: SB	1		1	1				1
2.1 Inhoud: EB	1			1				
2.2 Inhoud: SB	1		1	1				1
3.1 Stijl: EB	1			1				
3.2 Stijl: SB	1		1	1				1
4.1 Organisatie: EB	1			1				
4.2 Organisatie: SB	1		1	1				1
5.1 Inhoudselementen: EB	1			1				
5.2 Inhoudselementen: SB	1		1	1				1
Schrijftechniek								
6.1 Interpunctie: EB		1			1			
6.2 Interpunctie: SB		1	1		1			1
7.1 Spelling: EB		1				1		
7.2 Spelling: SB		1	1			1		1
8.1 Werkwoordspelling: EB		1					1	
8.2 Werkwoordspelling: SB		1	1				1	1
9.1 Grammatica 1: EB		1					1	
9.2 Grammatica 1: SB		1	1				1	1
10.1 Grammatica 2: EB		1					1	
10.2 Grammatica 2: SB		1	1				1	1

De factorgemiddelden voor stelvaardigheid en schrijftechniek variëren systematisch tussen scholen en tussen leerlingen binnen scholen. Deze verschillen zijn evenwel niet onafhankelijk van de schrijfo opdracht in kwestie. Zowel op school- als op leerlingniveau levert de taakfactor (TAAK) een significante bijdrage aan de beschrijving van de samenhang tussen de twintig toetscores. Dit betekent dat het onderscheiden van meer dan één schrijfo opdracht ondersteuning vindt in de data.

Verschillen tussen scholen

Uit de hiervoor gerapporteerde analyse kwamen op schoolniveau drie factoren naar voren: stelvaardigheid, schrijftechniek en een factor die het verschil tussen de beide brieven indiceert. Een belangrijke vraag betreft de verschillen tussen scholen (tweede onderzoeksvraag). In hoeverre is het - gegeven de instrumentatie en het onderzoeksdesign - mogelijk om te differentiëren tussen hoog en laag presterende scholen? Om daarover iets te kunnen zeggen moeten we weten hoe groot de tussen-scholenvariantie van een factor is ten opzichte van de totale variantie (de intraklassecorrelatie) of ten opzichte van de totale systematische variantie (de ware of 'error free' intraklassecorrelatie). De proportie tussen-scholenvariantie voor de op schoolniveau gevonden factoren (FAC1 en FAC2) is weergegeven in het bovenste gedeelte van Tabel 2.

In het voortgezet onderwijs bestaan relatief grote verschillen tussen scholen in het niveau van stelvaardigheid (zie Tabel 2). Ongeveer één vijfde van de totale variantie (21% voor de excuusbrief en 20% voor de sollicitatiebrief) in stelvaardigheid is toe te schrijven aan de school die de leerling bezoekt. Bij de schrijftechnische factor zijn de verschillen tussen scholen aanzienlijk kleiner: 5% voor de excuusbrief en 10% voor de sollicitatiebrief. Blijkens Tabel 2 wordt alleen

Tabel 2. Factorvarianties (op de diagonaal), covarianties (onder de diagonaal) en correlaties (boven de diagonaal) van de schrijffactoren voor de excuus- en sollicitatiebrief.

SCHRIJFFACTOREN	EXCUUSBRIEF				SOLLICITATIEBRIEF			
	SCHOOLNIVEAU							
	FAC1		FAC2		FAC1		FAC2	
FAC1 Stelvaardigheid	.21		-.79		.20		-.43	
FAC2 Schrijftechniek	--.08		.05 ¹		-.06		.10 ¹	
	LEERLINGNIVEAU							
	FAC1	FAC3	FAC4	FAC5	FAC1	FAC3	FAC4	FAC5
FAC1 Stelvaardigheid	.35 ¹	-.16	-.08 ²	-.15	.38 ¹	-.06 ²	-.03 ²	-.00 ²
FAC3 Interpunctie	-.05	.23	.73	.43	-.02 ²	.25	.75	.50
FAC4 Spelling	-.02 ²	.19	.30 ¹	.36	-.01 ²	.21	.33 ¹	.45
FAC5 Werkw.spel./gramm.	-.04	.09	.09	.20 ¹	-.00 ²	.12	.13	.24 ¹

¹ de variantie van de excuusbrief verschilt significant van die van de sollicitatiebrief op 5%-niveau;

² de covariantie of correlatie wijkt niet significant af van nul op 5%-niveau.

de tussen-scholenvariantie van de schrijftechnische factor beïnvloed door de opdracht in kwestie; voor het gemiddelde stelvaardigheidsniveau van de school maakt het niet uit met welke opdracht er gemeten wordt (het verschil tussen .05 en .10 is wel significant, maar het verschil tussen .21 en .20 niet).

Merk op dat deze gegevens een onderschatting zijn van de ware verschillen tussen scholen. Er is immers nog geen rekening gehouden met meetfouten op scoreniveau. Gecorrigeerd voor onbetrouwbaarheid zijn de proporties tussen-scholenvariantie aanmerkelijk hoger: .38 (excuusbrief) en .34 (sollicitatiebrief) voor stelvaardigheid en .17 (excuusbrief) en .27 (sollicitatiebrief) voor schrijftechniek. Scholen verschillen derhalve sterker qua stelvaardigheid dan wat betreft de schrijftechnische kant van het schrijven. Deze grote tussen-scholenverschillen zijn niet alleen te herleiden tot verschillen in de effectiviteit van het gegeven onderwijs, maar verwijzen ook naar het selectieve karakter van het voortgezet onderwijs waarbij leerlingen op grond van hun prestaties aan verschillende opleidingstypen worden toegewezen (Kreft, 1987)⁵.

Verschillen tussen leerlingen binnen scholen

In model III bleken op leerlingniveau vijf factoren aantoonbaar: stelvaardigheid, interpunctie, spelling, een 'zinsfactor' geïndiceerd door werkwoordspelling en de beide grammaticale maten, en een taakfactor voor het verschil tussen beide schrijfp opdrachten. In hoeverre zijn er wat deze factoren betreft verschillen tussen leerlingen binnen scholen (derde onderzoeksvraag)? De desbetreffende factorvarianties zijn gerapporteerd in het onderste gedeelte van Tabel 2. De verschillen tussen leerlingen blijken het grootst voor stelvaardigheid (.35 voor de excuusbrief en .38 voor de sollicitatiebrief), gevolgd door spelling (.30 en .33), interpunctie (.23 en .25) en de 'zinsfactor' voor werkwoordspelling en grammatica (.20 en .24). Gecorrigeerd voor onbetrouwbaarheid zijn de overeenkomstige proporties tussen-leerlingenvariantie aanzienlijk hoger (respectievelijk .63, .86, .82 en .80 voor de excuusbrief en .66, .77, .71 en .71 voor de sollicitatiebrief).

Op leerlingniveau speelt de taak bij drie van de vier schrijffactoren een rol. Het betreft stelvaardigheid, spelling en de zinsfactor. In alle drie gevallen is de variantie van de sollicitatiebrief het grootst. Voor de interpunctie maakt het niet uit met welke opdracht deze gemeten wordt.

Samenhang tussen de factoren op school- en leerlingniveau

Tabel 2 geeft ook inzicht in de correlaties tussen de schrijffactoren (vierde onderzoeksvraag). De correlatie tussen de *school*gemiddelden voor de stelvaardigheid en de technische kwaliteit bedraagt $-.79$ voor de excuusbrief en $-.43$ voor de sollicitatiebrief. Hoe beter de leerlingen van een school stellen, hoe minder schrijftechnische fouten de leerlingen van die school gemiddeld maken.

Hoe de factorgemiddelden van de *leerlingen* met elkaar samenhangen, is eveneens te zien in Tabel 2. Het correlatiepatroon is goed interpreteerbaar. De correlaties tussen de stelvaardigheidsfactor en de drie schrijftechnische factoren zijn alle negatief, maar slechts in twee gevallen significant. Voor de excuusbrief geldt dat hoe beter de leerlingen stellen, hoe minder zij zondigen tegen de regels van de interpunctie ($r = -.16$) en hoe minder schrijftechnische fouten op zinsniveau zij maken ($r = -.15$). Bij de sollicitatiebrief is bij leerlingen geen enkele relatie tussen de kwaliteit van het stellen en de schrijftechnische kwaliteit aantoonbaar. De vaardigheid in het stellen staat dus relatief los van de beheersing van deelvaardigheden als interpunctie, spelling en werkwoordspelling/grammatica.

De correlaties tussen de drie schrijftechnische factoren zijn alle positief en significant, maar moeten als vrij laag worden beoordeeld. Er is echter één uitzondering: de samenhang tussen interpunctie en spelling is relatief hoog ($r = .73$ voor de excuusbrief en $r = .75$ voor de sollicitatiebrief). Leerlingen die weinig interpunctiefouten maken, zondigen weinig tegen de regels van de spelling.

Betrouwbaarheid op school-, leerling- en scoreniveau

De residuele variantie, die niet door de factoren op school- en leerlingniveau verklaard wordt, geeft een eerste indicatie van de (on)betrouwbaarheid waarmee de factoren gemeten zijn (vijfde onderzoeksvraag). De proporties residuele variantie per aspect per opdracht zijn weergegeven in Tabel 3.

Het scoreniveau blijkt een groot deel van de geobserveerde variantie voor zijn rekening te nemen. Anders gezegd: de factoren weten maar weinig variantie te binden⁶.

Niet alle twintig toetsscores blijken hun factor even goed te indiceren. Al naar gelang het aspect en de opdracht trekken de verschillen tussen de scores binnen een leerling (de tussenscores-binnen-leerlingvariantie) 23% tot 93% van de totale schrijfvaardigheidsvariantie naar

Tabel 3. Residuele varianties per aspect per opdracht (model III).

	Excuusbrief	Sollicitatiebrief
Stelvaardigheid		
1 Globale kwaliteit	.23	.25
2 Inhoud	.57	.64
3 Stijl	.39	.61 ¹
4 Organisatie	.24	.26
5 Inhoudselementen	.62	.41 ¹
Schrijftechniek		
6 Interpunctie	.67	.66
7 Spelling	.64	.50 ¹
8 Werkwoordspelling	.88	.93
9 Grammatica 1	.71	.67
10 Grammatica 2	.76	.74

¹ De variantie van de excuusbrief verschilt significant van die van de sollicitatiebrief ($p < .05$).

Tabel 4. Betrouwbaarheid van de meting op school-, leerling- en scoreniveau.

		Excuusbrief	Sollicitatiebrief
Schoolniveau			
FAC1	Stelvaardigheid	.63	.60
FAC2	Schrijftechniek	.21	.35
Leerlingniveau			
FAC1	Stelvaardigheid	.81	.82
FAC3	Interpunctie	.26	.27
FAC4	Spelling	.32	.40
FAC5	Werkwoordspelling/grammatica	.43	.48
Scoreniveau			
1	Globale kwaliteit	.71	.70
2	Inhoud	.50	.48
3	Stijl	.59	.49
4	Organisatie	.70	.69
5	Inhoudselementen	.47	.59
6	Interpunctie	.29	.35
7	Spelling	.35	.46
8	Werkwoordspelling	.22	.27
9	Grammatica 1	.26	.34
10	Grammatica 2	.25	.31

zich toe. Opvallend is de heterogeniteit bij stelvaardigheid. De oordelen over de globale kwaliteit en organisatie van de brieven doen het verreweg het best. De twee inhoudelijke aspecten en stijl zijn de slechtste indicatoren voor de kwaliteit van het stellen. Van de schrijftechnische aspecten zijn de residuele varianties zonder uitzondering hoog (minimaal 50%). De meting van werkwoordspelling bestaat vrijwel volledig uit 'ruis' (88% versus 93% voor respectievelijk de excuus- en de sollicitatiebrief). De schrijftechniek is dus slechter gemeten dan de stelvaardigheid (voor een vergelijkbare bevinding, zie Breland e.a., 1987)⁷.

Met behulp van een paarsgewijze vergelijkingsprocedure (Goldstein, 1987, p. 29) is getoetst of de residuele varianties van de tien aspecten van opdracht tot opdracht verschillen. De hypothese dat deze gelijk zijn voor de beide opdrachten moet verworpen worden. Van de tien contrasten zijn er drie significant. Bij de excuusbrief is het aspect stijl een beduidend betere indicator voor stelvaardigheid dan bij de sollicitatiebrief. Bij de inhoudselementen en de spelfouten is het net andersom. Daar is de residuele variantie bij de excuusbrief het grootst.

Kortom: de mate waarin de scores de schrijffactoren indiceren, is niet invariant over aspecten en opdrachten. Het schatten van aparte residuele varianties per aspect per opdracht blijkt dus zinvol.

Informatie over de betrouwbaarheid van de meting op school-, leerling- en scoreniveau (zie Vergelijking 6 en 7) is weergegeven in Tabel 4.

Wanneer we het criterium voor een betrouwbare meting vrij arbitrair op .70 stellen, dan blijken de stelvaardigheid en schrijftechniek op schoolniveau onvoldoende betrouwbaar gemeten. De betrouwbaarheden van de schoolgemiddelden voor schrijftechniek zijn zonder meer laag te noemen.

Van de factoren op leerlingniveau is alleen de stelvaardigheidsfactor voldoende betrouwbaar gemeten (.81 voor de excuusbrief en .82 voor de sollicitatiebrief); de betrouwbaarheid van de factorgemiddelden voor interpunctie, spelling en de 'zinsfactor' bij leerlingen is ronduit laag.

De betrouwbaarheid van de toetsscores, de rapportage-eenheden van de taalpeiling, laat vaak te wensen over. Het varieert van .22 tot .71, met een gemiddelde van .45.

DISCUSSIE

In de recente taalpeilingen wordt het niveau van het schrijfonderwijs aan de hand van meerdere schrijfopdrachten op een groot aantal tekstenmerken (aspecten) beschreven. In de onderhavige studie is de efficiëntie van deze gedetailleerde rapportage onderzocht. De samenhang tussen de twintig toetsscores i.c. rapportage-eenheden blijkt met behulp van vijf schrijffactoren en één taakfactor adequaat beschreven te kunnen worden.

Het aantal en de aard van de factoren is niet op elk niveau gelijk. Op *schoolniveau* liggen aan de twintig toetsscores i.c. rapportage-eenheden drie factoren ten grondslag: twee schrijffactoren en een taakfactor. De eerste factor betreft de algemeen-inhoudelijke kwaliteit van het schrijven ofwel stelvaardigheid; de tweede heeft betrekking op de technische kwaliteit van het schrijven. Ongeveer één vijfde van de totale variantie (21% voor de excuusbrief en 20% voor de sollicitatiebrief) in stelvaardigheid is toe te schrijven aan de school die de leerling bezoekt. Bij de schrijftechnische factor zijn de verschillen tussen scholen aanzienlijk kleiner: 5% voor de excuusbrief en 10% voor de sollicitatiebrief. Deze factorvarianties op schoolniveau zijn toe te schrijven aan verschillen tussen scholen in onder meer instroom, het leerstofaanbod en de effectiviteit van de instructie en het schoolbeleid (Longford & Muthén, 1992).

De dimensionaliteit op *leerlingniveau* is niet gelijk aan die op schoolniveau. Op leerlingniveau zijn vijf factoren nodig om de samenhang tussen twintig toetsscores adequaat te beschrijven: vier schrijfvaardigheidsfactoren en een taakfactor. Gemeenschappelijk met het schoolniveau is de factor voor stelvaardigheid; de leerling voegt hier geen nieuwe dimensies toe. Daarentegen valt de schrijftechnische factor bij leerlingen uiteen in factoren voor interpunctie, spelling en de technische kwaliteit op zinsniveau. Overeenkomstig de modellering beschouwen we de structuur op leerlingniveau als een afwijking van de structuur op schoolniveau. De structuur op leerlingniveau geeft aan in hoeverre individuele leerlingen afwijken van deze gemeenschappelijke structuur. Het percentage tussen-leerlingenvariantie varieert met de factor en de opdracht van .20 tot .38. Deze factorvarianties verwijzen naar de unieke, individuele bijdrage van de leerling aan de schoolprestaties, bijvoorbeeld ten gevolge van zijn of haar intelligentie, inzet en motivatie. Met uitzondering van het verband tussen interpunctie en spelling is de samenhang tussen de factoren bij leerlingen laag; de kwaliteit van het stellen houdt nauwelijks verband met de technische kwaliteit.

De rapportage

De meting van schrijfvaardigheid blijkt duidelijk taakafhankelijk, zowel op school-, leerling- als scoreniveau. Niet alleen de varianties van de schrijffactoren verschillen vaak van opdracht tot opdracht, maar ook de samenhang tussen de schrijffactoren en de proporties residuele scorevariantie. Het onderscheiden van meer dan één schrijfopdracht in de rapportage lijkt derhalve gerechtvaardigd.

Het onderscheiden van tien beoordelingsaspecten i.c. rapportage-eenheden vindt daarentegen geen steun in de data. Wil men, zoals in peilingsonderzoek, een beschrijving geven van het gemiddelde niveau van schrijfvaardigheid in de populatie, dan zijn slechts drie factoren van belang (stelvaardigheid, schrijftechniek en een taakfactor die het verschil tussen de opdrachten beschrijft). Is men daarentegen ook geïnteresseerd in de verschillen tussen leerlingen, dan moeten vijf factoren worden onderscheiden (stelvaardigheid, interpunctie, spelling, de technische kwaliteit op zinsniveau en een taakfactor). Gezien vanuit een psychometrisch perspectief was de rapportage per opdracht per aspect in de peiling Nederlands (Kuhlemeier & Van den Bergh, 1989) dus te gedetailleerd.

Meting van schrijfvaardigheid

In de taalpeilingen wordt schrijfvaardigheid doorgaans gemeten met behulp van zogeheten bepaalde opdrachten (Schoonen, 1991). In tegenstelling tot de vrije opdracht zijn daarbij bepaalde aspecten van de communicatieve context, zoals het schrijfdoel en het lezerspubliek, voorgescreven. De inhoud, stijl, organisatie en de meer technische aspecten van de tekst moet de leerling grotendeels zelf bepalen. Deze toetsvorm lijkt valide (Burger & Burger, 1994). Bij de efficiëntie ervan voor de meting van schrijfvaardigheid kunnen echter twee vraagtekens worden geplaatst.

Ten eerste blijken de vijf indicatoren voor stelvaardigheid niet onderscheidbaar. De oordelen over de algemene kwaliteit, de inhoud, de stijl, de organisatie en de telling van het aantal relevante inhoudselementen vormen één factor. Voor de meting van de inhoudelijke, stilistische en organisatorische aspecten van stelvaardigheid zijn gerichte, sterker sturende schrijfoopdrachten vermoedelijk efficiënter. Deze zijn overigens niet per definitie minder valide dan vrije of bepaalde opdrachten (Breland e.a., 1987; Van Schooten & De Glopper, 1990; Schoonen, 1991). Desalniettemin pleiten wij niet voor afschaffing van de vrije en de bepaalde opdrachten in peilingsonderzoek, zoals Conry & Jerolsky (1980). Daarvoor worden deze taaltaken in het onderwijs te veel gebruikt. Wel verdient het aanbeveling de arbeidsintensieve en kostbare beoordeling op meerdere stelvaardigheidsaspecten in heroverweging te nemen. Uit overwegingen van efficiëntie kan worden volstaan met de meest betrouwbare indicator van stelvaardigheid, die tevens de minste training en beoordelingstijd vergt. Dat is het oordeel over de globale kwaliteit, geoperationaliseerd als de algemene indruk die de tekst op de lezer maakt na één keer lezen.

Ten tweede blijkt de bepaalde opdracht geen efficiënt instrument voor de meting van de technische kant van het schrijven. De factoren interpunctie, spelling en grammatica blijken bij leerlingen onderscheidbaar, maar de betrouwbaarheid is zo laag dat het gebruik van de bepaalde opdracht voor dit doel af te raden is. Zogeheten deelvaardigheidstoetsen, zoals de meerkeuze-opdracht, de correctie-opdracht of het dictee, lijken hiertoe meer geschikt (Culpepper & Ramsdell, 1982; Van Schooten & De Glopper, 1990).

Peilingsonderzoek en schooleffectiviteit

Voor onderzoek naar schooleffectiviteit is het kunnen aantonen van verschillen tussen scholen een minimale voorwaarde. De onderzochte scholen voor voortgezet onderwijs verschillen aanzienlijk in het niveau van stelvaardigheid, terwijl de verschillen in de technische kwaliteit van het schrijven vrij klein zijn. Dit laatste zal zeker ook te maken hebben met de lage betrouwbaarheid van de schrijftechnische scores.

Wanneer we het criterium voor een adequate meting op schoolniveau (tamelijk arbitrair) op .70 stellen, zijn de schoolgemiddelden voor stelvaardigheid onvoldoende betrouwbaar gemeten om secundaire analyse of vervolgonderzoek te rechtvaardigen. De technische kwaliteit van het schrijven is op schoolniveau ronduit onbetrouwbaar gemeten. Hieraan kunnen velerlei oorzaken ten grondslag liggen. Een daarvan is het geringe aantal leerlingen per school. Voor het verkrijgen van een nauwkeurige schatting van populatie-indices is een afnamedesign met één leerling per item per school het meest efficiënt. De in peilingsonderzoek gebruikte steekproef- en afnamedesigns dragen ertoe bij dat het negatieve effect van de clustering op de schattingsnauwkeurigheid van de populatie-indices geminimaliseerd wordt (Beaton, 1988; Wijnstra, 1988; Rust & Johnson, 1992). Enerzijds worden naar verhouding veel scholen en weinig leerlingen binnen scholen gesampeld. Anderzijds wordt gebruikt gemaakt van matrix sampling designs waarbij niet alle leerlingen binnen een school dezelfde taken krijgen voorgelegd. In de praktijk van peilingsonderzoek worden per school hooguit enkele tientallen leerlingen getrokken, terwijl elk item binnen een school door slechts enkele leerlingen gemaakt wordt. Gezien vanuit de primaire doelstelling van peilingsonderzoek - de nauwkeurige schatting van (sub)populatie-indices - is dit een begrijpelijke en verstandige keuze. Er moet echter wel een prijs voor worden betaald. Het kleine aantal leerlingen per school maakt het doen van uitspraken over de effectiviteit van de instructie of de school tot een moeilijke (Kuhlemeier, 1994) zo niet onmogelijk opgave (Rijlaarsdam, Van den Bergh & Zwarts, 1992). Het onderscheidingsvermogen voor de

significantietoetsing van de variantie tussen klassen en/of scholen is laag (Bosker & Snijders, 1990). Wil men onderzoek naar schooleffectiviteit aan peilingsonderzoek koppelen (Scheerens & Stoel, 1987; Bock, 1988), dan verdient het aanbeveling om het steekproef- en afnamedesign hierop af te stemmen (vgl. Kuhlemeier, Van den Bergh & Smeets, 1989).

Data-analyse

Een laatste aanbeveling betreft de analyse van peilingsdata. Dit gebeurt over het algemeen met technieken van unilevel analyse. Als men primair geïnteresseerd is in nauwkeurige schattingen van (sub)populatie-indices en er tevens in slaagt om het clustereffect te minimaliseren is hier niets mis mee. In de praktijk van peilingsonderzoek doen zich echter vaak aanzienlijke cluster-effecten voor. Als er al gecorrigeerd wordt voor deze clustereffecten, gebeurt dit met behulp van ingewikkelde wegingsprocedures (Johnson & Rust, 1992). Wil men peilingsonderzoek koppelen aan onderzoek naar schooleffectiviteit en het steekproef- en afname-ontwerp hierop afstemmen, dan verdient het aanbeveling de data ook op een andere manier te analyseren. Op grond van de bevindingen van de onderhavige studie lijkt het verstandig om de hiërarchische structuur van de data dan expliciet in een multilevel analyse te modelleren. Niet alleen vanwege de correcte schatting van standaardfouten (Goldstein, 1987), maar ook vanwege de extra interpretatiemogelijkheden die het onderscheiden van meer niveaus biedt.

NOTEN

1. Vanuit het perspectief van een zo zuinig en efficiënt mogelijke meting van *de* taalvaardigheid in peilingsonderzoek is er veel voor te zeggen om taakeffecten als ruis te beschouwen. Deze 'negatieve' benadering van taakeffecten is echter niet geheel onproblematisch. Vanuit het perspectief van een inhoudsva-lide meting zou men de variantie tussen taken ook als legitieme variantie kunnen beschouwen. Taal-taken kunnen immers verwijzen naar onderwijsdoelen die ieder voor zich het nastreven waard zijn en die in de onderwijspraktijk ook daadwerkelijk worden nagestreefd. Denk bijvoorbeeld aan het schrijven van verschillende soorten teksten, zoals een opstel, een betoog of een brief. In dat geval zou men taak-effecten niet zonder meer als ruis of onbedoelde variantie mogen interpreteren. Dit onderscheid tussen bedoelde en niet-bedoelde variantie hangt nauw samen met dat tussen fixed en random effecten. In de proefpeiling Nederlands zijn de prestaties voor elke taalkaak afzonderlijk gerapporteerd. Hiermee wor-den taaltaken impliciet als fixed effecten beschouwd. Zoals Clark (1973) laat zien, zijn er ook argumen-ten om taakeffecten als random effecten te beschouwen. In dat geval vormen de taken een aselechte steekproef uit het domein van mogelijke taken waarnaar gegeneraliseerd wordt. Voor de rapportage van peilingsdata zou dit bijvoorbeeld betekenen dat er geen gemiddelden per opdracht per vaardigheid wor-den gerapporteerd, maar vaardigheidsgemiddelden met een bandbreedte voor de spreiding ten gevolge van de verschillende taken.
2. Bij hun oordeel over de globale kwaliteit is de beoordelaars niet expliciet gevraagd om af te zien van de schrijftechnische aspecten. Er is dus sprake van enige inhoudelijke contaminatie met andere aspecten.
3. De in de tabellen gerapporteerde varianties op school-, leerling- en scoreniveau tellen niet altijd op tot 1.00. Dit is een gevolg van afrondingsfouten en van de wijze waarop de factorvarianties gemodelleerd zijn. Dit laatste vereist enige toelichting. De variantie van een factor op een bepaald niveau is een gewo-gen gemiddelde van de varianties van de indicatoren op datzelfde niveau. Bijvoorbeeld: de tussen-scho-lenvarianties van globale kwaliteit, inhoud, stijl, organisatie en inhoudselementen zijn bepalend voor de variantie van de gemeenschappelijke stelvaardigheidsfactor. De varianties van de indicatoren zijn echter niet alle gelijk. Zo is de tussen-scholenvariantie voor globale kwaliteit, inhoud en stijl significant groter dan voor inhoud en inhoudselementen. In de analyse is aan dergelijke verschillen in variantie recht gedaan door naast de factor nog een aparte variantie te modelleren. In ons voorbeeld is behalve de stelvaardigheidsfactor een aparte variantie voor inhoud/inhoudselementen geschat. Dit leidt tot een (ge-ringe) verandering van de factorvariantie, met als gevolg dat de totale variantie afwijkt van 1.00.
4. Een belangrijke assumptie is die van (multivariate) normaliteit van de random termen op elk niveau. De verdeling van de schrijftechnische fouten is nogal scheef (zie de Bijlage). Dit heeft te maken met het gegeven dat een kleine groep leerlingen verantwoordelijk is voor een zeer groot deel van de fouten in interpunctie, spelling en grammatica; de overgrote meerderheid van de leerlingen maakt geen of slechts enkele schrijftechnische fouten. Ter controle zijn de testcores afzonderlijk genormaliseerd via de nor-maliserende transformatie die beschreven wordt in Prosser, Rasbash and Goldstein (1991, p. 44) en

vervolgens geheranalyseerd. Het effect op de varianties en covarianties bleek minimaal te zijn; het patroon van uitkomsten bleef gelijk. Om deze reden wordt in dit artikel gebruik gemaakt van de originele, niet-genormaliseerde data.

5. Dit roept de vraag op naar de generaliseerbaarheid van de in de totale responsgroep gevonden structuur over opleidingstypen. Daartoe zijn de analyses nogmaals uitgevoerd, maar nu apart voor het VWO/HAVO, MAVO en LBO. Zoals verwacht zijn de tussen-scholenverschillen binnen deze opleidingstypen aanzienlijk kleiner dan in de totale responsgroep. Wel blijken in het VWO/HAVO, MAVO en LBO dezelfde factoren onderscheidbaar als in de totale responsgroep.
6. Dit kon op grond van de eerder gerapporteerde systematische factorvarianties op school- en leerlingniveau reeds vermoed worden. Omdat de scores per opdracht per aspect gestandaardiseerd zijn, met een gemiddelde van nul en een variantie van één, is de systematische variantie van een factor in beginsel gelijk aan de totale variantie minus het (gewogen) gemiddelde van de residuele varianties van de scores die de desbetreffende factor indiceren.
7. Een alternatieve verklaring voor de onbetrouwbaarheid van de foutentellingen is het gegeven dat de lengte van de excuus- en sollicitatiebrieven te kort is om de technische conventies betrouwbaar te kunnen meten. Een overeenkomstige analyse gebaseerd op langere teksten - een betoog over huiswerk en een opstel over een angstige ervaring - leidt echter tot vergelijkbare uitkomsten.

LITERATUUR

- Balke, G. (1991). *Multilevel factor analysis of proficiency in English as a foreign language*. Paper presented at the symposium 'Multilevel factor analysis: applications to education' at the annual meeting of the American Educational Research Association, Chicago, april 3-7, 1991.
- Beaton, A. (1988). *Expanding the new design: the NAEP 1985-1986 technical report*. Princeton, NJ: Educational Testing Service.
- Bentler, P.M. & Bonet, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bergh, H. van den (1988). *Examens geëxamineerd*. 's-Gravenhage: SVO.
- Bock, R.D. (1988). A design for a biennial national education assessment, reporting by state. *International Journal of Educational Research*, 12, 745-750.
- Bosker, R.J. & Snijders, T.A.B. (1990). Statistische aspecten van multiniveau onderzoek. *Tijdschrift voor Onderwijsresearch*, 15, 317-329.
- Breland, H.M., Camp, R., Jones, R.J., Morris, M.M. & Rock, D.A. (1987). *Assessing writing skill* (Research Monograph No. 11). New York: College Entrance Examination Board.
- Burger, S.E. & Burger, D.L. (1994). Determining the validity of performance-based assessment. *Educational Measurement: Issues and Practice*, 13, 1, 9-15.
- Clark, H.H. (1973). The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12, 335-359.
- Conry, E. & Jerolski, S. (1980). *The British Columbia assessment of written expression. General report*. Province of British Columbia: Ministry of Education.
- Culpepper, M.M. & Ramsdell, R. (1982). A comparison of a multiple choice and essay test of writing skills. *Research in the Teaching of English*, 16, 295-297.
- Gelderen, A. van (1987). *Taalmaten. Deel 3: Relaties tussen spreek-, schrijf-, lees- en luisterprestaties*. Amsterdam: SCO.
- Gipps, G. & Goldstein, H. (1983). *Monitoring children. An evaluation of the Assessment of Performance Unit*. London: Educational Books.
- Glopper, K. de (1988). *Schrijven beschreven. Inhoud, opbrengsten en achtergronden van het schrijfonderwijs in de eerste vier leerjaren van het voortgezet onderwijs*. 's-Gravenhage: SVO.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. Charles Griffin: London.
- Goldstein, H. & McDonald, R.P. (1988). A general model for the analysis of multilevel data. *Psychometrika*, 53, 4, 455-467.
- Hox, J.J. (1993). Factor analysis of multilevel data: gauging the Muthén model. In J.H.L. Oud & R.A.W. van Blokland-Vogelzang (Eds.). *Advances in longitudinal and multivariate analysis in the behavioral sciences* (pp. 141-155). Nijmegen: ITS.
- Johnson, E.G. & Rust, K.F. (1992). Population inferences and variance estimation for NAEP data. *Journal of Educational Statistics*, 17, 175-190.
- Kish, L. (1965). *Survey sampling*. New York: Wiley.
- Kreft, I. (1987). *Models and methods for the measurement of school effects*. Amsterdam: Universiteit van Amsterdam.

- Kuhlemeier, J.B. (1994). Effecten van schoolkenmerken en steldidactiek op functionele stelvaardigheid Nederlands in het MAVO. *Tijdschrift voor Onderwijsresearch*, 19, 2, 161-181.
- Kuhlemeier, J.B. & Bergh, H. van den (1989). *De Proefpeiling Nederlands: een onderzoek naar de haalbaarheid van peilingsonderzoek in het voortgezet onderwijs*. Specialistisch Bulletin Nr 74. Arnhem: Cito.
- Kuhlemeier, J.B. & Bergh, H. van den (1991). De correlatieve structuur van taalvaardigheid: een exploratie. *Tijdschrift voor Onderwijsresearch*, 16, 143-159.
- Kuhlemeier, J.B., Bergh, H. van den, Notté, H., Wagenaar, H., Verstralen, H. & Cappers, R. (1994). *Balans van het aardrijkskunde-onderwijs in het derde leerjaar van het voortgezet onderwijs*. Arnhem: Instituut voor Toetsontwikkeling.
- Kuhlemeier, J.B., Bergh, H.H. van den & Smeets, L. (1989). *Projectopzet peiling Aardrijkskunde*. Arnhem: Instituut voor Toetsontwikkeling.
- Linn, R.L. & Burton, E. (1994). Performance-based assessment: implications of task specificity. *Educational Measurement: Issues and Practice*, 13, 1, 5-8.
- Longford, N.T. (1990). Multivariate component analysis: an application in test development. *Journal of Educational Statistics*, 15, 2, 91-112.
- Longford, N.T. & Muthén, B.O. (1992). Factor analysis for clustered observations. *Psychometrika*, 57, 4, 581-597.
- Melse, L. & Kuhlemeier, J.B. (1993). De structuur van schrijfvaardigheid Duits: een toepassing van multilevel factoranalyse. In C. Blankenstijn & A. Scheper. *Taalvaardigheid: Symposiumbundel Werkverband Amsterdamse Psycholinguïsten. Wap Publikatie 2* (p. 139-152). Dordrecht: ICC Publications.
- Muthén, B.O. (1992). Multilevel factor analysis of class and student achievement components. *Journal of Educational Measurement*, 28, 338-354.
- Neville, M. (1988) *Assessing and teaching language: Literacy and oracy in schools*. London: Macmillan.
- Prosser, R., Rasbash, J. & Goldstein, H. (1991). *ML3: Software for three-level analysis. User's guide*. London: Institute of Education.
- Raudenbush, S.W., Rowan, B. & Kang, S.J. (1991). A multilevel, multivariate model for studying school climate with estimation via the EM algorithm and application to U.S. high-school data. *Journal of Educational Statistics*, 16, 4, 295-330.
- Rust, K.F. & Johnson, E.G. (1992). Sampling and weighting in the national assessment. *Journal of Educational Statistics*, 17, 111-129.
- Rymenans, R., Leroy, G. & Daems, F. (1991). *Geletterdheid op achttien jaar: peiling naar de lees- en schrijfvaardigheid bij het einde van het secundair onderwijs. Deel II: Functionele taalvaardigheid*. Antwerpen: Universitaire Instelling Antwerpen, Departement Didactiek en Kritiek.
- Rijlaarsdam, G., Bergh, H. van den & Zwarts, M. (1992). Incidentele transfer bij produktieve taalopdrachten: een aanzet tot een baseline. *Tijdschrift voor Onderwijsresearch*, 17, 55-66.
- Scheerens, J. & Stoel, W.G.R. (1987). Conceptuele en methodologische problemen bij onderzoek naar de effectiviteit van schoolorganisaties. In J.Scheerens & W.G.R. Stoel (Red.), *Effectiviteit van onderwijsorganisaties* (pp. 1-16). Lisse: Swets & Zeitlinger.
- Schoonen, R. (1991). *De evaluatie van schrijfvaardigheidsmetingen*. Amsterdam: Universiteit van Amsterdam.
- Schooten, E. van & Glopper, K. de (1990). De validiteit van meerkeuze- instrumenten voor het meten van schrijfvaardigheid. *Tijdschrift voor Taalbeheersing*, 12, 93-110.
- Sytstra, J. (1993). *Balans van het taalonderwijs halverwege de basisschool. Uitkomsten van de eerste taalpeiling medio basisonderwijs*. Arnhem: Cito.
- Sytstra, J. & Zwarts, M.A. (1991). *Constructvaliditeit van het model van domeinbeschrijvingen. Verslag van een secundaire analyse*. Arnhem: Cito.
- Tate, R.L. & King, F.J. (1994). Factors which influence precision of school-level IRT ability estimates. *Journal of Educational Measurement*, 31, 1, 1-15.
- Wijnstra, J.M. (1988). *Balans van het rekenonderwijs in de basisschool*. Arnhem: Cito.
- Zwarts, M. (1990). *Balans van het taalonderwijs aan het einde van de basisschool. Uitkomsten van de eerste taalpeiling einde basisonderwijs*. Arnhem: Cito.
- Zwarts, M. & Rijlaarsdam, G. (1991). *Verantwoording van de taalpeiling einde basisonderwijs 1988*. Arnhem: Cito.

BIJLAGE 1

Gemiddelden, standaarddeviaties, scheefheid, laagste en hoogste score, en aantal leerlingen per aspect per opdracht (EB: excuusbrief; SB: sollicitatiebrief)

Variabele	Gemiddelde	Stddev	Scheefheid	Minimum	Maximum	N
Stelvaardigheid						
Globale kwaliteit: SB	7.46	2.79	.36	3	15	1438
Globale kwaliteit: EB	8.90	2.89	-.13	3	15	1444
Inhoud: SB	9.55	4.44	-.41	3	15	1443
Inhoud: EB	13.10	2.38	-1.64	3	15	1437
Stijl: SB	5.94	2.49	.82	3	15	1443
Stijl: EB	7.88	2.60	.17	3	15	1430
Organisatie: SB	8.34	2.98	.17	3	15	1447
Organisatie: EB	9.09	2.91	-.09	3	15	1437
Inhoudselementen: SB	25.63	10.28	.19	1	50	1437
Inhoudselementen: EB	17.17	3.89	-.22	1	32	1425
Schrijftechniek						
Interpunctie ¹ : SB	100.10	77.40	2.00	0	800	1436
Interpunctie ¹ : EB	82.24	64.30	1.55	0	450	1434
Spelling ² : SB	4.76	4.12	1.66	0	31	1434
Spelling ² : EB	2.81	3.10	1.99	0	21	1432
Werkwoordspelling ³ : SB	2.49	7.70	4.87	0	100	1427
Werkwoordspelling ³ : EB	3.55	9.71	3.71	0	100	1433
Grammatica 1 ⁴ : SB	44.71	38.64	1.80	0	400	1425
Grammatica 1 ⁴ : EB	42.79	40.18	1.65	0	350	1436
Grammatica 2 ⁵ : SB	61.03	28.66	-.21	0	100	1429
Grammatica 2 ⁵ : EB	63.83	30.86	-.21	0	100	1435

¹ aantal interpunctiefouten per honderd zinnen;

² aantal spelfouten per honderd woorden;

³ aantal werkwoordspelfouten per honderd zinnen;

⁴ aantal grammaticale en idiomatische fouten per honderd zinnen;

⁵ aantal grammaticaal correcte zinnen per honderd zinnen (in de multilevel analyse is deze testscore vermenigvuldigd met minus één).

Een onderzoek naar leerstijlen

H.C. Schouwenburg

Afdeling Studie Ondersteuning, Rijksuniversiteit Groningen

ABSTRACT

Vermunt's (1992) four learning styles, partly based on English and partly on American factor-analytic research, have been defined as characteristic combinations of cognitive information processing activities, metacognitive control strategies, attitudes towards learning, and motives to learn. Learning styles may be conceived as relatively stable traits. It may be argued that optimal instruction would result from a certain fit of instructional methods to learning styles of students. On the other hand, learning styles may develop or not as a result of praise and blame administered by teachers and the educational system. As a consequence, learning styles may provide information about the quality of the educational setting.

In order to test Vermunt's basic ideas on the existence and importance of learning styles, correlational studies were performed among three independent samples of students at the University of Groningen. Research questions were: (1) Can Vermunt's four learning styles reliably be assessed among different groups of university students? (2) What predictive value do learning styles have for academic success? (3) What connection can be established between learning styles and other study approach variables?

Results indicate that (1) Vermunt's learning styles can indeed be demonstrated in different samples of university students. (2) Learning styles appear to be independent of self-regulative activities of students. Vermunt's most important learning style, meaning-oriented style, appears to be related mainly to intelligence-like learning strategies, such as analyzing and synthesizing. (3) In contrast to self-regulative activities, learning styles do not seem to predict academic success.

The results show a remarkable consistency among the three samples. Their implications for educational research and practice are critically evaluated.

INLEIDING

Toen vanaf 1992 de lovende besprekingen van Vermunt's proefschrift (Vermunt, 1992) in de pers verschenen, dacht menigeen betrokken bij het onderwijs in het door Vermunt ge(her)introduceerde begrip *leerstijl* het diagnosticum bij uitstek voor (in)effectieve studeergewoonten van studenten te hebben gevonden. Immers, het door hem geconcipeerde geheel van activiteiten, opvattingen en motieven was voor velen herkenbaar en sprak tot de verbeelding. Andere initiatieven om diagnostische instrumenten met betrekking tot ineffectief studeren te ontwikkelen waren in vergelijking met Vermunt's instrument nogal eenzijdig op gedrag gericht (Vorst, 1993) of richtten zich op slechts enkele aspecten van het studeren (Klein, Topman en van der Ploeg, 1994; Schouwenburg, 1994). Vermunt's totaalconceptie had daarom grotere *face* validiteit, terwijl de condensering van deze vele ingrediënten tot vier herkenbare leerstijlen een eenvoudige praktische toepassing leek te beloven.

Nu zijn er op het begrip leerstijl twee verschillende gezichtspunten mogelijk. Volgens het ene gezichtspunt, dat van Vermunt, is een leerstijl niet meer dan een bepaalde combinatie van activiteiten, opvattingen en motieven, ontstaan uit de interactie tussen persoons- en omgevingsgebonden invloeden (Vermunt, 1992). Daarmee is een leerstijl in principe (door onderwijs) veranderbaar. Volgens de andere, meer klassieke opvatting kan men een leerstijl opvatten als de concretisering van een voorkeur voor een bepaalde manier van denken (en handelen) in de

onderwijssituatie. In zijn meest algemene vorm wordt zo'n voorkeur *cognitieve stijl* genoemd (Cronbach & Snow, 1977). Een leerstijl is dus de expressie van een cognitieve stijl bij het leren of studeren. Het is de vraag of zulke voorkeuren zich wel gemakkelijk laten veranderen. Hoewel beide gezichtspunten theoretisch fundamenteel lijken te verschillen, zou men in dit geval kunnen stellen dat een leerstijl volgens Vermunt de meer klassieke leerstijlconceptie omvat (met name wordt dit duidelijk in de betekenisgerichte leerstijl – zie verder); Vermunt's leerstijlen zijn, althans ten dele, klassieke leerstijlen plus nog iets.

Hoe dit ook zij, de inrichting van het onderwijs kan het toepassen van een bepaalde leerstijl bevorderen of ontmoedigen. Is het onderwijs zo ingericht dat vooral het onthouden van feitjes wordt beloond, dan wordt de toepassing van een op reproductie gerichte leerstijl bevorderd en wordt de toepassing van mogelijke andere leerstijlen ontmoedigd. Op die manier heeft het onderwijs invloed op de selectie van leerstijlen door leerlingen en studenten.

Het constateren van een bepaalde frequent voorkomende (ongewenste) leerstijl bij studenten of leerlingen kan dus een indicatie opleveren van de (gebrekkige) kwaliteit van het onderwijs (Vermunt, 1995).

Een andere toepassing van de kennis omtrent de leerstijlen van studenten of leerlingen ligt in de mogelijkheid de onderwijsmethoden doelbewust te laten aansluiten bij die leerstijlen. Idealerweise zou dit betekenen dat het leerlingbestand wordt opgesplitst in naar leerstijl homogene subgroepen. Deze subgroepen zouden van bij hen passend onderwijs kunnen worden voorzien, wat de efficiëntie van het onderwijs zou kunnen verhogen. Dit is de gedachte die ten grondslag ligt aan het idee van *aptitude-treatment* interacties (Cronbach & Snow, 1977). Vermunt (1995) is daar nogal optimistisch over. De resultaten van Cronbach & Snow (1977) lijken zo'n optimisme echter niet te wettigen. Dat de situatie op dit moment niet veel veranderd is, moge blijken uit de van tijd tot tijd opvlammende discussies op Internet (bijvoorbeeld op discussielijst AERA-C; februari 1995) waarin de scepsis over de haalbaarheid of het nut van zo'n aanpak de boventoon voert.

Welke praktische toepassingen van leerstijldiagnostiek men ook overweegt, zulke toepassingen vallen of staan met (a) de robuustheid van de onderscheiden leerstijlen, (b) het aantoonbaar effect van bepaalde leerstijlen op de behaalde leerresultaten, en (c) de veranderbaarheid van bepaalde inadequate leerstijlen in andere, meer adequate leerstijlen. Het huidige onderzoek is bedoeld om in de kwesties (a) en (b) enige helderheid te verkrijgen. Daarnaast dient te worden onderzocht in hoeverre een diagnostiek van leerstijlen de voorkeur verdient boven andere, eveneens gangbare methoden om studeermethoden van studenten in kaart te brengen. Hiertoe zullen de volgende onderzoeksvragen worden beantwoord:

1. Zijn de leerstijlen zoals die door Vermunt (1992) worden gedefinieerd betrouwbaar bij verschillende groepen universitaire studenten terug te vinden?
2. Wat is de voorspellende waarde van deze leerstijlen voor de studievoortgang van studenten?
3. Hoe verhoudt zich de diagnostiek van leerstijlen tot die van enkelvoudige studeergedragingen en/of -problemen?

LEERSTIJLEN VOLGENS VERMUNT

Vermunt (1992) vat een leerstijl op als een conglomeraat van drie componenten:

1. Cognitieve verwerkingsstrategieën
2. Metacognitie: mentale leermodellen en regulatiestrategieën
3. Leeroriëntaties

Deze drie componenten zijn op verschillende wijze verweven in vier door hem onderscheiden leerstijlen: (a) de betekenisgerichte leerstijl, (b) de reproductiegerichte leerstijl, (c) de toepassingsgerichte leerstijl, en (d) een ongerichte leerstijl.

De *cognitieve verwerkingsstrategieën* gaan terug op de ideeën van Pask (1988) over een serialistische/analytische versus een holistische/synthetische aanpak en op het factoranalytische werk van Schmeck (1983), die vier stijlen onderscheidt: (1) diepe verwerking, (2) elaboratieve

verwerking, (3) onthouden van feiten, en (4) methodisch studeren/traditionele studievaardigheid.

Vermunt (1992) onderscheidt in dit kader:

- *diepteverwerking*, voornamelijk ontleend aan de diepe verwerking van Schmeck (1983) en de holistische/synthetische aanpak van Pask (1988). Vermunt onderscheidt hierin: (a) relateren/structureren, en (b) kritisch verwerken.
- *stapsgewijze verwerking*, voornamelijk ontleend aan de serialistische/analytische aanpak van Pask (1988) en zowel de elaboratieve verwerking als het onthouden van feiten van Schmeck (1983). Vermunt maakt hier een onderscheid tussen: (c) analyseren, en (d) memoriseren.
- *concrete verwerking*: (e) concretiseren/toepassen.

Aan *metacognitie*, in de opvatting van Vermunt, zijn twee aspecten te onderscheiden. Ten eerste gaat het hier om kennis en opvattingen over de eigen leer- en denkprocessen of *mentale leermodellen* (Marton & Säljö, 1984; Flavell, 1987). Individuele verschillen op dit punt zouden samenhangen met verschillend gedrag in leersituaties (Vermunt, 1995). Vermunt (1992) onder-

Tabel 1. Factorstructuur ILS volgens Vermunt (Vermunt, 1992).

ILS-schalen	F1	F2	F4	F3
<i>Verwerkingsstrategieën</i>				
Diepteverwerking				
V1: relateren/structureren	+++			
V5: kritisch verwerken	+++			
Stapsgewijze verwerking				
V2: memoriseren/herhalen		+++		
V4: analyseren		+++		
V3: Concrete verwerking	++		+	
<i>Regulatiestrategieën</i>				
Zelfsturing				
V7: leerproces/-resultaat	+++			
V8: leerinhoud	+++			
Externe sturing				
V6: leerproces		+++		
V10: leerresultaten		++		
V9: Stuurloos leergedrag				+++
<i>Mentale leermodellen</i>				
V13: opbouw van kennis	+++			
V14: opname van kennis	-	++		+
V15: gebruik van kennis			+++	
V12: stimulerend onderwijs				+++
V11: samen studeren				++
<i>Leeroriëntaties</i>				
V19: persoonlijk geïnteresseerd	++			
V18: certificaatgericht	-	+	+	
V16: testgericht				+
V17: beroepsgericht			+++	
V20: ambivalent				++
Percentage verklaarde variantie:	21%	15%	6%	10%

Noot: +++ Ladingen > .70

++ Ladingen > .50

+ Ladingen > .30

--- Ladingen < -.70

-- Ladingen < -.50

- Ladingen < -.30

scheidt de volgende vijf verschillende mentale leermodellen: (a) leren als het *opbouwen* van kennis teneinde de werkelijkheid beter te begrijpen, (b) leren als het *opnemen* en opslaan van de door het onderwijs aangeboden kennis, (c) leren als het *gebruiken* en toepassen van aanwezige kennis. Ook de opvatting (d) dat leren neerkomt op het consumeren van *stimulerend onderwijs*, of (e) dat leren vooral *samen* met anderen tot stand komt, behoort tot de mentale leermodellen.

Daarnaast behoren strategieën voor de *regulering of sturing* van leer- en denkprocessen tot de metacognitie. Het gaat hierbij enerzijds om het verschil tussen *zelfsturing* en *externe sturing*, en anderzijds om het überhaupt al dan niet toepassen van sturingsstrategieën; in het laatste geval spreekt Vermunt van *stuurloos leergedrag*. Zulke verschillen zouden volgens Brown (1987) en De Jong (1992) sterk samenhangen met leerprestaties (in: Vermunt, 1995).

De door Vermunt onderscheiden *leeroriëntaties* of studiemotieven tenslotte zijn terug te voeren op het werk van Entwistle & Ramsden (1983) en van Biggs (1984). In plaats van de door hen onderscheiden intrinsieke, reproductiegerichte, prestatiegerichte en gebrekkige studiemotieven, prefereert Vermunt een indeling in (a) persoonlijke interesse, (b) gerichtheid op het behalen van een diploma, (c) de wens om eigen capaciteiten uit te testen, (d) gerichtheid op een beroep, en tenslotte (e) ambivalentie.

Specifieke combinaties van al deze componenten (zie Tabel 1) leveren de leerstijlen op zoals Vermunt die onderscheidt. Daarmee krijgt het begrip *leerstijl*, zoals eerder vermeld, opzettelijk (Vermunt, 1992) een bredere betekenis dan het in de overige literatuur heeft. Wel wordt vooral aansluiting gezocht bij de wetenschappelijk georiënteerde Engelse en Amerikaanse onderwijspsychologie. Wat meer exotische denkbeelden, zoals die van Kolb (1976) en Dunn, Dunn & Price (1979) worden wel door hem besproken (Vermunt, 1992), maar niet nagevolgd.

De hier besproken componenten vinden hun operationalisering in de 20 schalen van de *Inventaris Leerstijlen ILS* (Van Rijswijk & Vermunt, 1987). De door Vermunt onderscheiden leerstijlen komen hieruit naar voren als de eerste vier (Varimax geroteerde) hoofdcomponenten. Tabel 1 reproduceert op schematische wijze de betreffende factorstructuur (Vermunt, 1992, tabel 5.8), zoals die werd gevonden bij 795 eerstejaarsstudenten van diverse studierichtingen aan de Katholieke Universiteit Brabant te Tilburg.

METHODE

Op grond van het proefschrift van Vermunt (1992) werd een verkorte versie van de *Inventaris Leerstijlen ILS* geconstrueerd, voor elk van de 20 schalen bestaande uit de naar factorzuiverheid en homogeniteit beste drie items (Vermunt, 1992, tabellen 5.1 t/m 5.4). Met slechts drie items per schaal is de interne consistentie van de schalen natuurlijk zwak; door de beste items te selecteren blijft de essentie van elke schaal echter tot op zekere hoogte behouden. De aldus verkorte ILS bestond dus uit 60 items; de betreffende schalen zijn gecodeerd als v1 t/m v20 (zie Tabel 1). Deze verkorte ILS werd, samen met andere vragenlijsten, afgenomen bij de volgende groepen universitaire studenten:

1. 176 deelnemers aan verschillende vaardigheidscursussen bij de centrale afdeling Studie Ondersteuning van de Rijksuniversiteit Groningen, van verschillende jaargangen en afkomstig uit verschillende studierichtingen (*steekproef 1*).
2. 169 eerstejaarsstudenten Bedrijfskunde (40% van het cohort 1993/94) aan de Rijksuniversiteit Groningen (*steekproef 2*).
3. 270 eerstejaarsstudenten Economie (91% van het cohort 1994/95) aan de Rijksuniversiteit Groningen (*steekproef 3a*).
4. 48 eerstejaarsstudenten Econometrie (het gehele cohort 1994/95) aan de Rijksuniversiteit Groningen (*steekproef 3b*). Voor de beantwoording van de eerste en derde onderzoeksvraag werden beide steekproeven 3a en 3b samengevoegd tot één steekproef 3 (n=318).

De toegevoegde vragenlijsten waren:

- De Studie Management en Academische Resultaten Test SMART (Klein, et al., 1994) met

als schalen Academische competentie (AC), Tentamencompetentie (TC), Time-management (TM) en Strategisch studeren (SS).

- De Vragenlijst Studeerproblemen VSP (Schouwenburg, 1994) met als schalen Lage werkdiscipline (LWD), Faalangst (FAA) en Desinteresse in de studie (DIS).
- De Nederlandse vertaling van Lay's Procrastination Scale LPS (Schouwenburg, 1994) die de neiging tot uitstellen meet.
- Negen zelfbeoordelingsschaaltjes betrekking hebbend op cognitieve verwerking en studeergedrag, ontleend aan Leren & Studeren (Vorst, 1993). Deze schaaftjes zijn: Effectief leesgedrag (EL), Structuur herkennen (SH), Argumenten analyseren (AA), Verbanden leggen (VL), Structuur Aanbrengen (SA), Kennis Controleren (KC), Planning en tijdbeheersing (PT), Werkomstandigheden inrichten (WI) en Zichzelf beheersen (ZB).

Tabel 2. Factorstructuur ILS in drie steekproeven.

		Betekenis- gericht	Reproductie- gericht	Toepassings- gericht	Ongericht
ILS-schalen	Steekproef:	-1-2-3-	-1-2-3-	-1-2-3-	-1-2-3-
<i>Verwerkingsstrategieën</i>					
Diepteverwerking					
V1: relateren/structureren		a b a			.-c .
V5: kritisch verwerken		a b b			
Stapsgewijze verwerking					
V2: memoriseren/herhalen			b a a		
V4: analyseren			a a b		
V3: Concrete verwerking		b b b		c c .	
<i>Regulatiestrategieën</i>					
Zelfsturing					
V7: leerproces/-resultaat		b b a	c . .		
V8: leerinhoud		b a b			
Externe sturing					
V6: leerproces			b b c	. . c	
V10: leerresultaten			c . .	c . c	. c .
V9: Stuurloos leergedrag					b a b
<i>Mentale leermodellen</i>					
V13: opbouw van kennis		b b b			
V14: opname van kennis		.-c .	c c b		c . .
V15: gebruik van kennis				c b b	c . .
V12: stimulerend onderwijs					b . b
V11: samen studeren				. c c	b . c
<i>Leeroriëntaties</i>					
V19: persoonlijk geïnteresseerd		c . c	. c .	c . c	-c .
V18: certificaatgericht		-b-b .	. . b	. c .	c . .
V16: testgericht			. c .	. b c	b . c
V17: beroepsgericht				b a b	c . .
V20: ambivalent				-b . .	c a a
Percentage verklaarde variantie:		18% 18% 18%	9% 13% 13%	7% 9% 8%	12% 7% 7%

Noot: a: Ladingen > .70
b: Ladingen .50 - .69
c: Ladingen .30 - .49

De vragenlijsten werden bij steekproef 1 in de loop van het eerste halfjaar van 1994 afgenomen, bij steekproef 2 in april 1994 en bij steekproef 3 in september 1994. De gegevens van steekproef 2 werden aan het eind van het studiejaar gecompleteerd met de behaalde *studieresultaten* over dat jaar; de gegevens van de steekproeven 3a en 3b werden aan het eind van het kalenderjaar 1994 aangevuld met de studieresultaten over het eerste trimester.

RESULTATEN

De repliceerbaarheid van leerstijlen

Op de ILS-data van elk van de drie steekproeven werd een componentenanalyse met Varimaxrotatie uitgevoerd waarbij telkens vier factoren werden geëxtraheerd. In schematische vorm wordt het resultaat daarvan weergegeven in Tabel 2. In alle drie gevallen zijn de vier factoren goed interpreteerbaar in termen van Vermunt's leerstijlen. Onderlinge vergelijking, alsmede vergelijking met de door Vermunt gevonden structuur (Tabel 1), laat zien dat de factorstructuren sterk overeenkomen. Daarmee is de repliceerbaarheid van Vermunt's leerstijlen aangetoond.

Een dergelijke vergelijking maakt het ook mogelijk de vier leerstijlen naar hun kernvariabelen operationeel te differentiëren. Zo kan de *betekenisgerichte* leerstijl worden gedefinieerd aan de hand van de 6 ILS-variabelen Relateren/structureren (V1), Concrete verwerking (V3), Kritisch verwerken (V5), Zelfsturing leerproces/-resultaat (V7), Zelfsturing leerinhoud (V8) en Opbouw van kennis (V13). De 4 ILS-variabelen Memoriseren/herhalen (V2), Analyseren (V4), Externe sturing leerproces (V6) en Opname van kennis (V14) definiëren dan de *reproductiegerichte* leerstijl. De twee ILS-variabelen Gebruik van kennis (V15) en Beroepsgericht (V17) definiëren de *toepassingsgerichte* leerstijl en de twee ILS-variabelen Stuurloos leergedrag (V9) en Ambivalent (V20) tenslotte de *ongerichte* leerstijl. Om leerstijl-scores te verkrijgen worden naar algemeen gebruik de scores op de betreffende definiërende variabelen bij elkaar opgeteld.

Een en ander betekent tevens dat 6 van de 20 ILS-schalen (V10, V11, V12, V16, V18 en V19) voor de differentiatie tussen de verschillende leerstijlen overbodig zijn. De leerstijlen, op deze manier gedefinieerd, blijken onderling maximaal $-.30$ en $.22$ te correleren en gemiddeld respectievelijk $.00$, $.01$ en $.01$. Ze zijn dus onafhankelijk te noemen.

Samenhang met overige studeervariabelen

Voor elk van de drie steekproeven werd vervolgens een componentenanalyse met Varimaxrotatie uitgevoerd over de als boven gedefinieerde vier leerstijl-scores plus alle overige studeergedrag-variabelen. Op grond van de scree-test werden vier factoren geëxtraheerd die met elkaar respectievelijk 55%, 56% en 56% van de variantie verklaren. De drie verkregen factorstructuren worden op dezelfde manier als bij Tabel 2 schematisch in Tabel 3 weergegeven.

Uit Tabel 3 blijkt dat alle time-management-achtige of zelfregulatie-variabelen clusteren in één factor (F1), die betrekkelijk onafhankelijk is van leerstijlen volgens Vermunt. Vanuit een persoonlijkheidspsychologisch perspectief is duidelijk dat het hier gaat om de persoonlijkheidsdimensie Zorgvuldigheid of *Conscientiousness* (Schouwenburg, 1994), naast Extraversie/introversie, Vriendelijkheid, Emotionele stabiliteit en Openheid één van de vijf gronddimensies van de persoonlijkheid, de *Big five* (zie o.a. Digman, 1990; John, 1990), een factor die duidelijk gerelateerd is aan studiesucces (zie o.a. Schouwenburg & Vugteveen, 1995).

Vermunt's *betekenisgerichte* leerstijl definieert de tweede factor (F2), samen met enkele studeervariabelen waarvan Argumenten analyseren en Verbanden leggen uit de vragenlijst *Leren & Studeren* (Vorst, 1993) de meest prominente zijn. Deze schalen worden door Vorst als volgt omschreven:

- *Argumenten analyseren*: Diep ingaan op de leerstof, kritisch kunnen lezen, de tijd nemen conclusies te controleren, argumenten te beoordelen en redeneringen op juistheid te toetsen.
- Verbanden leggen*: Verbanden leggen tussen verschillende cursussen onderling, gegevens uit andere bronnen en feiten uit het dagelijks leven.

Tabel 3. Leerstijlen en studeergedrag.

	F1	F2	F3	F4
<i>Steekproef:</i>	-1-2-3-	-1-2-3-	-1-2-3-	-1-2-3-
Lage werkdiscipline	-a-a-a			
Time-management	a a a			
Planning en tijdbeheersing	a a a			
Uitstelneiging	-a-b-a			
Zichzelf beheersen	a a a			
Werkomstandigheden inrichten	c b b			
Kennis controleren	c c c	b c b		
<i>Betekenisgerichte leerstijl</i>		b a a		c . .
Argumenten analyseren		a b b		
Verbanden leggen	. b .	b b b	. . c	c b .
Structuur herkennen	. . c	b b c	b . c	
Strategisch studeren	. c c	b c c		
Structuur aanbrengen		b . .		
Faalangst	-c . .		-b a-a	
Tentamencompetentie		c . .	b-a a	
<i>Reproductiegerichte leerstijl</i>	. b c		-c c-c	. . c
<i>Ongerichte leerstijl</i>	. .-c		-c-a-b	-b-c .
Effectief leesgedrag		. b .	b . c	
Academische competentie	. .c		.-b c	a b b
Desinteresse in de studie				-a-a-b
<i>Toepassingsgerichte leerstijl</i>			-c . .	. b a
<i>Percentage verkl. variantie:</i>	22%27%32%	16%14%10%	10%9%9%	8%7%6%

Noot: a: Ladingen > .70
b: Ladingen .50 - .69
c: Ladingen .30 - .49

De lineaire combinatie van deze twee schalen correleert respectievelijk .57, .60 en .60 met Vermunt's betekenisgerichte leerstijl. Gegeven de evidente lage betrouwbaarheid van deze twee 7-punts schaaltes, betekent dit waarschijnlijk dat de combinatie van Argumenten analyseren en Verbanden leggen de betekenisgerichte leerstijl goed benadert.

Vermunt's *reproductiegerichte* leerstijl en *ongericte* leerstijl laden vooral op een factor (F3) die gekenmerkt wordt door Faalangst en Tentamencompetentie. Deze factor lijkt overeen te komen met de *Big-five* dimensie Emotionele stabiliteit, een dimensie waarop faalangst steevast een zeer hoge lading heeft (Schouwenburg, 1994). Tentamencompetentie is daarbij als het ware het spiegelbeeld van faalangst, geïnterpreteerd als tentamenangst.

Een vierde factor tenslotte (F4) lijkt een factor te zijn van studiemotivatie; de factor wordt gekenmerkt door Academische competentie en door Desinteresse in de studie. Op deze factor laadt vooral Vermunt's *toepassingsgerichte* leerstijl. Dit duidt erop dat er enige systematische samenhang is tussen een min of meer beroepsgerichte kijk op studeren en interesse in de studie en het gevoel de studie aan te kunnen.

Voorspellende waarde voor studieresultaat

In drie van de vier steekproeven, steekproef 2 en de steekproeven 3a en 3b, zijn gegevens over studieresultaten voorhanden, in de vorm van het totaal aantal behaalde studiepunten op een

Tabel 4. Correlaties leerstijlen met studievoortgang.

Leerstijl volgens Vermunt:	Steekproef 2: Bedrijfskunde	Steekproef 3a: Economie	Steekproef 3b: Econometrie
- Betekenisgerichte leerstijl	-.04	-.01	.04
- Reproductiegerichte leerstijl	.12	.05	-.02
- Toepassingsgerichte leerstijl	-.16*	-.02	-.17
- Ongerichte leerstijl	-.25**	-.13	-.24

Noot: * significant ($p < .05$)

** significant ($p < .01$)

bepaald tijdstip. Het aantal behaalde studiepunten is een goede operationalisering van de studievoortgang. Tabel 4 toont de correlaties van de vier leerstijlen volgens Vermunt met deze studieresultaten.

Tabel 4 laat zien dat leerstijlen, gedefinieerd volgens Vermunt, geen voorspellende waarde blijken te hebben voor het studieresultaat in termen van voortgang. Dit op het eerste gezicht wat teleurstellende resultaat betekent natuurlijk dat *gegeven het huidige onderwijs* aan eerstejaarsstudenten bij Bedrijfskunde en Economie/Econometrie leerstijlen van de studenten zelf er nauwelijks toe doen; hoogstens het bezit van een ongerichte leerstijl wijst enigszins op het risico van studieovertraging.

Welke studeervariabelen dan wel met studiesucces samenhangen, blijkt uit Tabel 5, waarin de correlaties van de overige gemeten variabelen met studievoortgang worden weergegeven.

Uit deze tabel kunnen tenminste drie conclusies worden getrokken. Allereerst blijkt dat zelfregulatie-variabelen, zoals time-management, werkdiscipline, planning en zelfbeheersing in

Tabel 5. Significante ($p < .05$) correlaties studergedrag met studievoortgang.

Vragenlijst/variabele:	Steekproef 2: Bedrijfskunde	Steekproef 3a: Economie	Steekproef 3b: Econometrie
SMART:			
- Academische competentie	.23	—	.41
- Tentamencompetentie	.46	—	—
- Time-management	.42	.22	.32
- Strategisch studeren	—	—	—
VSP/LPS:			
- Lage werkdiscipline	-.45	-.29	-.38
- Faalangst	—	—	—
- Desinteresse in de studie	—	—	—
- Uitstelneiging	-.19	—	—
Leren & Studeren:			
- Effectief leesgedrag	.22	—	—
- Structuur herkennen	.17	—	—
- Argumenten analyseren	—	—	—
- Verbanden leggen	—	—	—
- Structuur aanbrengen	—	—	—
- Kennis controleren	.27	.14	—
- Planning en tijdgebruik	.34	.22	—
- Werkomstandigheden inrichten	—	—	—
- Zichzelf beheersen	.27	.24	.30

de drie steekproeven relatief sterk met studievoortgang samenhangen. Vervolgens blijken competentie-oordelen van studenten, zoals gemeten met de SMART, wel bij Bedrijfskunde en Econometrie, maar niet bij Economie met studieresultaten samen te hangen. Tenslotte blijken de cognitieve variabelen, zoals gemeten met het eerste gedeelte van Leren & Studeren, nauwelijks met de studieresultaten te correleren. Waarschijnlijk is er, gegeven het relatief homogene intelligentieniveau van universitaire studenten, weinig extra effect van individuele verschillen in cognitieve strategie te verwachten. Dit resultaat staat overigens in contrast met de bevindingen van Schouwenburg & Schilder (1994) die bij eerstejaars HBO-studenten Informatica wel een duidelijk effect van cognitieve variabelen op studieresultaten constateerden.

Behalve deze algemene trends tonen de drie profielen van Tabel 5 specifieke verschillen tussen de drie studierichtingen, waarvan de aard wellicht nadere bestudering verdient.

DISCUSSIE

De hier gepresenteerde onderzoeksresultaten vertonen in hun grote lijn een opmerkelijke constantie. In de Tabellen 2 t/m 5 overheerst de overeenkomst van de resultaten in de verschillende steekproeven. Dit is vooral opmerkelijk omdat de betreffende steekproeven geen van alle 'random' zijn maar op verschillende manieren nogal selectief of gebrekkig. De hiermee aangetoonde onderlinge replicerbaarheid laat zien dat de aangetroffen effecten behoorlijk robuust zijn. Dit stelt ons in staat de drie onderzoeksvragen met vertrouwen te beantwoorden.

1. De vier leerstijlen volgens Vermunt zijn betrouwbaar bij verschillende groepen universitaire studenten terug te vinden (de Tabellen 1 en 2). Daarbij zijn ze zonder overlap operationeel te definiëren; deze operationele definities blijken empirisch onafhankelijk.
2. De voorspellende waarde van de leerstijlen volgens Vermunt voor de studievoortgang van studenten is hooguit gering.
3. Deze leerstijlen zijn onafhankelijk van een zelfregulatiecluster van studeervariabelen als werkdiscipline, time-management, studieplanning etc. De betekenisgerichte leerstijl correleert sterk met cognitieve studeervariabelen als argumenten kunnen analyseren, verbanden kunnen leggen, structuur kunnen herkennen, etc. Daarentegen hangen zowel de reproductiegerichte als de ongerichte leerstijl samen met tentamencompetentie en faalangst. De toepassinggerichte leerstijl tenslotte blijkt samen te hangen met interesse in de studie en met het gevoel de studie aan te kunnen.

In essentie gaat het bij een leerstijl in engere zin om de expressie van een cognitieve stijl bij het leren en studeren. Cognitieve verwerkingsstrategieën vormen dus de kern van het leerstijlconcept. Zo gezien kunnen alleen Vermunt's betekenisgerichte en de reproductiegerichte stijl als 'echte' leerstijlen worden aangemerkt. Deze twee stijlen nu corresponderen met een gangbare normatieve opvatting over studeren aan de universiteit. Volgens deze opvatting zou universitair studeren vooral moeten betekenen: diep en kritisch op de stof ingaan, een brede kennis opbouwen, kortom, *mastery* verwerven. Bij zo'n opvatting past natuurlijk ook actief en intrinsiek gemotiveerd de studie aanpakken. Het huidige studeren zou daarentegen te passief plaatsvinden, teveel gericht zijn op memoriseren en op het kunnen reproduceren van het geleerde op het tentamen. In termen van Vermunt is betekenisgericht studeren 'goed' en reproductiegericht studeren 'slecht'.

Bij de betekenisgerichte leerstijl gaat het in essentie om de cognitieve verwerkingsvariabelen Argumenten Analyseren en Verbanden Leggen, dwz. intelligentie-achtige capaciteiten (Messick, 1994). Ook uit ander onderzoek (Van der Luit, 1994) blijkt een zeer hoge correlatie tussen de voor een betekenisgerichte leerstijl essentiële ILS-variabele Diepteverwerking en de variabelen Argumenten Analyseren en Verbanden Leggen.

Daarentegen gaat het bij reproductiegericht leren om volgzzaam, schools gedrag. Gegeven de genetische invloed op intelligentie (Carroll, 1993) zou men kunnen betogen dat het zinloos is een overwegend reproductiegerichte leerstijl trachten om te vormen tot een overwegend beteke-

nisgerichte leerstijl. Aan de andere kant kan betoogd worden dat de capaciteiten in kwestie behoren tot het eerste stratum in Carroll's theorie, als een onderafdeling van *fluid intelligence* uit het tweede stratum, capaciteiten die waarschijnlijk beter door training kunnen worden beïnvloed dan capaciteiten uit de hogere strata (Carroll, 1993).

Uit de hier toegepaste operationalisatie van Vermunt's leerstijlen blijkt wel dat het grote contrast zich vooral voordoet tussen de *betekenisgerichte* en de *reproductiegerichte* leerstijl. De betekenisgerichte leerstijl komt duidelijk overeen met een academisch ideale vorm van studeren, terwijl de reproductiegerichte leerstijl een min of meer laakbare instelling van studenten vertegenwoordigt. De beide overige door Vermunt onderscheiden leerstijlen verdienen dit predikaat eigenlijk niet, aangezien cognitieve verwerkingsstrategieën niet tot de definitie van deze stijlen behoren. Het zijn daarom geen echte leerstijlen, maar eerder motivatie-varianten: een beroepsgerichte en een ongemotiveerde variant. Daarmee verleent ook Vermunt ongewild grond aan het door Schmeck (1988) ooit uitgesproken vermoeden dat er eigenlijk maar sprake is van *twee* verschillende leerstijlen, die sterk overeenkomen met de twee oorspronkelijk door Pask (1988) voorgestelde.

De voorspellende waarde van leerstijlen volgens Vermunt voor studiesucces in de propaedeutische blijkt vrijwel nihil. Dat succes wordt veel beter voorspeld door het zelfregulatie-cluster en (in sommige studies) het gevoel de studie aan te kunnen (Academische competentie) (Tabel 5). Wil men het begrip *leerstijl* als manifestatie van persoonlijkheidsvariabelen bij het leren en studeren (Messick, 1994) handhaven, dan zou het de voorkeur verdienen zulke leerstijlen mede aan de hand van zelfregulatie-variabelen te definiëren.

Een interessante variant op het idee van *aptitude-treatment* interacties zou kunnen resulteren uit het vermoeden dat door studenten geselecteerde leerstijlen afhankelijk zijn van bemoediging of afkeuring in het onderwijs dat ze concreet volgen. Dat bijvoorbeeld de betekenisgerichte leerstijl zo laag correleert met studieresultaten zou dan betekenen dat deze leerstijl in het betreffende onderwijs voor een beter resultaat niet nodig is. De lage voorspellende waarde van de betekenisgerichte leerstijl zou dan ook een indicatie kunnen opleveren van de wenselijkheid dat betreffende onderwijs zodanig te herzien dat een betekenisgerichte leerstijl bij studenten wel wordt gehonoreerd. Leerstijldiagnostiek zou op die manier kunnen worden aangewend om de kwaliteit van het onderwijs te beoordelen.

Tenslotte kan het nuttig zijn zich te realiseren dat het huidige onderzoek naar leerstijlen zich (met uitzondering van steekproef 1) heeft beperkt tot studenten in de propaedeutische. Het kan daarom zo zijn dat in feite alleen de *output* van de leergeschiedenis in het Voorbereidend Wetenschappelijk Onderwijs is gemeten. Voor de aldaar aangeleerde manier van studeren zijn leerstijlen, en met name de betekenisgerichte leerstijl, misschien niet relevant maar gaat het meer om hard en gedisciplineerd werken. Dit zou kunnen betekenen dat, gegeven de pretenties van het universitaire onderwijs, een betekenisgerichte leerstijl in hogere studiejaren wel voor de studieresultaten relevant zou zijn. Nader onderzoek zal hierover uitsluitsel moeten geven.

DANKBETUIGING

De auteur dankt Wim Moes van de Faculteit Bedrijfskunde en Ineke Vugteveen van de Faculteit der Economische Wetenschappen aan de Rijksuniversiteit te Groningen voor hun participatie in het onderzoek en het beschikbaar stellen van het onderzoeksmateriaal. Ook gaat de dank van de auteur uit naar een anonieme reviewer voor het constructieve commentaar op een eerdere versie van dit artikel.

LITERATUUR

- Biggs, J.B. (1984). Learning strategies, student motivation patterns, and subjectively perceived success. In J.R. Kirby (ed.), *Cognitive strategies and educational performance*. New York: Academic Press.
- Brown, A.L. (1987). Metacognition, executive control, self-regulation and other more mysterious mechanisms. In F.E. Weinert & R.H. Kluwe (eds.), *Metacognition, motivation and understanding*. Hillsdale, NJ: Erlbaum.

- Carroll, J.B. (1993). *Human cognitive abilities; A survey of factor-analytic studies*. Cambridge: Cambridge University Press.
- Cronbach, L.J. & Snow, R.E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irving.
- Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417-440.
- Entwistle, N.J. & Ramsden, P. (1983). *Understanding student learning*. London: Croom Helm.
- Flavell, J.H. (1987). Speculations about the nature and development of metacognition. In F.E. Weinert & R.H. Kluwe (eds.). *Metacognition, motivation, and understanding*. Hillsdale NJ: Erlbaum.
- John, O.P. (1990). The 'big five' factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L.A. Pervin (ed.). *Handbook of personality: Theory and research*. New York: Guilford.
- Jong, F.P.C.M. de (1992). *Zelfstandig leren; regulatie van het leerproces en leren reguleren: een procesbenadering*. Katholieke Universiteit Brabant: Academisch proefschrift.
- Klein, W.C., Topman, R.M. & van der Ploeg, H.M. (1994). Cognities, studiegewoonten en academische prestatie: De ontwikkeling van de Studie-Management en Academische Resultaten Test (SMART). *Nederlands Tijdschrift voor de Psychologie*, 49, 231-232.
- Luit, H. Van der (1994). *Een vergelijking van twee vragenlijsten naar de studievoordigheid van studenten*. Intern rapport, Universiteit van Amsterdam, Faculteit der Psychologie, Vakgroep Psychologische Methoden.
- Marton, F. & Säljö, R. (1984). Approaches to learning. In F. Marton, D. Hounsell & N.J. Entwistle (eds.). *The experience of learning*. Edinburgh: Scottish Academic Press.
- Messick, S. (1994). The matter of style: Manifestations of personality in cognition, learning, and teaching. *Educational Psychologist*, 29, 121-136.
- Pask, G. (1988). Learning strategies, teaching strategies, and conceptual or learning style. In R.R. Schmeck (ed.). *Learning strategies and learning styles*. New York: Plenum.
- Rijswijk, F. van & Vermunt, J. (1987). *Vaardig studeren; Studie-aanpak adviespakket voor Ou-studenten. Deel 1: Inventaris Leerstijlen*. Heerlen/Tilburg: Open universiteit/Katholieke Universiteit Brabant.
- Schmeck, R.R. (1983). Learning styles of college students. In R. Dillón & R.R. Schmeck (eds.). *Individual differences in cognition, volume 1*. New York: Academic Press.
- Schmeck, R.R. (1988). Strategies and styles of learning: An integration of varied perspectives. In R.R. Schmeck (ed.) *Learning strategies and learning styles*. New York: Plenum.
- Schouwenburg, H.C. (1994). *Uitstelgedrag bij studenten*. Rijkuniversiteit Groningen: Academisch proefschrift.
- Schouwenburg, H.C. & Schilder, A. (1994). *Verslag van een experiment met studievoordigheidstraining voor het cohort 1993 van de opleiding Informatica van de Hanzehogeschool, Hogeschool van Groningen, sector Natuur & Techniek, november-januari 1993/94*. Intern rapport, Rijksuniversiteit Groningen, afdeling Studie Ondersteuning.
- Schouwenburg, H.C. & Vugteveen, T.H. (1995). Voorspellende waarde van de Vragenlijst Studieproblemen voor het studiesucces van eerstejaars studenten. In H.C. Schouwenburg & J.T. Groenewoud (red.). *Studievoordigheid en leerstijlen; Proceedings van de 14e Landelijke Dag Studievoordigheden te Groningen*. Groningen: Wolters-Noordhoff.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs: Naar procesgerichte instructie in zelfstandig denken*. Amsterdam/Lisse: Swets & Zeitlinger.
- Vermunt, J.D.H.M. (1995). Leerstijlen: Een overzicht. In H.C. Schouwenburg & J.T. Groenewoud (red.). *Studievoordigheid en leerstijlen; Proceedings van de 14e Landelijke Dag Studievoordigheden te Groningen*. Groningen: Wolters-Noordhoff.
- Vorst, H.C.M. (1993). *Leren & Studeren; Handleiding en verantwoording bij vragenlijst en cursus studievoordigheid L&S*. Nijmegen: Berkhout.

Manuscript ontvangen: 10-3-1995

Definitieve versie ontvangen: 4-9-1995

Gedrag van aanstaande docenten in aandachtseisende situaties in de klas

Wilfried Admiraal¹, Theo Wubbels en Fred Korthagen

Interfacultair Instituut voor Lerarenopleiding, Studievastigheden en Onderwijsontwikkeling (IVLOS), Universiteit Utrecht

ABSTRACT

In this study, we describe student teacher behaviour in response to stressful classroom events from three perspectives: the coping perspective, representing the meaning of the behaviour in the light of the situations in which it is shown; the pedagogical-methodological perspective, evident in an analysis of the instructional meaning of teacher behaviour; and the interpersonal perspective, meaning behaviour perceived as part of the teacher-pupil relationship.

Observation of lessons of student teachers yields descriptions of teacher behaviour in four types of stressful classroom events. Two dimensions of student teacher behaviour have been extracted: the 'active-passive'-dimension, describing the degree of intensity of a student teacher's actions when transferring (verbal) messages to the pupils; and the 'agitation-harmony'-dimension, defining the degree of tension the student teacher arouses in the relationship between teacher and pupils.

Two conclusions are developed. Firstly, in responding to pupils' off-task behaviour, student teachers generally show communication similar to the pupils' behaviour. This means that they are engaged in the same oppositional behaviour as the pupils, which can result in an aggressive symmetrically escalating relationship. Secondly, the newly developed descriptive framework with the two dimensions can be used in future research, and in helping supervisors to reveal student teachers' blind spots for what is going on in the classroom during stressful events.

INLEIDING

In het rapport van de Commissie Toekomst Leraarschap, 'Het Gedroomde Koninkrijk' (1993), is aandacht besteed aan de psychische belasting die het beroep van docent met zich meebrengt. De commissie concludeert in haar rapport dat vooral het directe contact met leerlingen het beroep zo zwaar maakt (p. 21). Zij merkt tevens op dat dit directe contact met leerlingen door veel docenten ook wordt gezien als een bron van arbeidsvreugde. Dit enigszins paradoxale gegeven zou kunnen betekenen dat niet zozeer gebeurtenissen tijdens de les gevoelens van belasting veroorzaken, maar inadequate reacties van docenten op deze situaties of een gebrekkige verwerking van de gebeurtenissen.

Het gegeven dat docenten vooral het contact met leerlingen als een belasting ervaren, is niet uniek voor Nederland. Onderzoek in Duitsland (Rudow, 1990), Japan (Ninomiya & Okato, 1990), Engeland (Borg, 1990) en de Verenigde Staten (Hock, 1988) maakt aannemelijk dat het directe contact met leerlingen ook daar één van de belangrijkste oorzaken is van gevoelens van belasting van docenten. Een voorbeeld hiervan is het gevoel van 'burnout', dat ondermeer kan resulteren in het verlaten van het beroep. Vergelijkbare resultaten zijn gevonden bij aanstaande docenten (zie bijvoorbeeld MacDonald, 1993).

Wanneer we aanstaande docenten meer adequate strategieën kunnen helpen ontwikkelen, kunnen gevoelens van belasting verminderen en zou het contact met leerlingen vooral als een bron van arbeidsvreugde kunnen worden ervaren. Om aanstaande docenten in de lerarenopleiding te kunnen helpen adequate strategieën te ontwikkelen, is beschrijvende en evaluerende informatie over gedrag van aanstaande docenten in aandachtseisende situaties in de klas essentieel. In deze

¹ Correspondentie: Wilfried Admiraal, IVLOS Universiteit Utrecht, Postbus 80127, 3508 TC Utrecht. Tel: 030-2534414; fax: 030-2532741; e-mail: W.Admiraal@ivlos.ruu.nl

bijdrage verschaffen we dergelijke beschrijvende informatie. Het maakt onderdeel uit van een uitgebreider onderzoeksproject naar de aard van de aandachtseisende situaties, de wijze waarop aanstaande docenten hierop reageren en gevolgen van die reacties (zie Admiraal, 1994).

ONDERZOEK NAAR REACTIES VAN DOCENTEN OP AANDACHTSEISENDE SITUATIES

Reacties van docenten op aandachtseisende situaties zijn veelvuldig onderzocht door het coping-gedrag van docenten binnen en buiten het klaslokaal in kaart te brengen. In deze paragraaf zullen we enkele voorbeelden kort bespreken. Eerst is een definitie op haar plaats. Coping-gedrag kan worden 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). Lazarus en Folkman benadrukken dat veel coping-gedrag beide functies vervult, afhankelijk van de context waarin het gedrag zich voordoet.

In onderzoek naar coping-gedrag van docenten zijn ook meer specifieke vormen van coping beschreven. Freeman (1987) heeft bijvoorbeeld bij docenten het coping-gedrag in situaties in de klas bestudeerd, alsmede factoren die gevoelens van stress verminderen of voorkomen. De laatstgenoemde factoren betreffen sociale steun en aspecten van de schoolorganisatie. Het coping-gedrag is door Freeman gemeten met een vragenlijst met 56 items, die vijf clusters van coping-gedrag vormen: 'rationeel probleemoplossen', 'verdringen' (bijvoorbeeld aan andere zaken denken), 'emotionele steun' (bijvoorbeeld het uiten van gevoelens), 'stoom afblazen en je op je werk storten', en 'positief denken en zoeken van afleiding'. Deze benamingen geven aan dat in deze vragenlijst veel items zijn opgenomen die verwijzen naar coping-gedrag *buiten* het klaslokaal.

Onderzoek naar coping-gedrag van docenten *in* het klaslokaal baseert zich dikwijls op andere theoretische ingangen dan gangbaar in psychologisch onderzoek naar stress en coping. Veel studies zijn uitgevoerd vanuit het perspectief van informatieverwerking tijdens het lesgeven (zie Clark & Peterson, 1986) of klasmanagement en -organisatie (zie Doyle, 1986). In dergelijk onderzoek is informatie over verbaal en nonverbaal gedrag van docenten dikwijls verzameld aan de hand van observaties van lessen, soms aangevuld met interviews volgens de 'stimulated recall'-methode. Evertson (1987) bijvoorbeeld, heeft clusters van management-gedrag van docenten onderscheiden, zoals 'management van de lesstof', 'regels en procedures', 'tegemoet komen aan de behoeften van leerlingen', en 'management van leerlinggedrag'. In het onderzoek van Evertson is, net zoals in veel ander onderzoek dat in het perspectief van informatieverwerking heeft plaatsgevonden, een nadruk gelegd op de betekenis van instructie in het docentgedrag. Créton en Wubbels (1984) hebben echter niet alleen een dergelijk methodisch-didactisch perspectief op gedrag onderscheiden, maar ook een interpersoonlijk perspectief, analoog aan het onderscheid door Watzlawick, Beavin en Jackson (1967) van het inhouds- en betrekkingsniveau van communicatie. In elke communicatie wordt zowel inhoudelijke informatie gerapporteerd (het inhoudsniveau) als een relatie tot uitdrukking gebracht (het betrekkingsniveau). Dit betekent dat gedrag van docenten in de klas en de interpretatie hiervan door leerlingen vanuit beide perspectieven te beschrijven is.

Den Hertog (1990) heeft in zijn proefschrift over reacties van docenten op problematische gebeurtenissen het coping- en het interpersoonlijke perspectief van gedrag gecombineerd. Hij heeft vijf typen gedragsintenties onderscheiden die respectievelijk verwijzen naar dominant-onvriendelijk, dominant-vriendelijk, meegaand, organiserend en vermijgend gedrag. Voor zover ons bekend zijn er geen studies uitgevoerd waarin het coping-, het methodisch-didactische en het interpersoonlijke perspectief van gedrag gezamenlijk zijn gebruikt om gedrag van docen-

ten in aandachtseisende situaties in de klas te beschrijven. Studies vanuit slechts één of twee perspectieven leveren naar onze mening beperkte informatie op over het gedrag van (aanstaande) docenten in aandachtseisende situaties in de klas.

THEORETISCH KADER VAN ONS ONDERZOEK

Een meer omvattende benadering beschrijft gedrag van aanstaande docenten in aandachtseisende klassesituaties vanuit zowel het coping-perspectief als het methodisch-didactische en het interpersoonlijke perspectief van docentgedrag. In deze paragraaf presenteren wij ons theoretisch kader, waarmee wij docentgedrag bestuderen vanuit deze drie perspectieven. Onder het coping-perspectief wordt de betekenis verstaan die gezien de aard van de aandachtseisende situaties aan gedrag kan worden toegekend. In het methodisch-didactische perspectief wordt gedrag beschreven in termen van lesorganisatie en didactisch handelen van docenten. Het interpersoonlijke perspectief van docentgedrag verwijst naar de relationele boodschap die door een docent in gedrag ten opzichte van leerlingen tot uiting wordt gebracht.

Coping perspectief van docentgedrag

Hoewel onderzoek naar coping-gedrag van docenten enkele aanwijzingen heeft gegeven over de vraag welk gedrag bij aanstaande docenten in aandachtseisende situaties in de klas kan worden waargenomen, kan literatuur over coping-gedrag in andere beroepsgroepen aanvullende informatie geven. Farrell (1983) heeft vier categorieën onderscheiden van gedragsopties van werknemers naar aanleiding van een als negatief ervaren werksituatie. Deze categorieën zijn gebaseerd op de taxonomie van Hirschman van reacties op de neergang in bedrijven, organisaties en staten (Hirschman, 1970). 'Vermijden' ('exit') is gelijk aan het vrijwillig verlaten van de betrekking, terwijl 'ingrijpen' ('voice') is gedefinieerd als elke poging tot veranderen in plaats van wegvlugten van de ongewenste situatie, zoals bijvoorbeeld praten met de chef over verbetering van de werkomstandigheden. 'Negeren' ('neglect') beschrijft laks en onachtzaam gedrag bij de werknemers, zoals zich ziek melden en voorbijgaan aan wat er aan de hand is. 'Afwachten' ('loyalty') tenslotte, is gedefinieerd als doorgaan met het werk alvorens te reageren op de problemen. Deze vier categorieën kunnen worden geordend aan de hand van twee (onafhankelijke) dimensies: 'opbouwend' versus 'afbrekend' en 'actief' versus 'passief'. Ingrijpen (actief) en afwachten (passief) kunnen worden beschouwd als opbouwende reacties, terwijl vermijden (actief) en negeren (passief) als afbrekend kunnen worden gezien. Hoewel deze vier categorieën van gedrag niet verwijzen naar gedrag van aanstaande docenten in de klas, zijn wij van mening dat dit model ons wel kan helpen bij het in kaart te brengen van dergelijk gedrag.

Methodisch-didactisch perspectief van docentgedrag

In een beschrijving van gedrag van aanstaande docenten in aandachtseisende situaties in de klas kan het methodisch-didactische perspectief van docentgedrag niet worden overgeslagen. Op basis van het 'Classroom Observation Instrument' (Gage, 1978), observatie van video-opnamen van lessen uit eerder onderzoek en vier case-studies heeft Westerhof (1989) een codeerschema ontworpen voor het observeren en coderen van docentgedrag in de klas. In een vervolgonderzoek bij ongeveer 100 docenten heeft hij drie hoofdcategorieën onderscheiden: 'leertaakgebonden instructie-gedrag', 'leertaakgebonden management-gedrag', en 'niet-leertaakgebonden gedrag'. Binnen deze drie categorieën worden acht subcategorieën onderscheiden: 'overdracht', 'uitnodigen', 'controleren', 'kritiseren', 'luisteren', 'arrangeren', 'reguleren' en 'niet-leertaakgebonden gedrag'. Op basis van 'time sampling' (eenheden van vijf seconden) heeft Westerhof reeksen van gedragingen van ervaren en onervaren docenten geanalyseerd en in verband gebracht met leeropbrengst bij leerlingen. Hoewel andere observatie-instrumenten zijn ontwikkeld (zie voor een overzicht Good & Brophy, 1991), is het instrument van Westerhof geschikt voor ons onderzoek, omdat vanuit het methodisch-didactische perspectief een breed bereik van gedragingen in kaart kan worden gebracht.

Interpersoonlijk perspectief van docentgedrag

Het onderscheid tussen het inhouds- (methodisch-didactisch) en het betrekkningsniveau (interpersoonlijk) van gedrag is ondermeer relevant voor een analyse van de relaties tussen docenten en leerlingen. Dit geldt met name voor aanstaande en beginnende docenten, die dikwijls worden geconfronteerd met leerlingen die hen 'uitproberen' (Brooks, 1985). Créton en Wubbels (1984) hebben een model, ontwikkeld door Leary (1957), bewerkt om docentgedrag vanuit het interpersoonlijke perspectief te beschrijven en te meten. Zij hebben acht categorieën onderscheiden, die worden geordend door twee onafhankelijke dimensies, de machts- (boven-onder) en de nabijheidsdimensie (samen-tegen). De eerste dimensie verwijst naar de mate waarin docenten de activiteiten van leerlingen bepalen (boven) of waarin de leerlingen zelf hun activiteiten bepalen (onder); de tweede dimensie verwijst naar de mate waarin docenten leerlingen en hun gedrag afwijzen (tegen) dan wel dit positief waarderen (samen). In elk van de acht categorieën van interpersoonlijk gedrag zijn beide dimensies min of meer vertegenwoordigd: 'leidend' (boven en enigszins samen), 'helpend/vriendelijk' (samen en enigszins boven), 'begrijpend' (samen en enigszins onder), 'ruimtegevend/ruimtelandend' (onder en enigszins samen), 'onzeker' (onder en enigszins tegen), 'onvriendelijk' (tegen en enigszins onder), 'corrigerend' (tegen en enigszins boven) en 'streng' (boven en enigszins tegen).

Dit model is in het verleden niet alleen gebruikt bij observaties van lessen, maar ook om percepties van docentgedrag van leerlingen en docenten in kaart te brengen. Hiertoe is de Vragenlijst Interpersoonlijk Leraarsgedrag (VIL) ontwikkeld en afgenomen bij docenten en leerlingen in Nederland, de Verenigde Staten en Australië (zie Wubbels, Créton, Levy & Hooy-mayers, 1993).

Doelstelling van het onderzoek

Ofschoon resultaten van onderzoek naar docentgedrag ons enig inzicht hebben gegeven in reacties van aanstaande docenten op aandachtseisende situaties, is een karakterisering van dergelijk gedrag onvolledig gebleven. In deze bijdrage beperken wij ons tot waarneembaar gedrag en beschrijven we activiteiten van aanstaande docenten in aandachtseisende situaties in de klas vanuit drie perspectieven van gedrag: het coping-, het methodisch-didactische en het interpersoonlijke perspectief. Het doel van hier beschreven onderzoek is een beschrijvend model van gedrag van aanstaande docenten te ontwikkelen met behulp van een beperkt aantal dimensies, waarbij de drie perspectieven worden gecombineerd. In toekomstig onderzoek zou dit model als uitgangspunt kunnen dienen voor een model voor reacties op aandachtseisende situaties in de klas van meer ervaren docenten. Tevens kan een dergelijk model worden gebruikt om begeleiders van aanstaande docenten te helpen met het signaleren van blinde vlekken van aanstaande docenten voor communicatie in de klas gedurende aandachtseisende situaties. We hebben ons beperkt tot *aanstaande* docenten, omdat we ook aanwijzingen willen verzamelen voor verbetering van de lerarenopleiding.

We hebben drie onderzoeksvragen geformuleerd:

- 1 Wat doen aanstaande docenten in aandachtseisende situaties in de klas, gezien vanuit het coping-, het methodisch-didactische en het interpersoonlijke perspectief van docentgedrag?
- 2 Wat zijn de overeenkomsten in gedrag van aanstaande docenten in verschillende typen aandachtseisende situaties?
- 3 Op welke wijze kunnen de scores op de verschillende perspectieven van docentgedrag in alle typen aandachtseisende situaties worden gereduceerd tot een beperkt aantal onderliggende dimensies van gedrag?

METHODE

Dataverzameling en codeereenheden

Het onderzoek is uitgevoerd bij 27 docenten-in-opleiding (DIO's) van de 60 DIO's die in 1990

en 1991 de universitaire lerarenopleiding van de Universiteit Utrecht volgden. Voor het selecteren van de respondenten is een gestratificeerde steekproeftrekking gehanteerd. De populatie van aanstaande docenten is eerst verdeeld in negen strata die verwijzen naar een combinatie van vakcluster (moderne vreemde talen, natuurwetenschappelijke vakken en maatschappijvakken) en stageperiode (voorjaar 1990, en najaar 1990 en 1991). Vervolgens zijn uit elk stratum aselect drie studenten in het onderzoek betrokken. Onze steekproef bleek niet te verschillen (bij $\alpha=0,05$) van de andere docenten-in-opleiding wat betreft sekse, leeftijd en vakgebied. De Utrechtse universitaire lerarenopleiding is een eenjarige postdoctorale opleiding, waarin naast andere curriculumonderdelen twee stages worden doorlopen: de triostage en de individuele eindstage. 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. Tijdens de individuele eindstage in de tweede helft van de opleiding neemt één DIO gedurende drie maanden alle lessen van een paar klassen over van een ervaren docent. In deze stage is de DIO verantwoordelijk voor de lessen, inclusief bijvoorbeeld de rapportcijfers en contacten met de schoolleiding. Zeven DIO's zijn tijdens hun triostage en twintig andere respondenten zijn tijdens hun individuele eindstage in het onderzoek betrokken. Koetsier (1991) geeft een uitgebreide beschrijving van de Utrechtse universitaire lerarenopleiding, in het bijzonder van de individuele eindstage.

In ons onderzoek zijn de gegevens op twee manieren verzameld: door middel van een video-opname van een les van de respondent en door middel van een interview volgens de 'stimulated recall'-methode. Van elke respondent is één les op video opgenomen. In de interviews hebben we de respondenten gevraagd ongeveer tien situaties uit de voorgaande les te omschrijven die tijdens de les hun aandacht hadden geëist. Naast deze uitspraken zijn in de interviews ook uitspraken ontlokt over de inschatting van aandachtseisende situaties door de respondent. Tijdens het interview is de respondent gevraagd de video stil te zetten, wanneer zij iets wilden melden over wat zij tijdens de les voelden en dachten. De uitspraken over de inschattingen zullen niet in deze bijdrage worden besproken.

In ons onderzoek verwijst gedrag van een aanstaande docent in een aandachtseisende situatie naar 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 zelf bepaald. Om het gedrag in de reacties in kaart te brengen is gebruik gemaakt van 'event sampling'. Elke eenheid heeft een lengte van één seconde tot twee minuten en vertegenwoordigt een voorbeeld van een activiteit van aanstaande docenten in aandachtseisende situaties. We hebben vier scores toegekend aan elke eenheid, waarvan er één verwijst naar het coping-perspectief, één naar het methodisch-didactische perspectief, en twee naar het interpersoonlijke perspectief van docentgedrag. Elke reactie op een aandachtseisende situatie bestaat derhalve uit één of meer sets van vier scores voor gedrag van aanstaande docenten. Aan de omschrijving van een respondent van een aandachtseisende situatie is één score toegekend. In totaal zijn 306 klassesituaties en hetzelfde aantal reacties geanalyseerd. Over het algemeen hanteert de respondent in de geobserveerde les als didactische werkvorm klassikale instructie of onderwijsleergesprek. Practicumlessen zijn voor dit onderzoek niet geobserveerd.

Vorm en inhoud van de operationalisaties

Aard van de aandachtseisende situatie

We hebben de aard van de classesituatie gedefinieerd in termen van de door de respondent waargenomen kenmerken van de aandachtseisende situatie in de klas. De situaties zijn gecodeerd aan de hand van negen categorieën (zie Tabel 1). De indeling in negen categorieën is gebaseerd op onderzoek van Peters (1985) naar problematische situaties van docenten. Op basis van een survey-onderzoek bij meer dan 200 docenten heeft hij 2.700 beschrijvingen van problematische situaties ingedeeld in 13 categorieën. Na clusteranalyse heeft hij drie categorieën van situaties in de klas onderscheiden: 'functioneren van de docent', 'ongewenst gedrag van leerlingen' en

Tabel 1. Categorieën van het instrument waarmee de aard van de aandachtseisende situaties wordt gemeten, elk geïllustreerd met een voorbeeld. In de rechter kolom is de proportie van voorkomen weergegeven. In totaal zijn 306 klassesituaties gecodeerd.

categorie	voorbeeld	percentage
1. lesstofoverdracht	docent Engels vergist zich in de uitleg van het gebruik van tijden	8
2. lesorganisatie	docent is bezig met het uitdelen van practicumspullen, stencils of opdrachten	9
3. onderwijsmiddelen	docent wil de video gebruiken, maar die werkt niet	4
4. niet-taakgericht gedrag van de klas	een klas is rumoerig aan het begin van het lesuur	22
5. niet-taakgericht gedrag van één leerling of van een groepje leerlingen	twee leerlingen zitten voortdurend met elkaar te praten	21
6. kritiek	een aantal leerlingen uit kritiek op de normering van een proefwerk	5
7. matige prestaties	een leerling spreekt een Frans woord niet goed uit	4
8. taakgericht gedrag	een leerling vraagt de docent uitleg te geven over een opdracht	18
9. weinig inzet	leerlingen antwoorden niet, als de docent twee leerlingen vraagt voor een dialoog	9

'hulpmiddelen en materialen'. Wij hebben de eerste twee categorieën verder opgedeeld, omdat deze naar onze mening teveel verschillende typen situaties omvatten. Hiertoe hebben wij de situatiebeschrijvingen van docenten uit het onderzoek van Wulk (1988) gebruikt en resultaten van onze voorstudie waarin wij een aantal lessen van aanstaande docenten hebben geobserveerd en hen hebben geïnterviewd over situaties die zij tijdens de les als aandachtseisend hadden ervaren. De eerste twee categorieën van ons instrument verwijzen naar situaties over het functioneren van de docent; categorie 4 tot 9 naar taakgericht en niet-taakgericht gedrag van leerlingen. Situaties die betrekking hebben op de omgeving van de klas, zoals krijgt dat ontbreekt of een video die niet functioneert, zijn ondergebracht onder categorie 3: 'hulpmiddelen en materialen'. De categorieën waarmee de aard van de aandachtseisende situatie wordt gemeten en de proportie van voorkomen van elke categorie in onze gegevens zijn weergegeven in Tabel 1.

Gedrag van aanstaande docenten

In ons onderzoek beperken we gedrag van aanstaande docenten tot de waargenomen activiteiten van docenten tijdens of direct volgend op een aandachtseisende situatie in de klas. We karakteriseren gedrag van aanstaande docenten vanuit drie perspectieven.

Ten eerste meten we gedrag vanuit het coping-perspectief, waarbij de score verwijst naar de betekenis die, gezien de aard van de situatie, aan gedrag kan worden toegekend. Wij hebben vier categorieën onderscheiden, gebaseerd op de bovenbeschreven studie van Farrell (1983) naar waargenomen gedragsopties van werknemers naar aanleiding van als negatief ervaren situaties op het werk. We hebben zijn vier categorieën bewerkt voor de klascontext en deze als volgt omschreven. De categorie 'ingrijpen' (I) houdt in dat een aanstaande docent een poging onderneemt de aandachtseisende situatie te veranderen. De categorie 'afwachten' (A) is van toepassing als de aanstaande docent een situatie afwijst, maar geen waarneembaar gedrag vertoont

waaruit kan worden opgemaakt dat hij gericht is op verandering van een situatie. De categorie 'vermijden' (V) is van toepassing als een aanstaande docent gedrag vertoont dat gericht is op verandering van de (klas-)omgeving, maar waarbij niet (direct) wordt ingegaan op de betreffende aandachtseisende situatie. De categorie 'negeren' (N) tenslotte is van toepassing als een aanstaande docent gedrag voortzet dat voorafgaand aan de aandachtseisende situatie is vertoond.

Ten tweede hebben we gedrag van aanstaande docenten gemeten in termen van lesorganisatie en didactisch handelen (methodisch-didactisch perspectief). Ons codeerschema is gebaseerd op het onderzoek van Westerhof (1989) naar gedrag in de klas van docenten in het basisonderwijs. We hebben zijn codeerschema met acht categorieën grotendeels overgenomen. Tevens hebben we zeven andere categorieën toegevoegd om meer gedetailleerde informatie over docentgedrag te verzamelen ('doorvragen', 'belonen', 'samenvatten en herhalen', 'zwijgen', 'discipline', 'helpen', en 'geen interactie'). We hebben ons hierbij gebaseerd op beschrijvingen van gedrag in de vier case-studies van Westerhof (1989) en onze observaties in ons vooronderzoek.

Ten derde hebben we gedrag gemeten vanuit het interpersoonlijke perspectief middels twee

Tabel 2. Categorieën en schalen van de instrumenten waarmee het docentgedrag wordt gemeten, elk geïllustreerd met een voorbeeld. In de rechter kolom zijn enkele beschrijvende kenmerken weergegeven (s.a. = standaardafwijking).

categorie		voorbeeld	N=2300	
Coping-perspectief			percentage	
I	ingrijpen	docent loopt langs de banken en maant iedere leerling tot rust	31	
A	afwachten	docent blijft voor de klas staan en kijkt ontevreden de klas in	21	
V	vermijden	docent maakt met een leerling een praatje over een voetbalwedstrijd	17	
N	negeren	docent gaat door met het overhoren van het huiswerk	31	
Methodisch-didactisch perspectief			percentage	
o	overdracht	docent legt een som uit aan de klas	11	
u	uitnodigen	docent vraagt een leerling te beginnen met het lezen van een tekst	8	
d	doorvragen	docent vraagt een leerling wat hij met een opmerking bedoelt	5	
c	controleren	docent informeert of iedereen twee stencils voor zich heeft	2	
k	kritiseren	docent zegt dat hij een antwoord van een leerling niet goed vindt	3	
b	belonen	docent zegt dat hij een antwoord van een leerling prima vindt	3	
s	samenvatten	docent vat een lang antwoord van een leerling samen	4	
a	arrangeren	docent pakt de overhead en zoekt naar een aantal transparanten	9	
l	luisteren	docent luistert naar een opmerking van een leerling	7	
z	zwijgen	docent staat zwijgend voor de klas en kijkt de klas rond	14	
r	reguleren	docent spoort de klas aan vooral aantekeningen te maken	11	
p	discipline	docent zet een leerling op een andere plaats	14	
h	helpen	docent helpt een leerling met een proef	2	
n	niet-inhoudgebonden	docent staat bij de deur te praten met binnenkomende leerlingen	6	
g	geen interactie	docent kijkt tijdens klassikale opdracht proefwerken na	1	
Interpersoonlijk perspectief			gem.	s.a.
m(1-5) machtsdimensie				
m1	onder	docent schrijft op het bord en de klas is rumoerig	2.7	1.1
m5	boven	docent legt klassikaal lesstof uit en de leerlingen luisteren		
n(1-5) nabijheidsdimensie				
n1	tegen	docent stuurt een leerling de klas uit	2.8	0.8
n5	samen	docent beloont enthousiast een antwoord van een leerling		

beoordelingsschalen die staan voor de machts- (boven-onder) en de nabijheidsdimensie (samen-tegen, zie Créton & Wubbels, 1984). De machtsdimensie verwijst naar de mate waarin uit het gedrag van aanstaande docenten kan worden afgeleid dat zij de activiteiten van leerlingen bepalen of dat de leerlingen zelf hun activiteiten bepalen. Aan het eerstgenoemde gedrag is een score 'boven' of 'zeer boven' toegekend, aan het laatstgenoemde gedrag een score 'onder' of 'zeer onder'. De nabijheidsdimensie verwijst naar de mate waarin uit het gedrag kan worden afgeleid dat een aanstaande docent leerlingen en hun gedrag afwijkt dan wel dit positief waardeert. In het eerste geval is een score 'tegen' of 'zeer tegen' toegekend, in het laatste geval is 'samen' of 'zeer samen' gescoord.

In Tabel 2 zijn de categorieën en/of beoordelingsschalen van onze instrumenten weergegeven, alsmede enkele beschrijvende kenmerken voor onze gegevens.

Kwaliteit van de operationalisaties

Betrouwbaarheid

Om de betrouwbaarheid te bepalen van het instrument waarmee de aard van de aandachtseisende situaties is gemeten, zijn de beschrijvingen van alle 306 classesituaties door twee codeurs gescoord. De interbeoordelaarsbetrouwbaarheid (Cohens κ) bedraagt 0,88.

Om de betrouwbaarheid van de instrumenten te bepalen waarmee het gedrag van aanstaande docenten is gemeten, zijn ongeveer 900 activiteiten of codeereenheden (voor zowel het coping- als het methodisch-didactische perspectief), en 120 activiteiten (voor het interpersoonlijke perspectief) gecodeerd door twee observatoren. Het aantal door twee observatoren gecodeerde activiteiten is afhankelijk van het aantal categorieën van het observatiesysteem: hoe meer categorieën, des te meer activiteiten zijn gecodeerd. De interbeoordelaarsbetrouwbaarheid voor het coping- en het methodisch-didactische perspectief (Cohens κ) bedraagt respectievelijk 0,90 en 0,93. De correlatie ('pooled within-groups-correlation coefficient') tussen beide observatoren ten aanzien van de scores op de machts- en de nabijheidsdimensie is respectievelijk 0,79 en 0,80.

Onze conclusie is dat de interbeoordelaarsbetrouwbaarheid voor het instrument waarmee we de aard van de aandachtseisende situaties meten en de instrumenten waarmee het gedrag van aanstaande docenten wordt gemeten bevredigend is (rond de norm van 0,80 of hierboven, zie Popping, 1988).

Validiteit

De uitspraken die een respondent heeft gedaan over de aard van de aandachtseisende situatie zijn door de codeur ook samengevat in één zin. Wat betreft de aard van de aandachtseisende situatie hebben we een aspect van de gelijktijdige validiteit gemeten door voor elk situatie de scores te vergelijken met de scores die zijn gebaseerd op de samenvattende zin van de uitspraken van een respondent. Beide manieren van meten geven vergelijkbare resultaten (Cohens κ bedraagt 0,87). Dit maakt aannemelijk dat de scores die verwijzen naar de aard van de aandachtseisende situaties ongevoelig zijn voor een dergelijk verschil in meetprocedure.

Wat betreft het gedrag van docenten hebben we de begripsvaliditeit van de metingen in kaart gebracht middels correlaties tussen categorieën en schalen van de instrumenten. We hebben correspondentie-analyse gebruikt om na te gaan of onze vier categorieën van het coping-perspectief zich tot elkaar verhouden zoals op basis van het model van Farell kan worden verwacht. In dit model worden de vier categorieën 'ingrijpen', 'afwachten', 'vermijden' en 'negeren' geordend aan de hand van twee onafhankelijke dimensies: 'opbouwend-afbrekend' en 'actief-passief'. De resultaten van de correspondentie-analyse geven aan dat dit het geval is voor zes van de negen typen van aandachtseisende situaties; in de andere drie typen wijkt één categorie enigszins af van de positie die op basis van het model verwacht kan worden (zie Admiraal, 1994).

Wat betreft het methodisch-didactische perspectief hebben we correspondentie-analyse toegepast om te achterhalen of elke categorie een proportie variantie verklaart die tenminste gelijk

is aan de norm van 1, gedeeld door het aantal categorieën dat in de analyses is betrokken (zie Jambu, 1991). Wij beschouwen categorieën die aan deze norm voldoen als noodzakelijk om het methodisch-didactische perspectief van gedrag te karakteriseren. Op de categorie 'niet-inhoudgebonden' na voldoen alle categorieën aan deze norm. De categorie 'niet-inhoudgebonden' is niet opgenomen in verdere analyses.

Met betrekking tot het interpersoonlijke aspect van gedrag is nagegaan of onze gegevens over de samenhang tussen de machts- en de nabijheidsdimensie het model van Créton en Wubbels (1984) ondersteunen, waarin deze dimensies als onafhankelijk worden beschouwd. Er blijkt geen significante ($\alpha=0,05$) samenhang te bestaan tussen de scores op de machts- en de nabijheidsdimensie. De 'pooled within-groups'-correlatiecoëfficiënt is $-0,12$.

Voor zowel het methodisch-didactische als het interpersoonlijke perspectief is de gelijktijdige validiteit nagegaan door onze resultaten te vergelijken met de resultaten van ander onderzoek. We hebben de frequenties van categorieën van het methodisch-didactische perspectief vergeleken met de categorieën van het oorspronkelijke instrument van Westerhof (1989). Resultaten van een toets op verschil in proporties geven aan dat de frequenties van de categorie 'discipline' van beide instrumenten verschillen: in ons onderzoek hanteren aanstaande docenten meer ordemaatregelen ($t=7,5$; $p<.001$). Wat betreft het interpersoonlijke perspectief van docentgedrag hebben de frequenties van de schaalpunten uit ons onderzoek vergeleken met die uit het onderzoek van Van Tartwijk (1993), die hetzelfde instrument gebruikt in zijn onderzoek naar interpersoonlijk leraarsgedrag. Het blijkt dat de scores op de nabijheidsdimensie, in vergelijking met het onderzoek van Van Tartwijk, zijn verschoven van het schaalgemiddelde naar de schaalpunten die verwijzen naar 'tegen'-gedrag. Een verklaring voor deze verschillen zou kunnen zijn dat we gedrag van aanstaande docenten in aandachtende situaties meten, waardoor dit gedrag dikwijls bestaat uit ordemaatregelen. Het ligt voor de hand dat dergelijk gedrag als 'tegen'-gedrag wordt gezien. Al met al stemmen de resultaten met betrekking tot de gelijktijdige validiteit ons tot tevredenheid.

Analyses

We zullen de eerste en tweede onderzoeksvraag beantwoorden op basis van frequenties van categorieën en schaalpunten van de instrumenten waarmee wij docentgedrag meten, alsmede combinaties van deze categorieën in elk type aandachtende situaties. Om onderliggende dimensies te extraheren (de derde onderzoeksvraag) hebben we principale componenten-analyses (PRINCALS, Gifi, 1983) toegepast. We hebben twee dimensies geëxtraheerd op basis van een kruistabel van categorieën en schaalpunten van de instrumenten waarmee het docentgedrag wordt gemeten. Deze twee dimensies verklaren ongeveer 50% van de variantie in scores in alle negen typen van aandachtende situaties. Dit percentage voldoet aan de norm die is gerapporteerd door Gifi (1983) van 38% verklaarde variantie bij twee dimensies.

RESULTATEN

Gezien het geringe aantal reacties in enkele typen aandachtende situaties zullen we het gedrag van aanstaande docenten beschrijven in vier typen classesituaties: situaties die te maken hebben met de lesstofoverdracht (type 1), situaties die betrekking hebben op de lesorganisatie (type 2), situaties waarin niet-taakgericht gedrag van leerlingen een belangrijke rol speelt (type 4), en situaties over taakgericht gedrag van leerlingen (type 8). De uitkomst van de PRINCALS-analyse voor het vijfde type situaties (niet-taakgericht van één leerling) is vrijwel gelijk aan die voor het vierde type situaties. De vijf typen aandachtende situaties vertegenwoordigen tezamen een verscheidenheid aan classesituaties (78% van alle situaties die door onze respondenten zijn genoemd). We hebben ervoor gekozen om het methodisch-didactische perspectief als uitgangspunt te nemen van onze beschrijvingen, omdat de betekenis van deze categorieën in vergelijking met die van de andere perspectieven, minder afhankelijk is van de context waarin het gedrag is vertoont (i.c. de aard van de classesituatie).

In het eerste deel van de beschrijving van de resultaten zullen we activiteiten beschrijven die aanstaande docenten ondernemen naar aanleiding van aandachtseisende situaties in de klas. Aan het eind van dit deel zullen wij overeenkomsten in gedrag in verschillende klassesituaties toelichten. In het tweede deel van de resultaten beschrijven we het gedrag van aanstaande docenten in termen van de twee in de principale componenten-analyse geëxtraheerde dimensies.

Activiteiten in termen van de drie perspectieven van docentgedrag

In Tabel 3 zijn de proporties weergegeven van de categorieën van zowel het coping- als het methodisch-didactische perspectief van gedrag. We hebben alleen de categorieën opgenomen met een proportie die hoger is dan vijf procent; categorieën met een lagere proportie van voorkomen zijn gecombineerd in de categorie 'rest'. Voor het interpersoonlijke perspectief hebben we de gemiddelde score op de machts- en de nabijheidsdimensie in de Tabel opgenomen.

Tabel 3. Proportie van voorkomen van elke categorie van het coping- en het methodisch-didactische perspectief op gedrag in de vier typen klassesituaties.

Categorieën met een proportie kleiner dan vijf procent in een type situaties zijn gecombineerd in één categorie 'rest'. Wat betreft het interpersoonlijke perspectief zijn de gemiddelde scores op de machts- en de nabijheidsdimensie weergegeven. In de onderste regel is het aantal codeereenheden (docentactiviteiten) aangegeven.

	typen klassesituaties			
	lesstofoverdracht	lesorganisatie	niet-taakgericht gedrag leerlingen	taakgericht gedrag van leerlingen
Coping-perspectief				
ingrijpen	12	25	27	38
afwachten	14	12	21	18
vermijden	25	26	18	20
negeren	49	37	34	24
	100	100	100	100
Methodisch-didactisch perspectief				
overdracht	24	6	5	21
uitnodigen	8	18	6	9
doorvragen	—	—	—	10
controleren	—	7	—	—
kritiseren	—	—	—	6
samenvatten	—	—	—	7
arrangeren	12	16	8	6
luisteren	8	9	—	11
zwijgen	11	14	22	6
reguleren	16	13	15	7
discipline	—	—	29	8
rest	21	18	15	10
	100	100	100	100
Interpersoonlijk perspectief				
macht	2.9	2.6	2.2	2.9
nabijheid	3.1	3.1	2.5	3.1
aantal codeereenheden	185	198	538	357

Om ook *combinaties* van categorieën en schaalpunten toe te lichten, gebruiken we de uitkomsten van de *principale componenten-analyse* voor elk type situaties. In figuur 1 tot en met 4 zijn de dimensiescores van de categorieën en de schaalpunten van elk perspectief grafisch weergegeven. Voor elk type situaties zijn alleen die categorieën in de figuren opgenomen die meer dan vijf keer voorkomen. Categorieën die vaak samen voorkomen, zijn dichtbij elkaar geplaatst in de figuur; categorieën die nauwelijks samen voorkomen, zijn ver van elkaar in de figuur weergegeven. De scores zijn zo gespiegeld en/of geroteerd ten opzichte van de twee dimensies dat de twee assen in alle vier figuren dezelfde dimensie vertegenwoordigen, alsmede per dimensie dezelfde oriëntatie hebben.

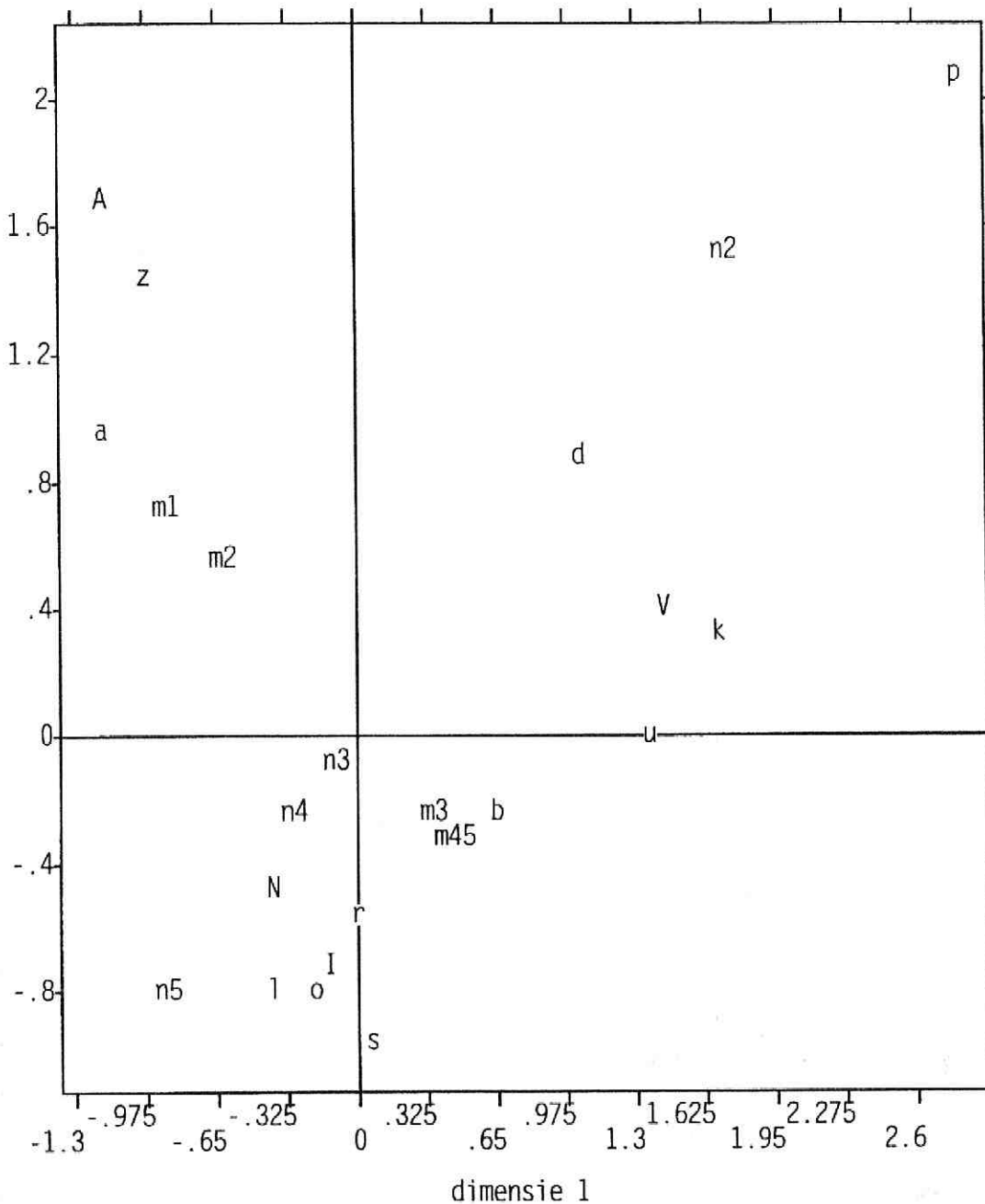


Fig. 1. Scores op de twee dimensies van categorieën en schaalpunten in situaties van het type 1. Voor de betekenis van letters en cijfers verwijzen we naar tabel 2. Wanneer aan schaalpunten van de machts- en de nabijheidsdimensie gelijke scores krijgen toegekend, zijn deze in de figuur gecombineerd (bijvoorbeeld 'm45').

Situatietype 1: lesstofoverdracht

Aandachteisende situaties in de klas van dit type handelen over problemen van aanstaande docenten met de lesstofoverdracht. Bijvoorbeeld, een aanstaande docent heeft een fout gemaakt in de uitleg van het gebruik van tijden in het Engels. Zoals we in Tabel 3 laten zien, is lesstofoverdracht het meest frequent in het gedrag van aanstaande docenten, gevolgd door 'reguleren' en 'arrangeren'. Deze drie categorieën vormen ongeveer 52% van het methodisch-didactische perspectief van het vertoonde gedrag. Lesstofoverdracht betreft zowel de juiste uitleg als meta-communicatie over de gemaakte vergissingen, zoals 'sorry, maar ik heb mij vergist in het gebruik van de verleden tijd'. 'Reguleren' heeft vooral betrekking op aanwijzingen van aanstaande docenten aan de leerlingen om in hun boek of op het bord te kijken, of om naar de juiste uitleg te luisteren. Figuur 1 geeft aan dat 'overdracht' ('o') en 'reguleren' ('r') dichtbij de categorieën 'ingrijpen' ('I') en 'negeren' ('N') zijn geplaatst. Dit betekent dat aanstaande docenten zowel lesstofoverdracht als organisatorische handelingen toepassen, soms om problemen met de lesstofoverdracht direct op te lossen, soms als een poging om door te gaan met de les. 'Arrangeren' van aanstaande docenten betreft vooral het opzoeken van de juiste uitleg in bijvoorbeeld een boek of eigen aantekeningen. Soms denken aanstaande docenten aan de juiste uitleg zonder hierbij te spreken ('zwijgen'). De positie van deze laatste twee categorieën in figuur 1 geeft aan dat aanstaande docenten een passieve houding aannemen ten aanzien van de problemen met de lesstofoverdracht ('A') en een 'onder'-positie innemen ten opzichte van de leerlingen ('m1' en 'm2'), wanneer zij in hun aantekeningen kijken of zwijgen. Combinaties van categorieën en schaalpunten anders dan hier beschreven, komen in veel mindere mate voor.

Situatietype 2: lesorganisatie

Klassesituaties van dit type betreffen problemen van aanstaande docenten met de organisatie van de les. Bijvoorbeeld, een aanstaande docent geeft de leerlingen de opdracht hun practicumspullen te pakken om in groepjes van twee aan een natuurkundepracticum te werken. Zoals aangegeven in Tabel 3 komt het uitnodigen van leerlingen om vragen te stellen en antwoorden te geven veelvuldig voor, gevolgd door de categorieën 'arrangeren', 'zwijgen' en 'reguleren'. Deze vier categorieën bevatten meer dan 60% van de scores van het methodisch-didactische perspectief van het vertoonde gedrag. Het uitnodigen van leerlingen betreft met name leerlingen de gelegenheid geven vragen te stellen over de lesstof, over de organisatie van de les, etcetera. In figuur 2 is aangegeven dat dergelijk gedrag van aanstaande docenten soms gebruikt wordt om problemen met de organisatie van de les op te lossen ('I'), soms om de problemen uit de weg te gaan ('V'). Regulerende activiteiten hebben vooral betrekking op het geven van aanwijzingen aan leerlingen in verband met de dagelijkse procedures in een les. Reguleren is in dit type situaties vooral gericht op het oplossen van de organisatieproblemen. 'Arrangeren' verwijst naar organisatorische handelingen van aanstaande docenten die soms op het oplossen en soms op het vermijden van de organisatorische problemen zijn gericht (zoals respectievelijk het uitdelen van materiaal of het uitvegen van het bord). De positie van 'arrangeren' en 'zwijgen' in figuur 2 houdt in dat aanstaande docenten een passieve houding vertonen ten aanzien van de lesorganisatie ('A') en een 'onder'-positie innemen ten opzichte van de leerlingen ('m1' en 'm2'), wanneer zij zwijgend de klas inkijken, het bord schoonvegen, experimenten voorbereiden of stencils uitdelen. Combinaties van categorieën en schaalpunten anders dan hier beschreven, komen in veel mindere mate voor.

Situatietype 4: niet-taakgericht gedrag van leerlingen

Situaties in de klas van dit type betreffen (ongewenst) niet-taakgericht gedrag van leerlingen en vormen meer dan 20% van het totale aantal aandachteisende situaties dat is genoemd (en meer dan 40%, wanneer situaties van het vijfde type die wat betreft aard van de situatie en de reactie vergelijkbaar zijn, worden meegeteld). Een voorbeeld is een klas waarin een aanstaande docent wil beginnen met de les, maar leerlingen met elkaar kletsen, ravotten of boeken naar elkaar gooien. Zoals in Tabel 3 is aangegeven, vormen de drie categorieën 'discipline', 'zwijgen' en 'reguleren' ongeveer 66% van het gedrag van aanstaande docenten vanuit het methodisch-

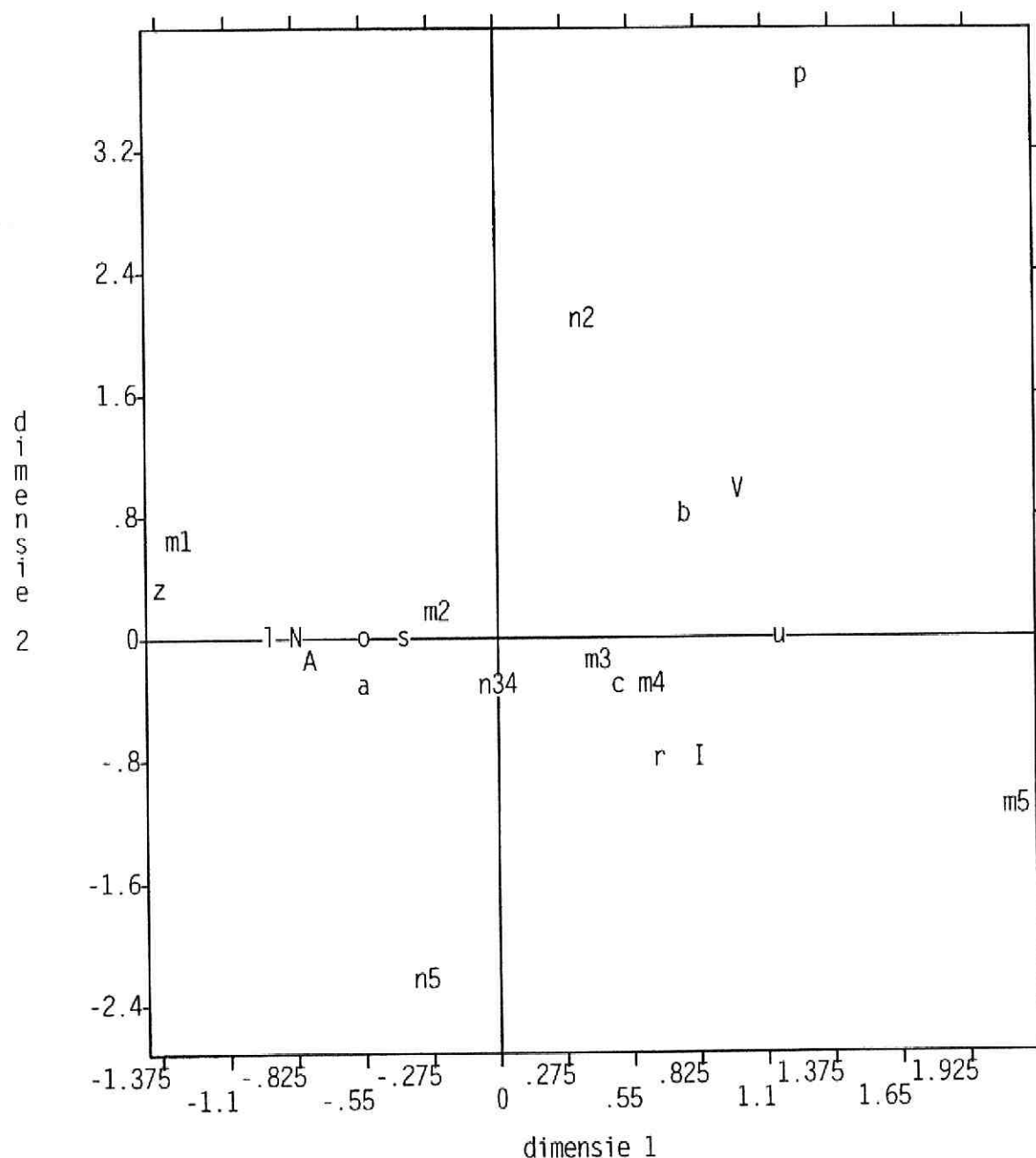


Fig. 2. Scores op de twee dimensies van categorieën en schaalpunten in situaties van het type 2. Voor de betekenis van letters en cijfers verwijzen we naar tabel 2. Wanneer aan schaalpunten van de machts- en de nabijheidsdimensie gelijke scores zijn toegekend, zijn deze in de figuur gecombineerd (bijvoorbeeld 'n34').

didactische perspectief. In deze situaties betreffen ordemaatregelen, die ongeveer 30% van het gedrag vanuit het methodisch-didactische perspectief in deze situaties vormen, het corrigeren van het niet-taakgerichte gedrag door leerlingen te vermanen, te waarschuwen of de klas uit te sturen. Zwijgen van aanstaande docenten houdt in dat ze de resultaten van hun ordemaatregelen afwachten of dat zij hun onderwijs onderbreken door af te wachten. De positie van de schaalpunten tegen ('n2') en zeer tegen ('n1') in figuur 3 geeft aan dat deze twee typen gedrag vanuit het interpersoonlijke perspectief gezien kunnen worden als het innemen van een afwijzende houding ten opzichte van leerlingen. Regulerende acties van aanstaande docenten houden vooral in dat de docent begint of wil beginnen met de les, en zijn met name gericht op het negeren

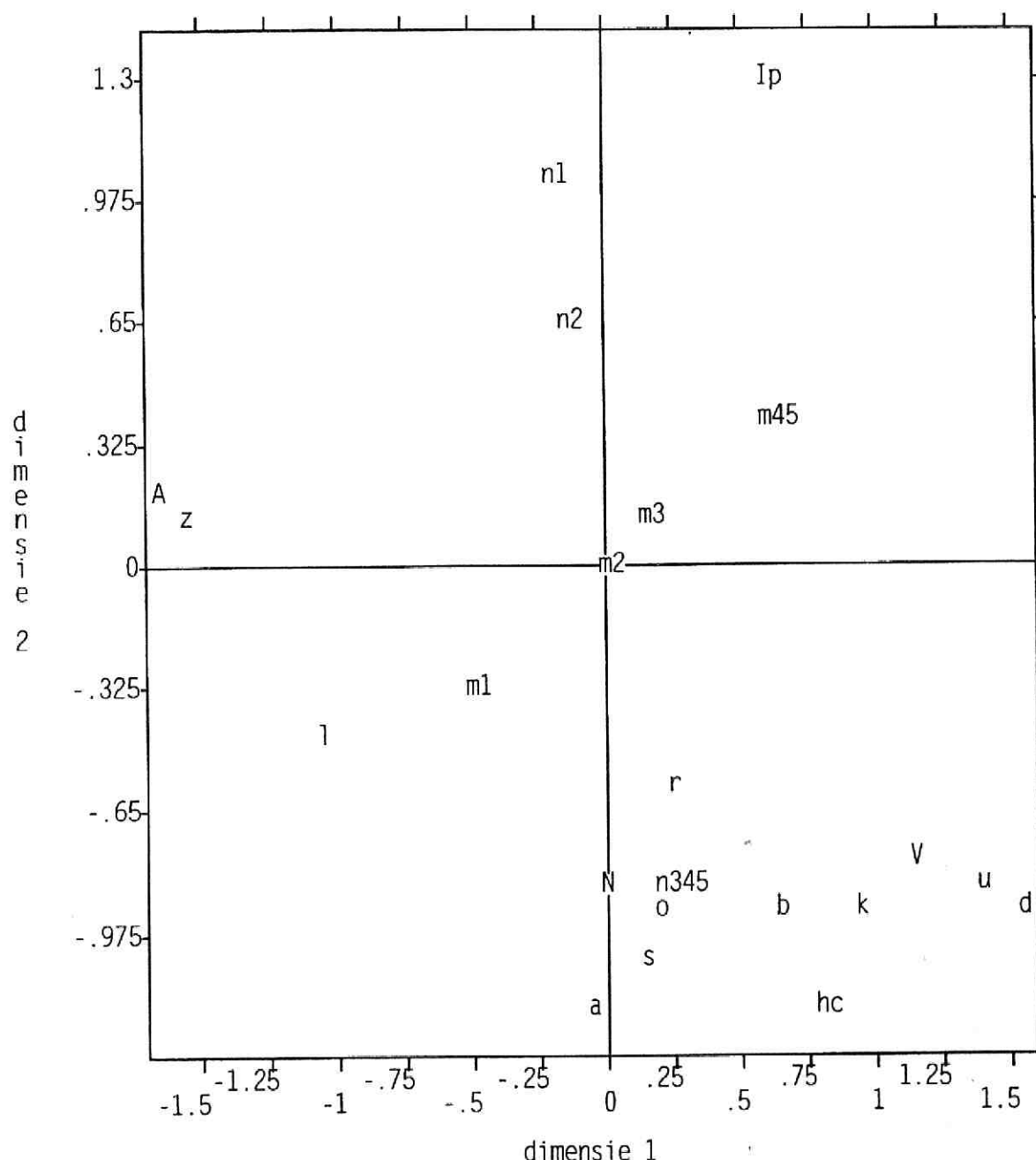


Fig. 3. Scores op de twee dimensies van categorieën en schaalpunten in situaties van het type 4. Voor de betekenis van letters en cijfers verwijzen we naar tabel 2. Wanneer aan schaalpunten van de machts- en de nabijheidsdimensie gelijke scores zijn toegekend, zijn deze in de figuur gecombineerd (bijvoorbeeld 'n345').

van het ongewenste gedrag van leerlingen ('N'). Andere combinaties van categorieën en schaalpunten dan die hier beschreven zijn, komen in veel mindere mate voor.

Situatietype 8: taakgericht gedrag van leerlingen

Klassesituaties van dit type betreffen taakgericht gedrag van leerlingen, zoals vragen en opmerkingen over de lesstof, en vormen 18% van alle aandachtseisende situaties die door de respondenten zijn genoemd. Een voorbeeld van een situatie is dat leerlingen vragen stellen over een scheikundesom. In Tabel 3 hebben we aangegeven dat in dit type situaties veel categorieën van het methodisch-didactische perspectief voorkomen. 'Lesstofoverdracht' maakt ongeveer 20%

uit van het methodisch-didactische gedrag, terwijl vrijwel alle andere categorieën met een proportie van vijf tot tien procent voorkomen. Lesstofoverdracht verwijst vooral naar het beantwoorden van vragen van leerlingen over de lesstof, maar is zo nu en dan ook gericht op het vermijden van (moeilijke of ongewenste) vragen van leerlingen. Zoals we in figuur 4 laten zien, zijn de meeste categorieën van het methodisch-didactische perspectief geplaatst in de buurt van 'ingrijpen' ('I'). Categorieën die hiervan veraf zijn geplaatst, zijn 'discipline' ('p'), 'luisteren' ('l') en 'zwijgen' ('z'). Met het gebruik van ordemaatregelen zijn aanstaande docenten niet gericht op het taakgerichte gedrag van leerlingen en scheppen een negatieve sfeer ten aanzien van de relatie met leerlingen. Wanneer aanstaande docenten niet spreken (zwijgen of luisteren),

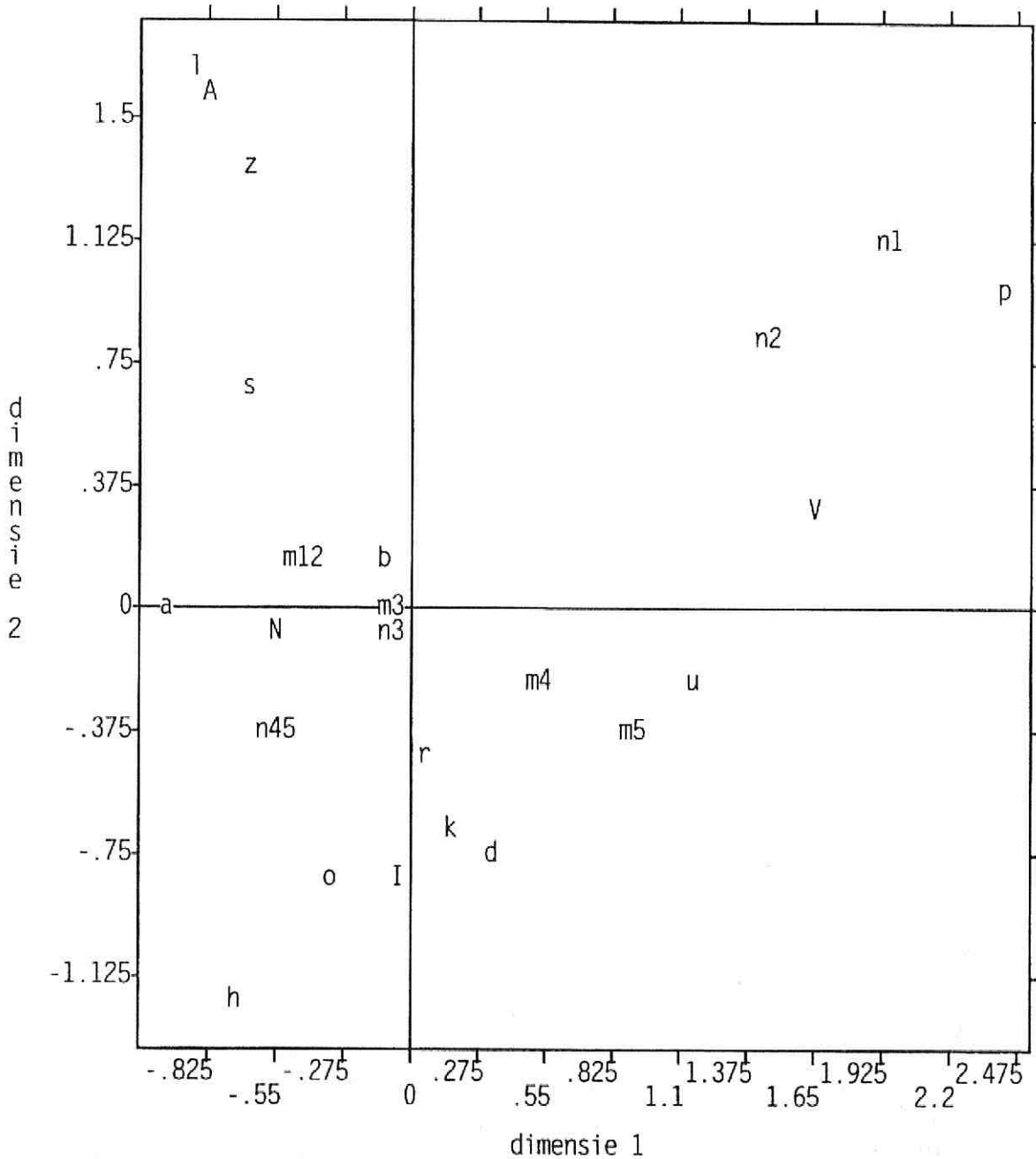


Fig. 4. Scores op de twee dimensies van categorieën en schaalpunten in situaties van het type 8. Voor de betekenis van letters en cijfers verwijzen we naar tabel 2. Wanneer aan schaalpunten van de machts- en de nabijheidsdimensie gelijke scores zijn toegekend, zijn deze in de figuur gecombineerd (bijvoorbeeld 'm12').

nemen zij een passieve houding aan, en een 'onder'- en 'tegen'-positie ten opzichte van het gedrag van leerlingen.

Het meest opmerkelijke resultaat van de vergelijking van de vier typen van aandachtseisende situaties in de klas heeft betrekking op de wijze waarop aanstaande docenten situaties aanpakken. Om problemen tijdens het onderwijs in de klas op te lossen, gebruiken aanstaande docenten vooral gedrag dat rechtstreeks te maken heeft met de aard van de aandachtseisende situaties waarin het wordt vertoont: lesstofoverdracht in situaties die problemen met de lesstofoverdracht betreffen en in situaties die gaan over het taakgerichte gedrag van leerlingen, organisatorische handelingen in situaties die met waargenomen moeilijkheden met de lesorganisatie te maken hebben, en ordemaatregelen naar aanleiding van niet-taakgericht gedrag van leerlingen. In situaties waarin een negatieve sfeer tussen docent en leerlingen is ontstaan (vooral wanneer niet-taakgericht gedrag van leerlingen als aandachtseisend is ervaren), impliceren acties van aanstaande docenten een afwijzende houding ten opzichte van hun leerlingen.

In alle vier typen situaties nemen aanstaande docenten een afwachtende houding aan door niets te zeggen, te luisteren en een 'onder'-positie in de klas in te nemen. Wanneer zij, ondanks de aandachtseisende situaties, proberen met hun les door te gaan, wordt hun gedrag gekenmerkt door overdracht van de lesstof en organisatie van de les.

Nadere bestudering van de resultaten in de andere vier typen classesituaties ('hulpmiddelen en materialen', 'kritiek', 'matige prestaties' en 'weinig inzet') geeft geen aanvullende inzichten in gedrag van aanstaande docenten in aandachtseisende situaties in de klas.

Gedrag in termen van de twee onderliggende dimensies

In figuur 5 presenteren we scores op de geëxtraheerde dimensies van elke categorie van het methodisch-didactische en het coping-perspectief, en van elk schaalpunt van de machts- en de nabijheidsdimensie van het interpersoonlijke perspectief van docentgedrag. Elke score is tot stand gekomen door de dimensiescores voor elke type van aandachtseisende situaties in de klas te middelen. Hoewel de beschrijving van de resultaten tot dusver verschillen heeft aangetoond tussen typen situaties, zijn de overeenkomsten van dien aard dat we van mening zijn dat een interpretatie van het gedrag van aanstaande docenten in aandachtseisende situaties in de klas op basis van de scores die in figuur 5 grafisch zijn weergegeven, geoorloofd is.

De eerste dimensie van gedrag van aanstaande docenten die is geëxtraheerd (de horizontale as in figuur 5) hebben we de 'activiteitsdimensie' (actief versus passief) genoemd. We hebben deze dimensie omschreven als de intensiteit waarmee een aanstaande docent in classesituaties de inhoudelijke informatie van het gedrag tracht over te brengen door middel van een actieve interactie met leerlingen. *Lage scores* op de activiteitsdimensie kunnen in alle typen van aandachtseisende situaties worden toegekend aan combinaties van afwachten met zwijgen, luisteren en arrangeren. Dergelijk gedrag verwijst naar een 'onder'-positie van aanstaande docenten ten aanzien van hun leerlingen, en kan worden beschouwd als een passieve houding van aanstaande docenten in communicatie in de klas. *Gemiddelde scores* op de activiteitsdimensie kunnen worden toegekend aan instructie-activiteiten ('lesstofoverdracht' en 'helpen') en organisatorische handelingen. In situaties die worden gekenmerkt door niet-taakgericht gedrag van leerlingen (type 4 en 5) zijn aanstaande docenten gericht op het doorgaan met de les, terwijl zij in andere typen classesituaties dit gedrag ook hanteren om waargenomen problemen aan te pakken. *Gemiddelde scores* verwijzen naar gedrag waarbij aanstaande docenten in de communicatie in de klas een passieve noch een actieve houding innemen. *Hoge scores* op de activiteitsdimensie kunnen worden toegekend aan gedrag van aanstaande docenten dat is gericht op het betrekken van leerlingen bij de les ('controleren', 'doorvragen', 'kritiseren' en 'uitnodigen') en op het handhaven van de orde in de klas. Met dergelijk gedrag activeren aanstaande docenten een interactie met hun leerlingen, waarbij zij zijn gericht op het vermijden van of ingrijpen in de situatie die als aandachtseisend is waargenomen.

De tweede dimensie van gedrag van aanstaande docenten die is geëxtraheerd (de verticale as in figuur 5) hebben we de 'spanningsdimensie' (spanning versus ontspanning) genoemd. We hebben de spanningsdimensie omschreven als de mate waarin er door het gedrag van aanstaande

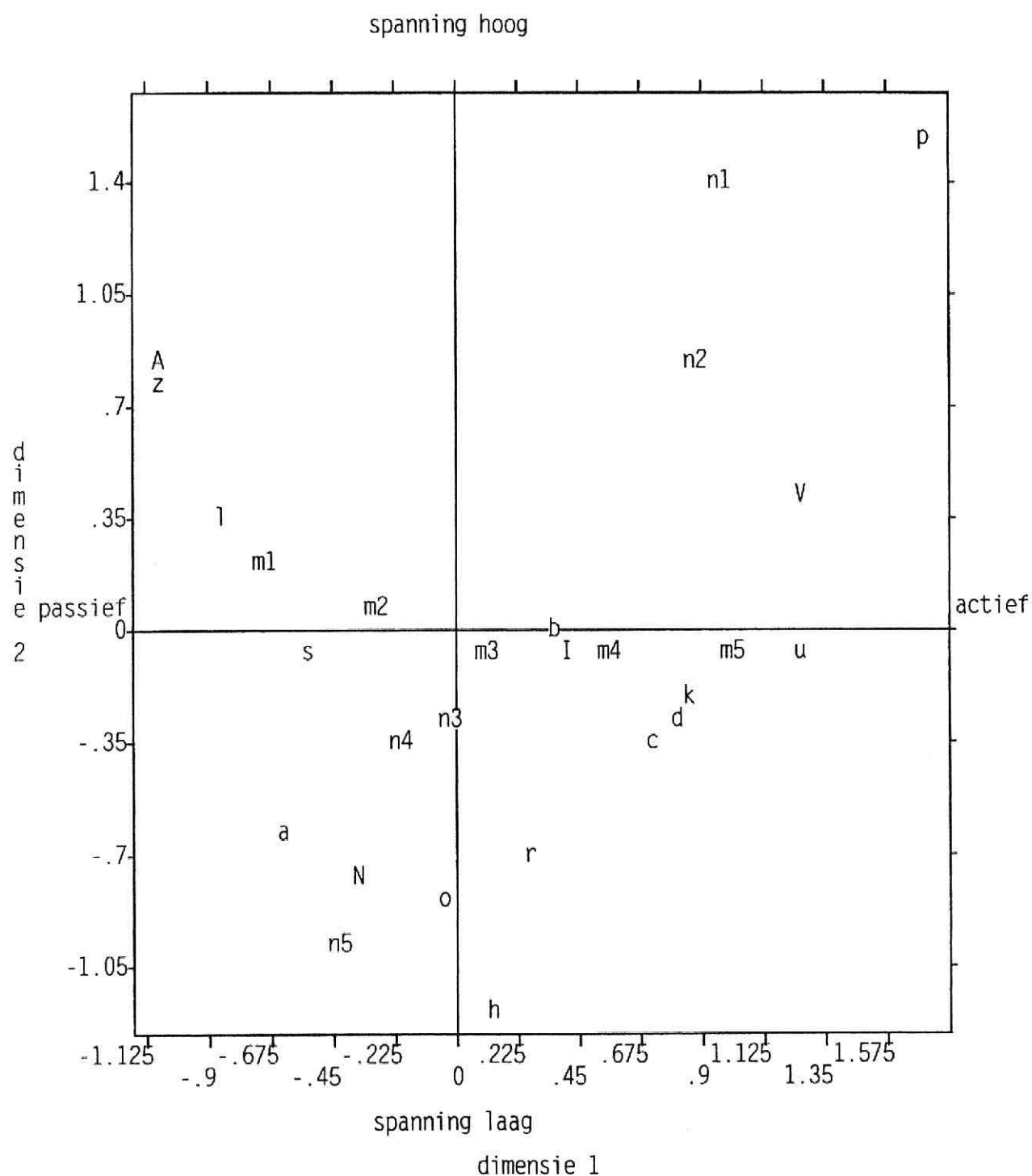


Fig. 5. Scores op de twee dimensies van de categorieën en schaalpunten, gemiddeld over de scores in alle typen van aandachtseisende situaties in de klas. Voor de betekenis van letters en cijfers verwijzen we naar tabel 2.

docenten spanning in de relatie met leerlingen wordt opgeroepen. *Lage scores* op de spanningsdimensie kunnen worden toegekend aan instructie-activiteiten ('lesstofoverdracht' en 'helpen') en organisatorische handelingen ('reguleren' en 'arrangeren') die aanstaande docenten hanteren om waargenomen problemen of aan te pakken of te negeren. Lage scores op deze dimensie houden in dat het gedrag van aanstaande docenten kan worden beoordeeld als gericht op een harmonieuze relatie met de leerlingen. *Gemiddelde scores* op de spanningsdimensie kunnen

worden toegekend aan de resterende categorieën van het methodisch-didactische perspectief, uitgezonderd 'zwijgen' en 'discipline'. De categorieën verwijzen naar gedrag waarmee aanstaande docenten een interactie met hun leerlingen aangaan, zoals leerlingen uitnodigen om antwoorden te geven, en antwoorden en opmerkingen van leerlingen belonen en bekritisieren. Dergelijk gedrag is vooral bedoeld om aandachtseisende situaties in de klas te vermijden, maar soms ook om deze aan te pakken. Door dergelijke gedrag zijn aanstaande docenten gericht op spanning noch op ontspanning in de relatie met hun leerlingen. *Hoge scores* op de spanningsdimensie kunnen worden toegekend aan ordemaatregelen en zwijgen door de aanstaande docenten. Dergelijk gedrag is opgevat als het blijf geven van een afwijzende houding ten aanzien van de leerlingen. Ordemaatregelen zijn gericht op het aanpakken van niet-taakgericht gedrag van leerlingen, maar zijn opgevat als vermijden wanneer niet-taakgericht gedrag van leerlingen niet als aandachtseisend is ervaren (vergelijk figuur 3 met de figuren 1, 2 en 4). Zwijgen door aanstaande docenten daarentegen, verwijst naar een passieve houding in reacties op aandachtseisende klassesituaties. Zowel met ordemaatregelen als met zwijgen geven aanstaande docenten blijf van een hoge mate van spanning in de relatie met hun leerlingen.

CONCLUSIES EN DISCUSSIE

In deze bijdrage hebben we gedrag van aanstaande docenten in het voortgezet onderwijs beschreven tijdens reacties op aandachtseisende situaties in de klas vanuit drie perspectieven: het coping-, het methodisch-didactische en het interpersoonlijke perspectief van gedrag. We hebben gedrag van aanstaande docenten in vier typen aandachtseisende situaties beschreven. Tevens hebben we twee dimensies van gedrag van aanstaande docenten in aandachtseisende situaties in de klas geëxtraheerd. Deze twee dimensies hebben we de activiteits- (actief versus passief) en de spanningsdimensie (spanning versus ontspanning) genoemd. De eerste dimensie hebben we omschreven als de intensiteit waarmee een aanstaande docent in klassesituaties de inhoudelijke informatie van het gedrag tracht over te brengen door middel van een actieve interactie met leerlingen; de tweede dimensie als de mate waarin er door het gedrag van aanstaande docenten spanning in de relatie met leerlingen wordt opgeroepen. Alvorens enkele interpretaties en conclusies te presenteren, willen we aandacht besteden aan twee kanttekeningen bij de betekenis van de resultaten van het onderzoek.

Ten eerste concentreren we ons in dit onderzoek op gedrag van aanstaande docenten en niet op leerlinggedrag of interacties tussen docent en leerlingen. Dit betekent dat we slechts één aspect van de communicatie in de klas tussen aanstaande docent en hun leerlingen hebben bestudeerd. Docent en leerlingen wisselen in communicatie in de klas voortdurend boodschappen uit. In deze cyclische communicatieprocessen (zie Wubbels, Créton & Holvast, 1988) is het bepalen van de oorzaak van een problematische uitwisseling niet erg zinvol. Hoewel we zijn gericht op gedrag van aanstaande docenten, zijn wij niet van mening dat docentgedrag de aanleiding van alle problemen is. Echter, in de lerarenopleiding is het gedrag van aanstaande docenten het meest voor de hand liggende aangrijpingspunt voor verandering.

Ten tweede bepalen in ons onderzoek de respondenten zelf de aard van de aandachtseisende situaties en de mate waarin een situatie aandacht heeft gekost. Voor deze wijze van selecteren van aandachtseisende situaties is gekozen om de ecologische validiteit van de resultaten te bevorderen. Echter, andere docenten zouden hetzelfde gedrag van aanstaande docenten en leerlingen anders kunnen benoemen. Bijvoorbeeld, de ene docent kan passief gedrag van leerlingen wanneer zij het materiaal voor een natuurkundepracticum moeten pakken, benoemen als een probleem met de lesorganisatie, terwijl de andere docent dergelijk gedrag als een teken van onverschilligheid van de leerlingen kan zien. Dit betekent dat we ons bij het beschrijven van gedrag in verschillende typen aandachtseisende situaties bewust moeten blijven van het feit dat de situaties zijn benoemd op basis van een subjectieve indruk van de betreffende aanstaande docenten.

Een resultaat van ons onderzoek lichten we toe vanuit de systeemtheorie van communicatie

(Watzlawick, Beavin & Jackson, 1967), die door Créton en Wubbels (1984) is bewerkt voor de onderwijscontext, namelijk het gevaar van symmetrisch reageren. De beschrijvingen van docentgedrag in aandachtseisende situaties in de klas geven aan dat aanstaande docenten problemen die tijdens het onderwijs ontstaan, aanpakken door gedrag te vertonen dat vergelijkbaar is met de aard van de aandachtseisende situatie. Aanstaande docenten hanteren met name instructie-activiteiten in situaties waarin problemen met de lesstofoverdracht aandacht kost, organisatorische handelingen in situaties waarin de lesorganisatie als problematisch ervaren wordt, ordemaatregelen in situaties die niet-taakgericht gedrag van leerlingen betreffen en alle typen van instructie- en organisatorische handelingen in reactie op taakgericht gedrag van leerlingen. Dergelijke manieren van reageren kunnen problematisch worden in situaties waarin niet-taakgericht gedrag van leerlingen als aandachtseisend is ervaren (dit is het geval in ongeveer 40% van alle klassesituaties in ons onderzoek). In dat geval kunnen ordemaatregelen, zoals het corrigeren van leerlingen door hen te vermanen en te waarschuwen, en leerlingen uit de klas te sturen, resulteren in een problematische interactie tussen docent en de leerlingen. Wubbels, Créton en Holvast (1988) maken aannemelijk dat in zulke situaties 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 een 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 versterken. Resultaten uit het onderhavige onderzoek geven aan dat de kans op een dergelijk symmetrisch destructief patroon bij aanstaande docenten reëel is. Créton, Wubbels en Hooymayers (1989) geven aanwijzingen over de wijze waarop dit patroon kan worden doorbroken.

Een tweede resultaat dat toelichting behoeft, betreft het nut van ons beschrijvend kader voor onderzoek naar gedrag van aanstaande docenten in aandachtseisende situaties in de klas. Door naast de aard van de aandachtseisende situatie drie perspectieven op docentgedrag te onderscheiden, kunnen onderlinge relaties worden onderzocht. Dit houdt in ons onderzoek bijvoorbeeld in dat ingrijpen door aanstaande docenten in een niet-taakgerichte klas vooral betekent dat de aanstaande docenten ordemaatregelen nemen waarbij de docent bepaalt wat er in de klas gebeurt en waarin de docent gedrag van leerlingen afwijst. De reductie tot een model met twee dimensies geeft de mogelijkheid om een meetinstrument te ontwikkelen om relatief eenvoudig gedrag van aanstaande docenten in aandachtseisende situaties vanuit het coping-, het methodisch-didactische en het interpersoonlijke perspectief in kaart te brengen. Een dergelijk instrument kan worden gebruikt in een meer complexe onderzoekssituatie, zoals onderzoek naar de ontwikkeling van docentgedrag gedurende een lange tijdsperiode (bijvoorbeeld gedurende het eerste jaar waarin een docent les geeft). Een dergelijk onderzoek zou ook de relatie kunnen verhelderen tussen reacties van docenten op aandachtseisende situaties en effecten daarvan op lange termijn.

Het kader met twee dimensies kan wellicht ook worden gebruikt in de begeleiding van aanstaande docenten. Het kan bruikbaar zijn om blinde vlekken van aanstaande docenten voor de gang van zaken in de klas tijdens aandachtseisende situaties te signaleren. Wubbels, Créton en Holvast (1988) rapporteren dat beginnende docenten in het voortgezet onderwijs dikwijls blind worden voor dat wat er in de klas gebeurt om vervolgens zichzelf en hun leerlingen ieder hun eigen gang te laten gaan. De docent is dan in beslag genomen door de lesstof en laat toe dat de meeste leerlingen iets anders doen. Hoewel het gedrag van leerlingen in zo een situatie de docent niet per se stoort, is dit geen produktieve situatie. Ons kader geeft een mogelijkheid om, bijvoorbeeld bij het nabespreken van lessen, aanstaande docenten te helpen zich bewust te worden van hun blinde vlekken. Vervolgens moet dan met behulp van resultaten van ander onderzoek, maar ook aan de hand van (praktijk)ervaring van begeleiders worden geprobeerd onproduktief gedrag te veranderen. Uiteraard is daarvoor confrontatie met de gegevens alleen niet voldoende. Intensieve lesnabespreking, oefening en begeleiding is over het algemeen noodzakelijk voor verandering van situaties waarin docenten blind zijn geworden voor aspecten van de communicatie in de klas (Créton, Hermans & Wubbels, 1990).

LITERATUUR

- Admiraal, W.F. (1994). *Reacties van docenten op aandachtseisende situaties in de klas*. Utrecht: W.C.C.
- Borg, M.G. (1990). Occupational stress in British educational settings: A review. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 10(2), 103-126.
- Brooks, D.M. (1985). The teachers communicative competence: the first day of school. *Theory into Practice*, 24(1), 63-70.
- Clark, C.M. & Peterson, P.L. (1986). Teachers' thought processes. In M.C. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd edition), (p. 255-296). New York: MacMillan.
- Commissie Toekomst Leraarschap (1993). *Het gedroomde koninkrijk*. Culemborg: Educatieve Partners Nederland.
- Créton, H., Hermans, J. & Wubbels, T. (1990). Improving interpersonal teacher behaviour in the classroom: a systems communication perspective. *South Pacific Journal of Teacher Education*, 18(2), 85-94.
- Créton, H.A., Wubbels, T. & Hooymayers, H.P. (1989). Escalated disorderly situations in the classroom and the improvement of these situations. *Teaching & Teacher Education*, 5(3), 205-215.
- Créton, H.A. & Wubbels, Th. (1984). *Ordeproblemen bij beginnende leraren*. Utrecht: W.C.C..
- Doyle, W. (1986). Classroom organization and management. In M.C. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd edition), (p. 392-431). New York: MacMillan.
- Evertson, C.M. (1989). Improving elementary classroom management: a school-based training program for beginning the year. *Journal of Education Research*, 83(2), 82-90.
- 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.
- Freeman, A. (1987). The coping teacher. *Research in Education*, 38, 1-38.
- Gage, N.L. (1978). *The scientific basis of the art of teaching*. New York: Teacher College Press.
- Gifi, A. (1983). *Principals user's guide*. Leiden: Department of Datatheory.
- Good, T.L. & Brophy, J.E. (1991). *Looking in classrooms* (5th edition). New York: Harper Collins.
- Hertog, P.C. den (1990). *Persoonlijkheidskenmerken en causale attributies*. Amsterdam: Thesis Publishers.
- Hirschman, A.O. (1970). *Exit, voice and loyalty: response to decline in firms, organizations, and states*. Cambridge (Mass.): Harvard University Press.
- Hock, R.R. (1988). Professional burnout among public school teachers. *Public Personnel Management*, 17(2), 167-189.
- Jambu, M. (1991). *Exploratory and multivariate data analysis*. San Diego: Academic Press.
- 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..
- Leary, T. (1957). *An interpersonal diagnosis of personality*. New York: Ronald Press Company.
- 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.
- Ninomiya, A. & Toshitaka, O. (1990). A critical analysis of job-satisfied teachers in Japan. *Comparative Education*, 26(2/3), 249-257.
- Peters, V.A.M. (1985). *Docenten en hun probleemsituaties*. Nijmegen: Academisch Proefschrift.
- Popping, R. (1988). On agreement indices for nominal data. In W.E. Saris & I.N. Gallhofer (Eds.), *Sociometric Research, volume I. Basingstoke*, (p. 90-105). Hampshire: MacMillan.
- Rudow, B. (1990). Empirische Untersuchungen zur Belastung und Beanspruchung in Lehrerberuf. *Zeitschrift für Pädagogische Psychologie*, 4(2), 75-85.
- Tartwijk, J.W.F. van (1993). *Docentgedrag in beeld*. Utrecht: W.C.C..
- Watzlawick, P., Beavin, J.H. & Jackson, D.D. (1967). *The pragmatics of human communication*. New York: Norton.
- Westerhof, K.J. (1989). *Effectiviteit van leerkrachtgedrag*. Groningen: Dissertatie.
- Wubbels, T., Créton, H.A. & Holvast, A. (1988). Undesirable classroom situations: a systems communication perspective. *Interchange*, 19(2), 25-40.
- Wubbels, T., Créton, H., Levy, J. & Hooymayers, H. (1993). The model for interpersonal teacher behavior. In T. Wubbels & J. Levy (Eds.), *Do you know what you look like?*, (p. 13-28). London: The Falmer Press.
- Wulk, J. (1988). *Lehrerbelastung*. Frankfurt: Peter Lang.

Manuscript ontvangen 21-7-1995

Definitieve versie aanvaard 2-2-1996

De constructie en evaluatie van een videotoets voor de beoordeling van bekwaamheid in professionele gespreksvoering

G.N. Smit¹ en H.T. van der Molen²

¹*Afdeling Persoonlijkheds- en Onderwijspsychologie, Rijksuniversiteit Groningen*

²*Faculteit Sociale Wetenschappen, Open universiteit Heerlen*

ABSTRACT

Courses in communication skills for the organizational setting are part of the curriculum of several educational institutions. The aim of these courses is to enhance students' professional communication skills in order to prepare them for their future work. After a course, students are often not tested on the level of their skills. In some educational institutions, simulations are used as assessment instruments, but a drawback is that such simulations are very time-consuming. Therefore, two efficient video-tape tests were developed and their reliability and validity were investigated. The results of this test form are promising. The test can be applied to test students' insight and skill in using separate communication skills. However, a simulation is needed to test students' ability to conduct a whole interview, such as the selection interview.

INLEIDING

Trainingen in professionele gespreksvoering voor de zakelijke setting zijn momenteel een onderdeel van tal van studierichtingen en (bedrijfs)opleidingen, zoals Bedrijfskunde, de HEAO en Psychologie. De trainingen hebben tot doel studenten deskundiger te maken in het voeren van een aantal gesprekstypen, zoals het functionerings- en het selectiegesprek. Sinds ongeveer 1980 wordt in de meeste trainingen gebruik gemaakt van een gestructureerde trainingsmethode (Ivey, 1971; Ivey & Authier, 1978; Lang & Van der Molen, 1992; Oomkes, 1980; Vrolijk, Dijkema & Timmerman, 1984). Meta-analyses (Baker & Daniels, 1989; Baker, Daniels & Greeley, 1990; Van der Molen, Smit, Hommes & Lang, in druk) tonen aan dat gestructureerde gespreksstrainingen over het algemeen effectief zijn.

Hoewel er veel tijd aan het ontwikkelen en uitvoeren van gespreksstrainingen wordt gespendeerd, wordt bij veel opleidingen na afloop van een training niet getoetst in hoeverre individuele studenten het gewenste *vaardigheidsniveau* hebben behaald. Toetsing vindt in Nederland wel plaats bij de medische opleidingen (bijv. Bögels, 1994; Dochy & Van Luyk, 1987; Kraan & Crijnen, 1987; Pieters, 1991; Vu & Barrows, 1994) en het Middelbaar Dienstverlenings- en Gezondheidszorg Onderwijs (Vermeulen, 1993). Daar gebruikt men gedragstoetsen in de vorm van rollenspelen, waarin een acteur de rol van patiënt of cliënt op zich neemt. In andere dan genoemde opleidingen wordt veelal volstaan met een aanwezigheidsplicht of een literatuurtentamen. Deze lacune is in de eerste plaats toe te schrijven aan het feit dat de constructie en evaluatie van vaardigheidstoetsen nog geen gemeengoed is en lastig blijkt te zijn (Baker, O'Neil & Linn, 1993; Quellmalz, 1991). Wat betreft de evaluatie van vaardigheidstoetsen is de afgelopen jaren veel vooruitgang geboekt, hetgeen blijkt uit de hoeveelheid literatuur waarin een bijdrage wordt geleverd aan de vorming van een standaard waarmee de kwaliteit van de vaardigheidstoetsen kan worden beoordeeld (o.a. Baker, O'Neil & Linn, 1993; Linn, Baker & Dunbar, 1991; McGaghie, 1991; Messick, 1989; Moss, 1992; Quellmalz, 1991). In de tweede plaats zijn

rollenspeltoetsen, waarin individuele studenten één voor één moeten worden getoetst, vaak niet haalbaar vanwege de hoeveelheid tijd die is gemoeid met hun afname en scoring.

De onvrede met de bestaande vormen van tentaminering bij gesprekstrainingen gecombineerd met de groeiende know-how op het gebied van de constructie en evaluatie van vaardigheidstoetsen vormden de aanleiding voor de afdeling Persoonlijkheds- en Onderwijspsychologie van de Rijksuniversiteit Groningen om in 1991 een project¹ van start te laten gaan met het doel toetsvormen voor gesprekstrainingen te ontwikkelen en deze te onderzoeken op kwaliteit. We richten ons in dit artikel op één training, namelijk een training gericht op zakelijke gesprekken in organisaties. In de eerste plaats zijn voor deze training rollenspeltoetsen ontwikkeld (Smit, 1994; Smit, 1995a; Smit, 1995b). Een zwaarwegend nadeel van deze toetsvorm is dat de score van een student sterk afhankelijk is van de specifieke inhoud van de gesimuleerde gesprekssituatie (o.a. Bögels, 1994; Smit, 1995b; Vu & Barrows, 1994), het probleem van 'casus-specificiteit', zodat meerdere rollenspeltoetsen moeten worden afgenomen om generaliseerbare uitspraken over de capaciteiten van een student te kunnen doen. Vanwege de hoeveelheid tijd die hiermee gemoeid is, is in de tweede plaats een meer efficiënte toetsvorm geconstrueerd: een videotoets. In dit artikel wordt de inhoud van de training en de inhoud en opbouw van de twee versies van een videotoets² (versie A en B) beschreven. De toetsvorm is echter ook bruikbaar voor andere gesprekstrainingen, zoals trainingen gericht op hulpverleningsgesprekken. Daarna wordt gerapporteerd over een onderzoek naar de betrouwbaarheid en validiteit van de videotoetsen.

De training "Gesprekken in organisaties"

De training "Gesprekken in organisaties" is een gevorderde training voor psychologiestudenten die wordt verzorgd door de Vakgroep Psychologie. In de training worden drie gesprekstypen geoefend die van belang zijn in de context van organisaties: het functionerings-, selectie- en slecht-nieuwsgesprek (Gramsbergen-Hoogland & Van der Molen, 1993). De aparte basisgespreksvaardigheden, zoals vragen stellen en samenvatten, worden geïntegreerd geoefend in volledige gespreksmodellen, zoals bijvoorbeeld een model voor het selectiegesprek. In zo'n model wordt het gesprek in fasen onderverdeeld waarbij per fase wordt gespecificeerd welke gedragsstappen en gespreksvaardigheden daarin thuis horen. De basisvaardigheden worden bekend verondersteld, omdat die een verplicht onderdeel zijn van het predocoraal programma.

Er wordt in de training gewerkt volgens de methode van cumulatieve microtraining (Lang & Van der Molen, 1992), wat inhoudt dat een vast stamien wordt gevolgd: theoretische instructie, het bekijken van videovoorbeelden, zelf oefenen van een gespreksmodel met een medestudent en het krijgen van feedback van een observator. De gehele training beslaat 10 bijeenkomsten van drie uur. De volgende eindtermen (De Groot, 1982) moeten aan het einde van de training door studenten zijn bereikt:

- 1) het verwerven van kennis en inzicht in de drie gesprekstypen.
- 2) het verkrijgen van vaardigheid in het voeren van de drie gesprekstypen, waarbij een bij het gesprekstype passende houding aangenomen dient te worden.

Een passende houding voor bijvoorbeeld een selecteur in een selectiegesprek is dat deze zakelijk en doelgericht is, zodat in beperkte tijd voldoende relevante informatie over de capaciteiten van de sollicitant kan worden verzameld.

De videotoets

Bij de constructie van de videotoets hebben we gebruik gemaakt van ervaringen die elders zijn opgedaan. In Assessment Centers worden video-opnamen in incidentele gevallen gebruikt als vervanger van een "live" acteur om de sociale vaardigheden van een sollicitant te beoordelen (Eyckman, 1990/1991). RPD Advies maakt bijvoorbeeld gebruik van deze vorm als onderdeel van hun selectie-instrumentarium. De sollicitant moet daarbij mondeling reageren op hetgeen de acteur op de video-opname zegt. De reacties van de sollicitant worden op hun beurt ook weer op video opgenomen en achteraf beoordeeld.

Bij Gezondheidswetenschappen aan de Rijksuniversiteit Limburg is in het kader van onder-

wijs waarin studenten oefenen met het leiden van een groepsgesprek ervaring opgedaan met het vertonen van video-opnamen. Na elk fragment worden vragen gesteld die kennis en inzicht in het gebruik van de vaardigheden betreffen (Dochy & Van Luyk, 1987). Wij hebben in onze toets de video-opnamen op de twee hierboven beschreven manieren ingezet, alleen laten we studenten antwoorden schriftelijk noteren, zodat de toets klassikaal kan worden afgenomen.

De videotoets heeft tot doel de eindtermen kennis, inzicht en vaardigheid te dekken in het functionerings-, selectie- en slecht-nieuwsgesprek. De toets bestaat uit vier korte en drie lange opnamen van gespreksfragmenten die het functionerings-, selectie-, en slecht-nieuwsgesprek betreffen. Het aantal vragen behorende bij de verschillende gesprekstypen is evenredig verdeeld. Van de lange fragmenten ontvingen de studenten een transcript, zodat ze rustig de gelegenheid hebben om het gesprek te volgen. Naar aanleiding van de fragmenten worden drie typen vragen gesteld. In de eerste plaats zijn dat kennis- en inzichtvragen met betrekking tot een gespreksmodel. In de tweede plaats zijn er vragen geconstrueerd die het herkennen van het adequaat en niet adequaat gebruiken van de basisvaardigheden en gevorderde vaardigheden, zoals kritiek geven in het functioneringsgesprek en doorvraagtechnieken in het selectiegesprek, betreffen. In de derde plaats hebben we getracht de beheersing van basisgespreksvaardigheden en gevorderde vaardigheden binnen één gespreksmodel te toetsen door studenten te laten opschrijven wat zij naar aanleiding van een gespreksactie van de werknemer, zoals dat wordt getoond in een fragment, zouden zeggen. We geven van elk vraagtype een voorbeeld in tabel 1.

Er wordt na elk fragment de gelegenheid gegeven voor het beantwoorden van de vragen doordat het beeld gedurende de antwoordtijd blauw is. In totaal zijn er 19 vragen, waarvan sommige zijn onderverdeeld in een a- en een b-vraag.

In de beoordelingsinstructie wordt per vraag aangegeven voor welke onderdelen van een antwoord punten toegekend mogen worden. Tevens worden voorbeelden gegeven van goede en foute antwoorden. De score van een student wordt berekend door alle punten bij elkaar op te tellen; het bereik loopt van 0 tot 42. Deze score hebben we met behulp van een lineaire transformatie omgezet naar een tienpuntsschaal.

Dekking doelen training

Het primaire doel van de videotoets is dat daarmee kan worden nagegaan in hoeverre studenten voldoen aan de trainingsdoelen ten aanzien van kennis, inzicht en vaardigheid in de centraal staande gesprekstypen. De videotoets leent zich als toetsvorm goed voor het stellen van kennis- en inzichtvragen. Daarnaast is het mogelijk vragen te stellen die een beroep te doen op het toepassen van verscheidene deelvaardigheden, zoals bijvoorbeeld het toepassen van de vaardigheid "reageren op kritiek" naar aanleiding van een kort videofragment. Het is echter niet correct om te concluderen dat iemand die de deelvaardigheden beheerst ook een geheel gesprekstype adequaat kan voeren (De Groot, 1982). Messick (1994) besteedt in een artikel over de validatie van vaardigheidstoetsen uitgebreid aandacht aan het onderscheid tussen het toetsen van de complexe gehele vaardigheid versus het toetsen van deelvaardigheden. Voorstanders van toetsing van de complexe vaardigheid in haar geheel gebruiken bij voorkeur rollenspeltoetsen, omdat die vragen om demonstratie van bekwaamheid in gespreksvoering in een realistische

Tabel 1. Voorbeelden toetsvragen.

-
- 1) Welke fouten maakt Wendy bij het meedelen van het slechte nieuws?
 - 2) De leidinggevende reageert niet adequaat op de kritiek van de werknemer. Geef aan wat hij fout doet.
 - 3) De selecteur vraagt de sollicitant om verduidelijking van een "gat" in zijn curriculum vitae.
 - a) Noem de vaardigheid die de selecteur zou kunnen gebruiken naar aanleiding van de reacties van de sollicitant.
 - b) Schrijf woordelijk uit wat de selecteur kan zeggen, als hij de bij a) genoemde vaardigheid toepast.
-

gesprekssituatie. Ze verdedigen hun keuze met het volgende argument: “complicated skills and competencies owe their complexity not just to the number of components they engage but also to the interactions among the components and heuristics calling for them” (Resnick & Resnick, 1991, p. 42). Messick pleit er echter voor ook het toetsen van deelvaardigheden een plaats te geven in het onderwijs. Het voordeel van toetsen op het niveau van deelvaardigheden is dat valt na te gaan op welke punten de groep of een individu nog problemen heeft, zodat aanvullende training op deze punten kan worden aangeboden. Daarnaast vormt de beheersing van de deelvaardigheden een voorwaarde voor het kunnen uitvoeren van de complexe vaardigheid.

Afgezien van bovengenoemd onderscheid levert een ander onderscheid, dat Messick (1994) haarscherp benoemt, dezelfde tweedeling in toetsvormen op. Bij het toetsen van vaardigheden zijn twee wegen te volgen. Een toets kan zich richten op de beoordeling van de uitvoering van een taak sec of op beoordeling van het domein van kennis, deelvaardigheden en strategieën dat een voorwaarde vormt voor het kunnen uitvoeren van de vaardigheid. Bij de eerste benadering, de taak-benadering, wordt de taak als uitgangspunt genomen en worden later criteria opgesteld. Het probleem van casus-specificiteit dient zich bij deze benadering nadrukkelijk aan. Daarnaast is een tweede probleem dat er construct-irrelevante zaken kunnen worden beoordeeld. In het geval van de rollenspeltoets kan een beoordelaar bijvoorbeeld impliciet in de beoordeling mee laten wegen of iemand goed formuleert, ook al is dat geen beoordelingscriterium. Bij de tweede benadering, de construct-benadering, wordt eerst het domein aan kennis en deelvaardigheden gespecificeerd en ter dekking van dit domein worden vragen en opdrachten geconstrueerd. Het gevaar dat hier speelt is dat het complexe construct “beheersing van een gespreksmodel” niet volledig wordt gedekt. Bij de constructie van de videotoets is de constructbenadering toegepast.

METHODE

Proefpersonen

Aan het onderzoek hebben 69 psychologiestudenten die de training volgden (trainingsgroep) en 34 psychologiestudenten die nog geen training volgden (vergelijkingsgroep) meegewerkt. De trainingsgroep bestond uit gevorderde psychologiestudenten die in de collegejaren 1992/1993 en 1993/1994 aan de training deelnamen. De personen uit de vergelijkingsgroep waren - net als de personen uit trainingsgroep - gevorderde psychologiestudenten die eveneens een basistraining in gespreksvoering hadden gevolgd. Personen uit deze groep ontvingen als tegenprestatie voor deelname een bescheiden financiële vergoeding en na de nameting een terugkoppeling van hun resultaten. De trainingsgroep telde 20 mannen (29%) en 49 vrouwen (71%). In de vergelijkingsgroep zaten 15 mannen (44.1%) en 19 vrouwen (55.9%). De gemiddelde leeftijd in beide groepen was 24 jaar. In de trainingsgroep gaven iets meer studenten aan (54%) ervaring met gespreksvoering te hebben opgedaan in werk of een andere opleiding dan psychologie dan personen in de vergelijkingsgroep (44%).

Onderzoeksopzet

De studenten uit de trainings- en vergelijkingsgroep hebben de toetsen op twee tijdstippen gemaakt: tijdstip 1 viel voor de training en tijdstip 2 viel na de training. Bij de voormeting kreeg

Schema 1. Afnameschema toetsen

Trainingsgroep	video-toets	A B	training	video-toets	B A
Vergelijkingsgroep	video-toets	A B	geen training	video-toets	B A

A= versie A; B=versie B

de helft van de personen uit de trainings- en de vergelijkingsgroep versie A van de toets te maken en de andere helft kreeg versie B. Bij de nameting kreeg men een andere toetsversie. In schema 1 wordt de opzet weergegeven.

Na de afnames hebben drie getrainde beoordelaars de toetsen onafhankelijk van elkaar nagekeken.

Ten behoeve van het onderzoek naar de begripsvaliditeit stuiten we op het probleem dat voor inhoudelijke validatie van de toets een goed criterium, een 'gouden standaard', ontbreekt (Vu & Barrows, 1994): van studenten zijn immers geen scores voorhanden hoe goed ze in een werksituatie de in de training geoefende gesprekken kunnen uitvoeren. Een maat die wel haalbaar is in een onderwijssituatie is het trainersoordeel dat we als criterium voor het valideren van de toetsen hebben gebruikt. Het trainersoordeel dat wordt gebaseerd op de geobserveerde prestaties van een student op verschillende momenten in verschillende omstandigheden is voor velen een geloofwaardige maat, maar wordt volgens Hopkins, George en Williams (1985) verbazingwekkend weinig gebruikt. Elke trainingsgroep werd door twee trainers begeleid. De trainers hebben onafhankelijk van elkaar de studenten een cijfer op een tienpuntsschaal gegeven, waarmee ze uitdrukten in hoeverre een student aan het einde van de training de te leren gespreksvaardigheden en -modellen beheerste. Met het oog op die onafhankelijkheid was de trainers gevraagd om gedurende de training onderling niet te bespreken hoe goed ze een bepaalde student vonden. Wel bespraken ze in het algemeen didactische methoden voor studenten die moeite hadden om bepaalde onderdelen onder de knie te krijgen. We rapporteren de ongecorrigeerde en voor onbetrouwbaarheid van de videotoets en trainersoordeel gecorrigeerde correlaties. Als maat voor de betrouwbaarheid is de intraclass-correlatiecoëfficiënt (ICC) (Shrout & Fleiss, 1979) gebruikt.

Om verder zicht te krijgen op de begripsvaliditeit zijn twee hypothesen opgesteld waaraan valide toetsen dienen te voldoen, die we empirisch hebben getoetst. Deze methode levert geen data op ten aanzien van de inhoudelijke validering van de toets; indien de opgestelde hypothesen echter empirisch worden bevestigd, dan biedt dat wel een ondersteuning van de begripsvaliditeit van de videotoetsen (Nunnally, 1978). Hofstee (1971), Messick (1994) en de NIP-richtlijnen (NIP, 1988) stellen in de eerste plaats dat valide toetsen in staat moeten zijn een toename in kennis, inzicht en vaardigheid te registreren. Daartoe hebben we onderzocht of studenten na de training hoger op de toetsen scoren dan ervoor. In de tweede plaats moet een valide toets in staat zijn om te differentiëren tussen groepen die in niveau van beheersing van gespreksvaardigheden verschillen. Daartoe is nagegaan of studenten die training hebben gevolgd meer vooruitgang boekten (verschil scores voor- en nameting) dan de personen uit de vergelijkingsgroep.

Tot slot is de inhoudsvaliditeit van de toets onderzocht. Zeven trainers hebben daartoe hun oordeel gegeven in hoeverre de videotoetsen kennis van, inzicht in en vaardigheid in de drie gesprekstypen dekken. Ze dienden hun oordeel te geven op een schaal van 1 tot 5, waarbij één punt betekende dat een aspect in de toets niet wordt gedekt, 3 dat dat aspect in redelijke mate wordt gedekt en 5 dat dat aspect in zeer sterke wordt gedekt.

RESULTATEN

Studenten met ervaring boekten niet significant meer vooruitgang op de videotoets dan personen zonder ervaring ($p > .05$). Daarom behoeften we bij de analyses niet voor ervaring te corrigeren.

Betrouwbaarheid

Interne consistentie

De interne consistentie is per toetsvorm berekend voor de hele groep, dat wil zeggen over de toetsscores van voor- en nameting van alle studenten. In tabel 2 worden de interne consistenties weergegeven. In de tabellen vermelden we ook steeds de standaarddeviaties, de standaardmeetfouten en het aantal toetsen waarover de alfa is berekend.

Tabel 2. Interne consistentie, standaarddeviaties, standaardmeetfouten en steekproefgrootte voor de videotoets (24 items).

	alfa	sd	se	N
gehele groep versie A	.87	1.45	.52	105
gehele groep versie B	.86	1.43	.53	101

De interne consistenties, berekend voor de hele groep, zijn hoog. De standaarddeviaties en standaardmeetfouten lopen voor versie A en B zeer weinig uiteen. De standaardmeetfouten liggen in de orde van grootte van een halve punt op een schoolcijferschaal.

Als overeenstemmingsmaten hebben we de Pearson produkt-moment-correlatiecoëfficiënt en de getransformeerde Gower berekend (Gower, 1971; Hofstee & Zegers, 1991) die de overeenstemming tussen beoordelaars op absoluut niveau weergeeft. De range van de getransformeerde Gower, loopt net als bij de correlatiecoëfficiënt, van -1 tot 1. De gevonden waarden worden in tabel 3 weergegeven.

Uit Tabel 3 blijkt dat de interbeoordelaarsovereenstemming bij de videotoets voor alle beoordelaarsparen hoog tot zeer hoog te noemen is.

Tabel 3. Interbeoordelaarsovereenstemming videotoets met behulp van de correlatiecoëfficiënt (r) en de getransformeerde Gower (G*).

	Psychologie	
	r	G*
beoordelaar 1 en 2	.91	.89
beoordelaar 1 en 3	.90	.87
beoordelaar 2 en 3	.90	.88

Validiteit

De overeenstemming tussen trainersoordelen en observatoroordelen is .23 ($p < .05$). Deze waarde wordt na correctie voor onbetrouwbaarheid niet veel hoger, namelijk .28, omdat de videotoets een betrouwbaar instrument is en ook de trainers een goede interbeoordelaarsovereenstemming bereiken ($ICC = .73$). De gevonden correlatie biedt enige ondersteuning met betrekking tot de begripsvaliditeit van de videotoets.

In Tabel 4 worden de descriptieve gegevens voor de videotoets vermeld in het kader van de beantwoording van de vragen of de videotoetsen een toename in kennis-, inzichts- en vaardig-

Tabel 4. Steekproefgrootte (N), gemiddelden (M) en standaarddeviaties (sd) voor de videotoets.

	Voormeting			Nameting	
	N	M	sd	M	sd
Trainingsgroep	69	3.18	.88	5.44	.99
Vergelijkingsgroep	34	2.93	.84	3.10	1.07

Tabel 5. Inhoudsvaliditeit videotoets: Gemiddelden (M), standaarddeviaties (Sd) voor de verschillende gesprekstypen wat betreft kennis, inzicht en vaardigheden (schaal van 1 tot 5).

	Kennis		Inzicht		Vaardigheid	
	M	Sd	M	Sd	M	Sd
FG	2.00	.00	3.29	1.11	3.29	1.11
SL	2.57	.79	3.57	.79	3.00	1.15
SN	2.71	.95	3.71	.76	3.00	1.41

Noot: FG=functioneringsgesprek, SL=selectiegesprek, SN=slecht-nieuwsgesprek

heidsniveau kunnen registreren en of met de toetsen valt te differentiëren tussen getrainde en niet getrainde studenten.

Bij de voor- en nameting wordt verschillend op versie A en B gescoord; hiermee houden we rekening bij verdere analyses. De trainingsgroep scoort bij de nameting gemiddeld hoger dan bij de voormeting (volgorde A-B: $T=-18.87$, $df=35$, $p<.001$; volgorde B-A: $F=-9.65$, $df=32$, $p<.001$). De trainingsgroep gaat tevens significant meer vooruit dan de vergelijkingsgroep (volgorde A-B: $F(1,50)=68.58$, $p<.001$; volgorde B-A: $F(1,49)=47.85$, $p<.001$).

In tabel 5 worden de gegevens voor de inhoudsvaliditeit van de videotoets weergegeven.

De videotoets dekt volgens de docenten in redelijke tot sterke mate inzicht en vaardigheid in de drie gesprekstypen. Kennis wordt in geringe tot redelijke mate gedekt.

CONCLUSIES

Het onderhavige onderzoek levert aanwijzingen dat videotoetsen voor de beoordeling van 'bekwaamheid in het voeren van gesprekken' voldoen aan de eisen van betrouwbaarheid en validiteit. Uit het betrouwbaarheidsonderzoek bleek dat de interne consistenties voor beide versies van de videotoets hoog zijn en daarnaast is de interbeoordelaarsbetrouwbaarheid zeer hoog. Dit laatste resultaat wijst erop dat de beoordelingscriteria eenduidig zijn geformuleerd.

De validiteitseis is onderverdeeld in begripsvaliditeit en inhoudsvaliditeit. De samenhang tussen trainersoordelen en de scores op de videotoets biedt in de eerste plaats enige steun aan de begripsvaliditeit van de videotoetsen. Een verdere ondersteuning van de begripsvaliditeit zien we in het gegeven dat met de toetsen een toename in kennis, inzicht en vaardigheid in het functionerings-, selectie- en slecht-nieuwsgesprek valt te registreren. Tevens is met de toetsen te differentiëren tussen groepen waarvan valt te verwachten dat ze in niveau van beheersing verschillen. Uit het onderzoek naar de inhoudsvaliditeit van de videotoetsen blijkt dat inzicht en vaardigheid met betrekking tot een bepaald gesprekstype in redelijke tot sterke mate worden gedekt, terwijl kennis in mindere mate wordt gedekt.

Over het geheel genomen, blijken de onderzoeksresultaten een gunstig beeld van de betrouwbaarheid en de validiteit op te leveren. Daarbij moet wel de volgende kanttekening worden geplaatst. In de toets wordt wat betreft "vaardigheid" alleen het kunnen toepassen van deelvaardigheden getoetst. Daarentegen wordt in rollenspeltoetsen wel de complexe vaardigheid van het voeren van een gesprek in haar geheel getoetst. In hoeverre er een verband bestaat tussen het kunnen toepassen van geïsoleerde deelvaardigheden en het geïntegreerd kunnen gebruiken van die vaardigheden in een gesprek moet met vervolgonderzoek nader worden onderzocht.

De videotoets kent in vergelijking met de rollenspeltoets echter ook voordelen. Het eerste voordeel is dat op eenvoudige wijze verscheidene gesprekssituaties aan de student kunnen worden voorgelegd, waardoor het probleem van casus-specificiteit wordt tegengegaan. Het

tweede voordeel is de efficiëntie van de videotoets, waardoor deze eenvoudig in een onderwijs-situatie is te gebruiken. Het gebruik van de videotoets als toetsingsmethode is een aanzienlijke verbetering ten opzichte van de situatie waarin niet of met een literatuurtentamen wordt getoetst: met de videotoets kan immers beter worden nagegaan of studenten in voldoende mate voldoen aan de trainingsdoelen van inzicht en vaardigheid.

Op basis van bovenstaande overwegingen en de onderzoeksresultaten komen we tot de volgende aanbevelingen. Voor het beoordelen van het geïntegreerd gebruik van vaardigheden in een geheel gespreksmodel is het aan te bevelen om een rollenspeltoets te gebruiken, terwijl de videotoets is aan te raden als een efficiënt instrument voor de beoordeling van inzicht in gespreksmodellen en -vaardigheden en het kunnen toepassen van deelvaardigheden in verschillende gesprekssituaties. Tot slot willen we opmerken dat het toepassen van rollenspel- en videotoetsen vanuit onderwijskundig standpunt wenselijk is, omdat ze studenten stimuleren om zich – naast het bestuderen van de literatuur – te richten op het oefenen van gespreksvaardigheden (Bouhuijs, Van der Vleuten & Van Luyk, 1987; Smit, 1995b). Op deze wijze kunnen dit soort toetsen bijdragen aan de effectiviteit van onderwijs in gespreksvoering.

NOTEN

1. De goede suggesties en het commentaar van G. Lang en W.K.B. Hofstee bij de begeleiding van dit project hebben bijgedragen aan de totstandkoming van dit artikel.
2. We willen G.Dam, G. Oosterveld en P. de Vries bedanken voor hun bijdrage aan het onderzoek dat in dit artikel wordt beschreven.

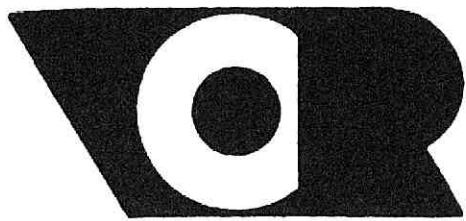
LITERATUUR

- Baker, E. L., O'Neil, H. F. & Linn, R. L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist*, 48, 1210-1218.
- Baker, S. B. & Daniels, T. G. (1989). Integrating research on the microcounseling program: A meta-analysis. *Journal of Counseling Psychology*, 36(2), 213-222.
- Baker, S. B., Daniels, T. G. & Greeley, A. T. (1990). Systematic training of graduate-level counselors: Narrative and meta-analytic reviews of three major programs. *Counseling Psychologist*, 18, 355-421.
- Bögels, S. M. (1994). *Teaching and assessing diagnostic interviewing skills. An application to the mental health field*. Dissertatie. Rijksuniversiteit Limburg, Maastricht.
- Bouhuijs, P. A. J., Vleuten, C. P. M. van der & Luyk, S. J. van der (1987). The OSCE as a part of a systematic skillstraining approach. *Medical Teacher*, 9, 183-191.
- Dochy, F. J. & Luyk, S. J. van (red.). (1987). *Handboek voor vaardigheidsonderwijs*. Lisse: Swets & Zeitlinger.
- Eykman, M. (1990/1991). De sollicitant op de buis. *Vrouw & Bedrijf. Magazine voor de Ondernemende Vrouw*, 5(10), 27-29.
- Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 27, 857-874.
- Gramsbergen-Hoogland, Y. & Molen, H. T. van der (1993). *Gesprekken in organisaties* (2e druk). Groningen: Wolters-Noordhoff.
- Groot, A. D. de (1982). 'Dekkingsproblemen' bij de constructie van examenprogramma's en curricula. In *Handboek voor de onderwijspraktijk: afl. 14* (pag. 2.1.Gro.1. - 2.1.Gro.25). Deventer: Van Loghum Slaterus.
- Hofstee, W. K. B. (1971). Begripsvalidatie van studietoetsen. *Nederlands tijdschrift voor de psychologie*, 26, 491-500.
- Hofstee, W. K. B. & Zegers, F. E. (1991). Idiographic correlation: Modeling judgments of agreement between school grades. *Tijdschrift voor onderwijsresearch*, 16, 331-336.
- Hopkins, K. D., George, C. A. & Williams, D. D. (1985). The concurrent validity of standardized achievement tests by content area using teachers' ratings as criteria. *Journal of Educational Measurement*, 22, 177-182.
- Ivey, A. E. (1971). *Microcounseling*. Springfield, IL: Charles C. Thomas.

- Ivey, A. E. & Authier, J. (1978). *Microcounseling. Innovations in interviewing, counseling, psychotherapy and psychoeducation*. Springfield: Charles C. Thomas.
- Kraan, H. F. & Crijnen, A. (1987). *The Maastricht history taking and advice checklist: Studies of instrumental validity*. Dissertatie. Amsterdam: Lundbeck.
- Lang, G. & Molen, H. T. van der (1992). *Methodiek van gesprekst raining*. Baarn: H. Nelissen.
- Linn, R. L., Baker, E. L. & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, vol. 20(8), 15-21.
- McGaghie, W. C. (1991). Professional competence evaluation. *Educational Researcher*, Vol. 20(1), 3-9.
- Messick, S. (1989). Validity. In R.L. Linn (Ed.). *Educational measurement* (3rd ed., pp.13-103). New York: Macmillan.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.
- Molen, H. T. van der, Smit, G. N., Hommes, M. A. & Lang, G. (in druk). Two decades of microtraining in the Netherlands: A meta-analysis. *Educational Research and Evaluation*.
- Moss, P. A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Review of Educational Research*, 62(3), 229-258.
- Nederlands Instituut van Psychologen (1988). *Richtlijnen voor ontwikkeling en gebruik van psychologische tests en studietoetsen*. Assen: Van Gorcum.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York [etc]: McGraw-Hill Book Company.
- Oomkes, F. R. (1980). *Handboek voor gesprekst raining: programmeren, werkwijzen, hulpmiddelen, procedures en oefeningen* (5e druk). Meppel: Boom.
- Pieters, H. M. (1991). *De Utrechtse consult evaluatie methode. Vaardigheden in consultvoering van huisartsen in opleiding getoetst*. Dissertatie. Lelystad: Meditekst.
- Quellmalz, E. S. (1991). Developing criteria for performance assessments: The missing link. *Applied Measurement in Education*, 4(4), 319-331.
- Resnick, L. B. & Resnick, D. P. (1991). Assessing the thinking curriculum: New tools for educational reform. In R. B. Gifford & M. C. O'Connor (Eds.). *Changing assessments: alternative views of aptitude, achievement and instruction* (pp. 37-75). Boston [etc.]: Kluwer Academic Publishers.
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.
- Smit, G. N. (1994). De beoordeling van professionele gespreksvaardigheden. *De Psycholoog*, 7/8, 266-269.
- Smit, G. N. (1995a). Het rollenspel als toetsvorm. In *Handboek Effectief Opleiden* (9.2-1.01 - 9.2-1.19). 's-Gravenhage: Delwel.
- Smit, G. N. (1995b). *De beoordeling van professionele gespreksvaardigheden. Constructie en evaluatie van rollenspel-, video- en schriftelijke toetsen*. Baarn: H. Nelissen.
- Vermeulen, W. (1993). *Toetsing van communicatieve vaardigheden. Constructie en evaluatie van gedragstoetsen voor professionele gespreksvaardigheden*. Dissertatie. Arnhem: Cito.
- Vrolijk, A., Dijkema, M. F. & Timmerman, G. (1984). *Gespreksmodellen. Een geprogrammeerde instructie* (4e druk). Alphen a/d Rijn: Samsom.
- Vu, N. V. & Barrows, H. S. (1994). Use of standardized patients in clinical assessments: Recent developments and measurement findings. *Educational Researcher*, 23(3), 23-30.

Manuscript ontvangen 12-01-1995

Definitieve versie aanvaard 08-12-1995



BULLETIN

ORGAAN VAN DE VERENIGING VOOR ONDERWIJSRESEARCH

Redactie-adres

Dr. M. Mulder,
Faculteit der Toegepaste Onderwijskunde,
Universiteit Twente,
Postbus 217, 7500 AE Enschede,
tel. 053-4893652
fax 053-4893759

Jrg. 20, nr. 1

1996

Inhoudsopgave

Verslag VOR-ledenvergadering

192

Verslag VOR-ledenvergadering 20 juni 1995: 12.45 - 13.45 uur te Groningen (Martinihal tijdens ORD'95).

Aanwezig: Leden: T. Bruining, W. Hoebe, J.L. van der Linden, H. Lodewijks, P. Kirschner, S. Klatter, E. Moen, M. Peters-Sips, G. Teurlings, J. Slavenburg, R. Voeten, K. Westerhof, A. van Wieringen.
Bestuursleden: J. van den Akker, J. Ax, J. Dronkers, H. van Hout, N. Lagerweij (voorzitter), F.J.J. van Bussel (penningmeester), M. Mulder, J. Scheerens, P.J.J. Stijnen (secretaris).

1. Opening

De voorzitter opent de vergadering om 12.45 uur.

Bericht van verhindering ontvangen van J. Pieters en N. Verloop.

2. Mededelingen

- De voorzitter deelt mede dat het nieuwe ere-lid, prof. dr. A. Knoers, zijn benoeming gisteravond met genoegen heeft aanvaard.
- De EERA komt goed van de grond. De oprichtingsvergadering zal in september in Bath plaatsvinden. Voor het probleem van de contributiegelden voor Oost-Europese landen zal wellicht een oplossing gezocht worden in de richting van het aanpassen van de hoogte van de contributie aan het prijspeil van het land.
- In verband met de gelden voor onderzoek zal binnenkort een gesprek plaatsvinden met de heer Hutter van NWO. Er is ook een vertegenwoordiging van de instituten voor onderzoek voorzien.

3. Verslag van de VOR-ledenvergadering van 25 mei 1994

Het verslag wordt onder dankzegging aan de secretaris ongewijzigd vastgesteld.

4. Vaststelling jaarrekening 1994

De jaarrekening wordt vastgesteld. Besloten wordt voor de reservering van Fl. 15.000,- (voor alle divisies te zamen) een verdeling van gelden per divisie per jaar te maken (naar rato van het ledenaantal van elke divisie).

5. Rapport kascommissie en decharge kascommissie

De heer Voeten rapporteert dat nog niet alle leden van de commissie kennis hebben kunnen nemen van de cijfers. De heer Voeten heeft de cijfers wel gezien. Zijn persoonlijke indruk is dat de penningmeester uitstekend werk heeft verricht en dat de financiële situatie van de vereniging goed is. De financiële gegevens zijn door hem in orde bevonden. Hij stelt voor de penningmeester te dechargeren. Later kunnen de cijfers dan nog eens bekeken worden.

Het voorstel van de heer Voeten wordt door de vergadering vastgesteld.

Onder dankzegging van de voorzitter wordt de kascommissie (de heren Voeten, Berger en Claessen) gedechargeerd.

6. Instelling nieuwe kascommissie

Mevrouw Peters-Sips en de heren Van Wieringen en Stijnen worden bereid gevonden de nieuwe kascommissie te vormen.

7. Begroting 1996

De begroting wordt vastgesteld. Besloten wordt - in afwijking van het voorstel van de penningmeester - de contributie voor studenten niet te verhogen.

8. Wisseling bestuurssamenstelling

Conform het voorstel van het bestuur treden de heren Van Bussel (penningmeester) en Stijnen (secretaris) af. De voorzitter dankt beiden voor hun positieve inzet.

De vergadering gaat akkoord met de benoeming van de heren Mulder (secretaris) en Imbos (penningmeester).

De voorzitter deelt mede dat hij voornemens is in het begin van 1996 af te treden. Hij verzoekt de leden suggesties voor mogelijke kandidaten aan de bestuursleden door te geven.

9. Bestuursstandpunt in verband met positie tijdschriften; toelichting TOR-redactie

De heer Voeten geeft een toelichting op het nieuwe redactiebeleid van TOR. Er zal onder meer gestreefd worden naar meer themanummers en naar meer discussies rond 'leading articles'. Ook zal per divisie een contactpersoon voor TOR worden gezocht.

Na de toelichting wordt het voorstel van het bestuur door de vergadering vastgesteld.

10. Rondvraag

In verband met de kosten voor TOR wordt gewezen op de aantrekkelijkheid van goedkope elektronische produktie van het blad. In de toekomst zou elektronische uitgave sterk kostenverlagend kunnen werken.

11. Sluiting

Niets meer aan de orde zijnde sluit de voorzitter onder dankzegging om 13.40 uur de vergadering.

N.B.: De volgende Algemene Ledenvergadering vindt plaats tijdens de ORD 1996 te Tilburg.

Rendement van het beroepsonderwijs: Beleid en onderzoek

W.Th.J.G. Hoebe

*GION: Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling,
Rijksuniversiteit Groningen*

ABSTRACT

The output of vocational training is disappointingly low. Average yields of 50% (the qualified students as percentages of the yearly influx) are reported for different kinds of vocational training, with some as low as 20% and others above 80%. There are also complaints about the qualitative linkages between vocational training and job requirements. Educational policy tries to raise the output and to improve the qualitative linkages by taking the following actions: mergers of different local schools into regional community colleges so as to offer broad training opportunities; reform of all training into a dual system of practical and theoretical training after the pattern of the apprenticeship system; and regular consultations with representatives of the employers' associations. The goals and means of the educational policy are compared with research that is relevant for the proposed reforms in vocational training and its linkages with job requirements. Empirical data justify severe doubts about the validity of programme theories that may be implied by the educational policy.

INLEIDING

Met het rendement van het secundaire beroepsonderwijs en volwasseneneducatie in Nederland (MBO, KMBO, LLW, PBVE) is het sinds jaren slecht gesteld. Bij een discussie over en onderzoek naar het probleem moet op school- en systeemniveau onderscheid worden gemaakt tussen problemen van intern en van extern rendement, en op individueel niveau tussen slaagkansen van leerlingen en arbeidsmarktkansen van schoolverlaters.

In beleidsdiscussies en -adviezen staat sinds het begin van de jaren tachtig dualisering van het beroepsonderwijs op de agenda. Dualisering wordt bepleit om de aansluiting tussen onderwijs en arbeid te verbeteren en ten dele als oplossing van de rendementsproblemen. In het begin van de jaren negentig is dit uitgekristalliseerd tot een op verdere dualisering van het beroepsonderwijs gericht beleid. Dat beleid is aanvankelijk door de overheid veilig gesteld in convenanten met sociale partners en met onderwijsinstellingen en later enigszins genuanceerd in de Wet Educatie en Beroepsonderwijs (WEB).

Het is de vraag in hoeverre dit beleid door onderzoeksresultaten wordt ondersteund. Tegenover de beleidsvoornemens staat weinig onderzoek. In dit themanummer wordt over de relatie tussen dualisering en rendement van beroepsonderwijs gerapporteerd vanuit vier onderzoeksprojecten. Elk onderzoek leidt tot conclusies die als kritische kanttekeningen bij de beleidsvoornemens kunnen worden geïnterpreteerd.

In deze inleiding wordt kort ingegaan op het onderscheid tussen intern en extern rendement, op het beleid van verdere dualisering van het beroepsonderwijs en op het beschikbare onderzoek. Vervolgens worden de conclusies van de vier in dit themanummer gerapporteerde onderzoeken samengevat. Tenslotte volgt een discussie over de betekenis van deze conclusies voor de bijdrage aan de oplossing van rendementsproblemen die van een verdergaande dualisering kan worden verwacht.

INTERN EN EXTERN RENDEMENT

Rendement van het beroepsonderwijs heeft betrekking op de opbrengsten van dit onderwijs minus de kosten er van. Het is een begrip dat wordt gebruikt om de verhouding tussen kosten en baten aan te geven op het niveau van de onderwijsinstellingen of op het niveau van het landelijke onderwijssysteem. In de discussie wordt meestal alleen met de opbrengsten rekening gehouden. Vanwege de financieringssystematiek van het Nederlandse onderwijs – een vast bedrag per leerling of student – kunnen de kosten immers in beginsel als een constante worden beschouwd. In het beroepsonderwijs is rendementsbekostiging geïntroduceerd in de vorm van een bescheiden verhoging van het vaste bedrag per leerling voor instellingen die verhoudingsgewijs veel gediplomeerden afleveren.

Rendement van het beroepsonderwijs is een begrip dat in de discussies vooral wordt betrokken op de onderwijsopbrengsten. Daarbij wordt gesproken over twee soorten opbrengsten: allereerst opbrengsten in termen van het slagen van leerlingen of hun in de loop van het onderwijs bereikte kwalificaties (intern rendement); en vervolgens opbrengsten in termen van de arbeidsmarktkansen van afgestudeerden, dat wil zeggen wat zij na het onderwijs met hun kwalificaties kunnen doen (extern rendement) (Van Batenburg & Haanstra, 1992; Van Batenburg & Lokman, 1991; Den Boer & Meesterberends-Harms, 1992; De Bont-Hoek e.a., 1989; Hoeben, 1993).

Intern rendement heeft betrekking op kwaliteit van het onderwijs en op de kosten en baten van het realiseren van deze kwaliteit. Intern rendement omvat begrippen als efficiency en effectiviteit van het onderwijs. Rendementsverhoging omvat daarmee uiteenlopende problemen als het bestrijden van voortijdige schooluitval, het realiseren van efficiënte doorstroming binnen en tussen diverse opleidingen, en het waarborgen van het niveau van de kwalificaties bij het einde van het onderwijs. Van Batenburg en Den Boer rapporteren in dit themanummer in deze termen over intern rendement van het middelbaar beroepsonderwijs, waarbij voortijdig schooluitval, efficiënte doorstroming en slagen voor het diploma zijn gecombineerd tot een opbrengstmaat. Het beroepsonderwijs kent kwalificaties van meerdere niveaus. Zonder te letten op deze gedifferentieerde kwalificatiestructuur is zelfs voor relatieve buitenstaanders zichtbaar dat het interne rendement van het beroepsonderwijs laag is. Door de bank genomen valt de helft van de leerlingen voortijdig uit. Gerapporteerde uitvalspercentages variëren tussen 20% en 80% al naar gelang het type onderwijs en de beroepssector. Dergelijke hoge percentages van voortijdige schooluitval vormen het meest zichtbare criterium voor het (lage) rendement van het beroepsonderwijs. Het gevaar bestaat dat voortijdige uitvallers geen primaire startkwalificatie behalen, zodat onderzoek naar de oorzaken van uitval geboden is. Den Boer en Harms rapporteren in dit themanummer over onderzoek naar deze oorzaken.

Extern rendement heeft primair betrekking op de aansluiting van het beroepsonderwijs op de arbeidsmarkt en secundair op doorstroming naar andere vormen van onderwijs (van korte naar lange beroepsopleidingen, van middelbaar naar hoger beroepsonderwijs). Het extern rendement van beroepsonderwijs omvat problemen als de kansen van schoolverlaters op de arbeidsmarkt, de overeenkomsten tussen hun opleidingskwalificaties en de functiekwalificaties enerzijds en vereisten van vervolgonderwijs anderzijds, en de wijze waarop schoolverlaters op de arbeidsmarkt en in het vervolgonderwijs worden gewaardeerd. De opbrengsten van beroepsonderwijs kunnen niet vanuit een onderwijscontext alleen worden beoordeeld, maar ook vanuit arbeidsmarktfactoren en bedrijfscontexten. Onderzoeksgegevens relativeren klachten dat de aansluiting op de arbeidsmarkt slecht is. Werkloosheid onder gediplomeerde schoolverlaters van het secundaire beroepsonderwijs is verhoudingsgewijs laag (Hoeben, 1993; Meesters, 1992; Van der Velden & Wieling, 1994). Onduidelijk is echter of de lage werkloosheidscijfers van gediplomeerde schoolverlaters zijn geflatteerd door hun toegenomen doorstroming naar andere vormen van onderwijs en door hun verdringing van lager geschoolden op de arbeidsmarkt (Van der Velden & Wieling, 1994). Duidelijk is wel dat het beleid recentelijk meer het accent legt op secundair beroepsonderwijs als eindonderwijs en dat doorstroming naar ander onderwijs meer in termen van stapelen en omwegen en daarmee gepaard gaande kosten wordt geanalyseerd. De

Boer en Hoebe rapporteren in dit themanummer vanuit onderscheiden invalshoeken over extern rendement.

NIEUW BELEID VOOR BEROEPSONDERWIJS

Sinds 1991/92 staat nieuw overheidsbeleid in de steigers om studenten van het beroepsonderwijs beter op de arbeidsmarkt voor te bereiden. Dit nieuw beleid is in eerste instantie door middel van convenanten onderschreven door de sociale partners en door de instellingen voor beroepsonderwijs. De overheid trekt zich terug van haar traditionele bemoeienissen met gedetailleerde procesvoorschriften: zij dereguleert en vergroot de autonomie van de onderwijsinstellingen. Het beroepsonderwijs heeft in dit opzicht een voorsprong op het voortgezet en het basisonderwijs. Daarnaast wil men komen tot een grotere coördinatie en samenwerking tussen de verschillende opleidingen in het beroepsonderwijs en de volwasseneneducatie (BVE) in de vorm van regionale opleidingscentra (ROC's). In 1995 zijn deze beleidsontwikkelingen in de Wet Educatie en Beroepsonderwijs (WEB) vastgelegd.

Met schaalvergroting en ROC-vorming wil men bereiken dat onderwijsinstellingen meer dan vroeger in staat worden gesteld om op eigen kracht aan kwaliteitszorg te doen, om een breed opleidingsaanbod te verzorgen, om leerlingen naar geëigende opleidingstrajecten door te verwijzen en om de interne en externe rendementsproblemen op te lossen. Van Batenburg en Den Boer rapporteren over het effect van de breedte van het opleidingsaanbod op het interne rendement. Met een aantal beleidsinstrumenten rond marktorientatie en met verschillende vormen van overleg en informatie-uitwisseling tussen onderwijs en bedrijfsleven denkt men primair de problemen van extern rendement aan te pakken. Beleidsinstrumenten rond kwaliteitszorg, verbreding van het aanbod en leerlingbegeleiding zijn in eerste instantie gericht op de verhoging van intern rendement, dat wil zeggen op verbetering van efficiency en effectiviteit; deze beleidsinstrumenten zijn in tweede instantie ook gericht op de verbetering van extern rendement, dat wil zeggen op een betere voorbereiding van de leerlingen op de arbeidsmarkt.

Versterking van de praktijkcomponent en een verdere dualisering van het beroepsonderwijs spelen in alle beleidsinstrumenten een rol. In navolging van de commissies Wagner (1983) en Rauwenhoff (1990) wordt de dualisering in het leerlingwezen ten voorbeeld gesteld aan de rest van het beroepsonderwijs. In eerste instantie werd naar aanleiding van de adviezen van de commissie Rauwenhoff de dualisering vastgelegd in regelgeving aangaande een minimumpercentage praktijkervaring voor het hele beroepsonderwijs. Naar aanleiding van de adviezen van de commissie Van Veen (Commissie Dualisering, 1993) wordt dualisering in de WEB gedifferentieerd en flexibel benaderd. Voor zover aan de WEB een beleidstheorie over dualisering ten grondslag ligt, kan deze – overige omstandigheden gelijkblijvend – als volgt worden samengevat. Leertrajecten die op enigerlei wijze dual zijn, bereiden schoolverlaters beter voor op de arbeidsmarkt en verbeteren zodoende het extern rendement van het beroepsonderwijs; zij versterken ook de motivatie van studenten of houden deze in stand zodat zij indirect het intern rendement verhogen.

ONDERZOEK NAAR RENDEMENT EN DUALISERING

Uit de gegevens over de werkgelegenheid van schoolverlaters uit het secundaire beroepsonderwijs valt niet zonder meer af te leiden dat zij beter op de arbeidsmarkt zouden moeten worden voorbereid. Integendeel, vergeleken met andere schoolverlaters en vergeleken met landelijke werkloosheidscijfers doen de schoolverlaters uit het secundaire beroepsonderwijs het goed op de arbeidsmarkt (Hoebe, 1993; Meesters, 1992; Van der Velden & Wieling, 1994). Het zijn vooral de laag- en ongeschoolden die problemen ondervinden op de arbeidsmarkt (Veendrick, 1993). De kwantitatieve aansluiting van het secundaire beroepsonderwijs op de arbeidsmarkt is met andere woorden goed.

Er zijn verschillende theorieën over de aansluiting van het onderwijs op de arbeidsmarkt. Voorbeelden zijn theorieën over uit onderwijsdiploma's blijkend menselijk kapitaal (Becker, 1964), over de door het niveau van diploma's bepaalde plaats in de wachtrij op de arbeidsmarkt (Thurow, 1975) of over de uit het niveau van hun opleiding blijken de relatieve trainbaarheid van kandidaten (Glebbeeck, 1988). Deze theorieën geven aanleiding tot de voorspelling dat elke kwalificatieverbetering aan de betrokken individuen meer kansen op de arbeidsmarkt geeft dan ze zonder die kwalificaties zouden hebben. Voor wat betreft de kwantitatieve aansluiting van het secundaire beroepsonderwijs op de arbeidsmarkt leiden deze theorieën zonder uitzondering tot de voorspelling van lage werkloosheidscijfers bij de gediplomeerde schoolverlaters en dus tot de voorspelling van een hoog extern rendement van het beroepsonderwijs. Deze voorspellingen zijn in overeenstemming met de feitelijke werkloosheidscijfers van gediplomeerde schoolverlaters van verschillende beroepsopleidingen. De theorieën verschillen van elkaar voor wat betreft de voorspelde collectieve voorwaarden en effecten. Voorbeelden daarvan zijn: de productiviteitsgroei die gepaard gaat met groeiend menselijk kapitaal; de verhoging van het gemiddelde opleidingsniveau in de samenleving met de daarbijbehorende verdringingseffecten op de arbeidsmarkt waardoor plaatsen in de verschillende wachtlijsten verschuiven; en de groei van het aantal functies op de arbeidsmarkt, die naschoolse training op de werkplek vereisen (Ganzeboom, 1993).

Het beleid moet dus op andere rendementsproblemen zijn gericht dan op de bestrijding van werkloosheid van schoolverlaters uit het secundaire beroepsonderwijs. De rendementsproblemen die overblijven, zijn problemen van intern rendement en – voor wat betreft het extern rendement – problemen van kwalitatieve aansluiting op de arbeidsmarkt. Het beleid lijkt dan ook vooral gericht op de oplossing van problemen van efficiency en effectiviteit en van kwalitatieve aansluiting. In het beroepsonderwijs is nog relatief weinig onderzoek naar dergelijke problemen gedaan. Onderzoek van intern rendement in termen van efficiency en effectiviteit van het beroepsonderwijs vraagt om het zoeken van oorzaken van uitvalcijfers en (vertragingen in) het behalen van een diploma.

De kwalitatieve aansluiting kan rechtstreeks worden onderzocht, als rekening wordt gehouden met niveau en kwaliteit van de door schoolverlaters gevonden arbeidsplaatsen en als wordt gekeken naar de fit tussen de geleerde specifieke beroepskwalificaties en de eisen die aan functies op de arbeidsmarkt worden gesteld. Van der Velden en Wieling (1994) hebben een begin gemaakt met dergelijk onderzoek voor wat betreft de arbeidsmarktpositie van gediplomeerde schoolverlaters. Pas als meer van dergelijk onderzoek heeft plaats gevonden, kan rechtstreeks worden gezocht naar oorzaken van en oplossingen voor een eventuele gebrekkige kwalitatieve aansluiting. Dat geldt in het bijzonder voor oplossingen voor het kwalitatieve aansluitingsprobleem die worden gezocht in dualisering en het inrichten van het beroepsonderwijs met sterke praktijkcomponenten.

Een beleidstheorie dat dualisering op een of andere manier rendementsverhoging met zich brengt, kan ook indirect worden onderzocht. In dit themanummer wordt deze beleidstheorie geconfronteerd met onderzoeksresultaten over de relatie tussen de mate van dualisering c.q. de omvang van de praktijkcomponent van opleidingen en hun rendement. Dit gebeurt vanuit verschillende invalshoeken, voor verschillende beroepsopleidingen, en voor verschillende beroepssectoren.

In het Leerlingwezen is de dualisering het meest systematisch uitgewerkt. Met een diploma van het leerlingwezen hebben mensen goede kansen op de arbeidsmarkt. De interne rendementscijfers zijn daarentegen gemiddeld het laagste van alle beroepsopleidingen, hoewel er aanzienlijke verschillen zijn tussen beroepssectoren. Harms en Den Boer gaan in op de oorzaken voor dit lage interne rendement. In het bijzonder gaan zij in op de vraag welke rol de verschillende aspecten van dualisering spelen in het door hen aangegeven complex van oorzaken.

In de Primaire Beroepsgerichte Volwasseneneducatie (PBVE) kunnen leertrajecten worden onderscheiden met en leertrajecten zonder een praktijkcomponent. Den Boer analyseert effecten van de aanwezigheid van een praktijkcomponent en van de inrichting daarvan op de kansen van de cursisten: zowel hun kansen om te slagen als hun kansen op de arbeidsmarkt.

Verschillen in dualisering van beroepsopleidingen (de verschillen tussen enerzijds opleidingen in het primaire en het voortgezette leerlingwezen en anderzijds de korte en de lange opleidingen in het MBO) blijken niet van invloed op de kansen van schoolverlaters op de arbeidsmarkt. De manier waarop werkgevers beroepsopleidingen waarderen naar hun geschiktheid voor functies blijkt daarentegen een grote invloed te hebben op kansen van schoolverlaters op de arbeidsmarkt (Hoeben, 1993). Ook deze factoren zijn van invloed op de kwalitatieve aansluiting van schoolverlaters op de arbeidsmarkt en op het extern rendement van opleidingen. Hoeben analyseert in het laatste artikel de verschillen in waardering van meer en minder gedualiseerde opleidingen in relatie tot verschillende functieniveaus en in relatie tot de selectiecriteria en zoekkanalen die werkgevers hanteren bij personeelswerving.

CONCLUSIES

In de artikelen van dit themanummer worden de gegevens en conclusies bediscussieerd in de context van het uitgevoerde onderzoek en wordt ingegaan op de consequenties ervan voor de door het beleid gelegde relaties met rendement. De getrokken conclusies worden hier samengevat om na te gaan of de gerapporteerde onderzoeken in combinatie verdergaande conclusies toelaten.

Van Batenburg en Den Boer gaan na of de breedte van het opleidingsaanbod van een school gunstige effecten heeft op het interne rendement. Zij concluderen dat de verschillen die tussen leerlingen bestaan aangaande criteria van intern rendement in overwegende mate afhankelijk zijn van de individuele verschillen tussen de leerlingen en in geringe mate van verschillen tussen scholen. Zij vinden grote effecten van de vooropleiding van de leerlingen en van de interactie daarvan met de opleidingssectoren. Zij vinden geen effecten van de breedte van het opleidingsaanbod.

Harms en Den Boer concluderen dat de oorzaken voor de hoge uitval en het lage intern rendement van het primaire leerlingwezen van velerlei aard zijn en in verschillende contexten gezocht moeten worden. Het betreft de contexten van bedrijfsleven, de beroepsopleidingen en de leerlingen zelf. In de context van het bedrijfsleven zijn de bedrijfstak, het beroep, de bedrijven, de werkgevers, de praktijkplaatsen, de praktijkbegeleiding, en de kwaliteit van de leermeester van belang. In de context van beroepsopleidingen noemen zij de kwalificatiestructuur, de organisatie van het leerlingwezen, de organisatie van de instellingen, en de kwaliteit van lesmateriaal, van begeleiding en van docenten. Bij de leerlingen zelf spelen hun achtergrondkenmerken zoals thuismilieu een rol, evenals hun persoonskenmerken zoals aanleg, hun beroeps- en opleidingsbeelden die op persoonlijke ervaringen zijn gebaseerd, hun attitude en motivatie en de kosten- en batenafwegingen die zij maken. Door het gewicht dat de context van het bedrijfsleven heeft, ligt de oplossing van problemen van efficiency en effectiviteit in het beroepsonderwijs in het algemeen en in het leerlingwezen in het bijzonder veel ingewikkelder dan in andere onderwijssectoren. In het bijzonder leggen zij de vinger op de zwakke plek dat zowel de beroepsopleidingen (landelijke organen en streekscholen) als de bedrijven het in feite aan de individuele leerlingen hebben overgelaten om de aansluiting tussen theorie en praktijk tot stand te brengen. Opleidingen en bedrijven lijken daar in dit meest gedualiseerde onderdeel van het secundaire beroepsonderwijs weinig aan te hebben gedaan.

Den Boer gaat na welke variabelen van cursisten, van opleidings(in)richting en van dualisering invloed hebben op slaagkansen van cursisten en daarmee op het intern rendement van de beroepsgerichte volwasseneneducatie. Hij constateert dat individuele variabelen als ethniciteit, werkervaring, en vooropleiding effect hebben: autochtonen met veel werkwervaring en een relatief hoge vooropleiding hebben hogere slaagkansen dan cursisten met andere kenmerken. Hij concludeert dat de aanwezigheid van een praktijkcomponent in de opleiding niet van significante invloed is op de slaagkansen, de lengte van een opleiding daarentegen wel. De korte Oriëntatie & Schakelcursussen bieden tevens hogere slaagkansen dan reguliere beroepsgerichte cursussen. Vervolgens gaat hij na welke variabelen arbeidsmarkteffecten hebben en daarmee

het extern rendement van de opleidingen beïnvloeden. Hij constateert opnieuw dat individuele variabelen effect hebben. Het betreft variabelen als sekse, ethniciteit, werkervaring, lengte van de werkloosheid en motivatie: vrouwen, allochtonen en langdurig werklozen hebben een ongunstige arbeidsmarktpositie in tegenstelling tot mannen, autochtonen en cursisten met veel werkervaring, een korte periode van werkloosheid en een hoge motivatie. Hij concludeert dat de aanwezigheid van een praktijkcomponent wel en de lengte van een opleiding niet een significant arbeidsmarkteffect hebben. Dualisering als zodanig heeft in Den Boers conclusies geen effecten op intern rendement en gunstige effecten op extern rendement. Hij heeft geen externe effecten gevonden van verschillen in inrichting van de praktijkcomponent. De lengte van een opleiding heeft gunstige effecten op intern rendement en heeft geen effecten op extern rendement van beroepsgerichte volwasseneneducatie.

Hoeben constateert dat de verschillen in waardering die werkgevers hebben voor beroepsopleidingen in hun sector, niet samenhangen met het al dan niet gedualiseerd zijn van de beroepsopleidingen, maar met hun ingeschatte kwalificatieniveau. Min of meer onafhankelijk van het niveau van de functies (van simpel uitvoerend tot specialistisch niveau) vinden werkgevers de lange opleidingen van het MBO het beste en de korte opleidingen het slechtste aansluiten bij de eisen die zij aan functies stellen; de waardering voor gedualiseerde opleidingen ligt in het algemeen tussen deze twee uitersten; alleen in de sectoren van de bouw en de metaal wint het voortgezette leerlingwezen het van de lange MTO-opleiding voor wat betreft aansluiting bij de lagere functieniveaus; alleen in de administratieve sector wint de korte MEAO-opleiding het van de ECABO-opleiding in het primaire leerlingwezen. Tussen hun waarderingen voor opleidingen en de specifieke kwalificatie-eisen die werkgevers aan functies stellen, bleken geen of slechts lage correlaties te bestaan, waarin bovendien geen patroon te ontdekken was. Bij de werving van hun personeel geven werkgevers vaak de voorkeur aan werkervaring boven de opleiding van kandidaten. Zij zoeken voor hun vacante functies daarentegen het minste langs kanalen die hen de meest betrouwbare en gedetailleerde informatie geven over werkervaring, zoals personeel dat in het bedrijf een andere functie bekleedt, een leerovereenkomst heeft, stage loopt of via een uitzendbureau werkzaam is. Hoewel veel nog onduidelijk is, laten werkgevers zich bij hun werving en selectie van personeel zeker niet leiden door een voorkeur voor gedualiseerde opleidingen.

DISCUSSIE

In de conclusies van het in dit themanummer gerapporteerde onderzoek wordt duidelijk dat verdergaande dualisering van de opleidingen de verschillende rendementsproblemen van het beroepsonderwijs niet zal oplossen. Een beleidstheorie dat dualisering bijdraagt aan rendementsverhoging krijgt weinig ondersteuning. Theorieën die het rendement verklaren vanuit kenmerken van de arbeidsmarkt, krijgen wel ondersteuning.

In het MBO blijken verschillen tussen scholen nauwelijks en individuele verschillen tussen leerlingen in overwegende mate verantwoordelijk te zijn voor de schooluitval en het tempo van slagen voor een diploma; de breedte van het opleidingsaanbod voegt daar niets aan toe. In de gedualiseerde opleidingen van het leerlingwezen doen de partijen er weinig aan om de theorie- en praktijkcomponenten op elkaar af te stemmen. In de beroepsgerichte volwasseneneducatie heeft een duale opleiding geen effecten die bijdragen aan de verhoging van intern rendement, maar daarentegen wel effecten die bijdragen aan de verhoging van extern rendement. In de waardering van beroepsopleidingen door werkgevers en in hun werving en selectie van personeel kan geen steun worden gevonden voor de hypothese dat een verdergaande dualisering van beroepsonderwijs van invloed zal zijn op de arbeidsmarktpositie van schoolverlaters.

Meer nog dan het lage intern rendement van het leerlingwezen rechtvaardigt de conclusie van Harms en Den Boer dat de leerling "de enige verbindende schakel tussen school en praktijk (lijkt) te zijn", een vraagteken achter de in dit opzicht geprezen voorbeeldfunctie van het leerlingwezen. Opleidingen en bedrijven lijken weinig te hebben gedaan om de eventuele kwalita-

tieve aansluitingsproblemen tussen beroepsonderwijs en arbeidsmarkt op te lossen en met name niet wat betreft de beoogde cruciale inbreng van dualisering. De partijen lijken te hebben gehandeld alsof er geen kwalitatieve aansluitingsproblemen bestaan die tot onderwijsverbetering zouden moeten leiden; zij hebben gehandeld alsof het marktmechanisme zijn werk zal doen. Deze conclusie vindt een gedeeltelijke bevestiging in de bijdrage van Hoeben over de manier waarop werkgevers de diverse opleidingen waarderen: werkgevers lijken geen problemen te hebben met onderbenutting van de kwalificaties van hun personeel en lijken het verdringingseffect op de arbeidsmarkt actief te stimuleren.

In de beroepsgerichte volwasseneneducatie bleek dualisering evenmin effecten te hebben op het intern rendement. Dualisering bleek daarentegen wel arbeidsmarkteffecten te hebben en dus extern rendement te beïnvloeden; de wijze van inrichting van de praktijkcomponent bleek daarbij niet terzake te doen. Den Boer interpreteert dit arbeidsmarkteffect van de praktijkcomponent op zichzelf – in tegenstelling tot de wijze van inrichting – als een gebruik van de praktijkcomponent door werkgevers ten behoeve van een verlengde proefperiode van de betreffende werknemers; cursisten zonder praktijkcomponent in hun opleiding zijn zodoende in het nadeel. Dit is op zijn best een onbedoeld neveneffect van dualisering. Deze interpretatie komt op het eerste gezicht niet overeen met de gegevens uit het werkgeversonderzoek van Hoeben: bij het zoeken van nieuw personeel wordt weinig gebruik gemaakt van eventuele interne bedrijfsinformatie over mensen met een leerovereenkomst of een stage. Daar staat tegenover dat werkgevers de voorkeur geven aan nieuw personeel dat al werkervaring heeft.

De met dualisering bedoelde effecten, waarop met name door de inrichtingsvariabelen van de beroepsgerichte volwasseneneducatie (gerichte werkritme-training, ervaringsleren, aansluiting tussen theorie en praktijk, begeleiding op de praktijkplaats) wordt geanticipeerd, traden niet op. De plaats die in het beleid aan de opleidingen wordt toegekend, wordt zodoende niet waargemaakt. Ook het neveneffect van een verlengde proefperiode treedt waarschijnlijk niet zozeer op tengevolge van de beroepsopleiding maar door het functioneren van de arbeidsmarkt. Dit suggereert eens te meer dat de arbeidsmarkt mogelijk wel belangrijker is in de verklaring van extern rendement van beroepsonderwijs dan dit onderwijs zelf.

De in dit themanummer gerapporteerde onderzoeksresultaten rechtvaardigen veel van de adviezen van de Commissie Dualisering in haar rapport Beroepsvorming langs vele wegen. Het lijkt inderdaad zo te zijn dat er niet één koninklijke weg voor een adequate beroepsvoorbereiding bestaat. Het beroepsonderwijs moet de oplossing van zijn rendementsproblemen zeker niet zoeken in een bepaalde wijze en een bepaalde mate van dualisering. De commissie constateert dat de problemen van kwalitatieve aansluiting op de arbeidsmarkt reeds grotendeels zijn opgelost in de voorbije periode zonder dat het beroepsonderwijs is gedualiseerd in de strikte betekenis die daaraan in eerste instantie is gegeven (Commissie Dualisering, 1993). In die zin adviseert de Commissie de overheid om haar beleidstheorie voor het beroepsonderwijs bij te stellen. In een later aangebrachte wijziging van de WEB heeft de minister dit advies overgenomen.

Dit impliceert uiteraard niet dat een beleidstheorie die benadrukt dat beroepsonderwijs langs vele wegen effectief kan zijn, daarmee de oplossing heeft voor de problemen waarvoor het beroepsonderwijs zich ziet geplaatst. Integendeel, een dergelijke beleidstheorie zal er toe neigen afstand te nemen, het marktmechanisme zijn werk te laten doen en het beroepsonderwijs zijn eigen problemen te laten oplossen. Het mechanisme van de arbeidsmarkt zal daardoor nog belangrijker worden als verklaring voor rendementseffecten van het beroepsonderwijs.

Onderzoekers van het verschijnsel dualisering en van de effecten ervan op het interne en het externe rendement van het beroepsonderwijs doen er verstandig aan de arbeidsmarkt en de daarnaar verwijzende theorieën uitdrukkelijk in hun onderzoek te betrekken. Zij doen er eveneens verstandig aan de contextgebondenheid van de bestudeerde effecten als uitgangspunt te nemen. In een contingentie-theoretische benadering zullen zij de differentiële effecten per opleidingssector moeten verklaren. Vooralsnog lijkt de relevante sector van de arbeidsmarkt de belangrijkste contextuele of contingentie-factor (Meesters, 1992; Onstenk & Wilbrink, 1992; Van der Velden & Wieling, 1994). Zelfs de invloed van regionale verschillen binnen eenzelfde sector van de arbeidsmarkt is aangetoond (Nieuwenhuis & Mulder, 1994).

Al met al komt uit deze onderzoeksresultaten de vraag naar voren of de scholen voor beroepsonderwijs waarde toevoegen aan de invloeden van de arbeidsmarkt en van de individuele leerlingen. Bij toegevoegde waarde moet worden gedacht aan de rendementsverhoging die het gevolg zou kunnen zijn van de theoretisch denkbare verschillen in efficiency en effectiviteit van de inrichting van (praktijkcomponenten in) het beroepsonderwijs in het algemeen en gedualiseerd beroepsonderwijs in het bijzonder. Effecten lijken vooral toegeschreven te moeten worden aan het functioneren van de arbeidsmarkt (met name de vraag- en aanbodverhoudingen en de selectie van personeel) en aan de inzet, de bekwaamheden en de a-scriptieve kenmerken van de individuele leerlingen. In die zin komen de hier gerapporteerde resultaten overeen met het tot nu toe spaarzaam verrichte onderzoek naar verbetering van de kwaliteit van het beroepsonderwijs (Hoeben, Visser & Van Batenburg, 1994).

LITERATUUR

- Batenburg, Th.A. van & Haanstra, F. (1992). *Instroom en rendement van het deeltijd middelbaar beroepsonderwijs*. Groningen: RION.
- Batenburg, Th.A. van & Lokman, A.H. (1991). *Vooropleiding en rendement in het MBO*. Groningen: RION.
- Becker, G. (1964). *Human capital; a theoretical and empirical analysis, with special reference to education*. New York: National Bureau of Economic Research.
- Boer, P.R. den & Meesterberends-Harms, G.J. (1992). *Het intern rendement van het leerlingwezen, een vooronderzoek*. Groningen: RION.
- Bont-Hoek, J.G.M. de, Haas, Th.H.C. de, & Kraft, P. (1989). *Aanpak van de uitval. Resultaten van een vooronderzoek naar de uitval in het leerlingwezen*. Den Bosch: Pedagogisch Centrum Beroepsonderwijs Bedrijfsleven.
- Commissie Dualisering (1993). *Beroepsvorming langs vele wegen*. Zoetermeer: Ministerie van O&W.
- Ganzeboom, H.B.G. (1993). Gedoemd om door te leren. In *NRC Handelsblad* 21 oktober 1993, p. 5.
- Glebbeek, A.C. (1988). De arbeidsmarktpositie van opleidingen. *Tijdschrift voor arbeidsvraagstukken* 4, 75-89.
- Hoeben, W.Th.J.G. (1993). *Beroepsonderwijs en arbeidsmarkt; Een vergelijking van de arbeidsmarktpositie van schoolverlaters uit verschillende vormen van secundair beroepsonderwijs*. Groningen: RION.
- Hoeben, W.Th.J.G., Visser, E.P. & Batenburg, Th.A. van (1994). Kwaliteitsbepaling en -bevordering: onderzoek in het BVE. In B.P.M. Creemers (ed.). *Deregulering en de kwaliteit van het onderwijs* (pp. 215-241). (Monografieën onderwijsonderzoek 17). Groningen: RION.
- Meesters, M.J. (1992). *Loopbanen in het onderwijs en op de arbeidsmarkt. Vertikale en horizontale differentiatie in het voortgezet onderwijs: oorzaken en gevolgen voor de arbeidsmarktpositie van Nederlandse jongeren*. Proefschrift U.v.A., Nijmegen: ITS.
- Nieuwenhuis, A.F.M. & Mulder, R.H. (1994). Training policy at branch level: The pharmacist's assistants case. In W.J. Nijhof & M. Mulder (Eds.). *Book of summaries of the Third Conference of the International Research Network for Training and Development. Education and Training for Work*. Milan: University of Milan.
- Onstenk, J.H.A.M. & Wilbrink, B. (1992). *Arbeidsmarktrelevantie van beroepsgerichte volwasseneneducatie. De beoordeling van arbeidsmarktrelevantie in de PBVE-praktijk*. Amsterdam: SCO.
- Thurow, L. (1975). *Generating inequality*. New York: MacMillan.
- Veendrick, L. (1993). *Het loon van de last. Ongeschoolde arbeid en de veranderende identiteit van omgeschoolde jongeren*. Proefschrift RUG. Groningen: Wolters-Noordhoff.
- Velden, R.K.W. van der & Wieling, M.H. (1994). Arbeidsmarktpositie van schoolverlaters uit het voortgezet onderwijs. *Pedagogische Studiën*, 71, 5, 326-339.

Manuscript ontvangen 5-12-1994

Definitieve versie ontvangen 11-4-1996

Schoolloopbaansucces en de breedte van het opleidingsaanbod in het MBO

Th.A. van Batenburg en P.R. den Boer

GION: Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling, Rijks-universiteit Groningen

ABSTRACT

One of the goals of the government's senior secondary vocational education policy is to reduce the percentage of early school leavers. It is thought that this goal might be reached by means of integrating all schools for vocational education in one region into a large-scale 'community college'. In large-scale colleges pupils have more opportunities to choose the right course and will be able to correct a wrong choice more easily.

The data presented in this paper concern the period just before these school integration processes took place. Thus the hypothesis could be tested that more opportunities of choosing a course as a result of the scale of the school leads to a lower percentage of early school leavers. A second hypothesis could be tested concerning the size of municipality and percentage of early school leavers. The results show no relations with early school leaving for either the number of courses within a school or the size of municipalities.

PROBLEEMSTELLING

De overheid staat ten aanzien van het beroepsonderwijs een beleid voor gericht op de vorming van brede scholengemeenschappen voor beroepsonderwijs. In eerste instantie werd concentratie van de voltijd- en deeltijd-opleidingen afzonderlijk beoogd, de voltijd-opleidingen in de SVM- en deeltijdopleidingen in de WCBO-operatie. Na of eigenlijk tijdens de uitvoering van deze beide operaties werd het creëren van brede scholengemeenschappen voor beroepsonderwijs verder gestimuleerd, gericht op de vorming van Regionale Opleidings Centra (ROC's). Daarin worden alle vormen van voltijd- en deeltijd secundair beroepsonderwijs onder één dak verenigd. Dit beleid heeft inmiddels definitief zijn beslag gekregen in de Wet Educatie en Beroepsonderwijs (WEB). De achtergrond achter de vorming van brede scholengemeenschappen is in de eerste plaats een verbetering van de diensverlening aan de klant (de leerling c.q. cursist). Deze zou door een 'één-loketten-systeem' beter – meer 'op maat' – bediend kunnen worden. In de tweede plaats wordt beoogd het rendement van het beroepsonderwijs te verbeteren, zowel het interne rendement als het externe rendement. Verdergaande dualisering van (ook het voltijds) beroepsonderwijs maakt op dit punt deel uit van dit beleid gericht op brede scholengemeenschappen (zie het inleidende artikel van Hoebe in dit nummer).

De verbetering van het interne rendement werd noodzakelijk geacht vanwege de teleurstellende rendementscijfers voor zowel het Leerlingwezen als het MBO. Een van de middelen om een verbetering van het interne rendement te realiseren is het terugdringen van het voortijdig schoolverlaten (vsv). Op factoren die een rol spelen bij het voortijdig schoolverlaten in het Leerlingwezen wordt in het artikel van Harms en Den Boer (in dit nummer) nader ingegaan. Uit onderzoek in het MBO is bekend dat het merendeel van de leerlingen die dit schooltype voortijdig verlaten, dat in het eerste leerjaar doen¹ en dat ruim de helft van hen vervolgens een andere dagopleiding gaat volgen (Van Batenburg, Lokman en Nieuwenhuis 1990). Een belangrijk deel van het probleem van het voortijdig schoolverlaten heeft dus vermoedelijk te maken met een

verkeerde opleidings- en beroepskeuze. Van verbreding van het onderwijsaanbod kan voor deze groep een verbeterde opvang binnen dezelfde school worden verwacht en daarmee een daling van het aantal vsv-ers per school.

Met het creëren van brede scholengemeenschappen voor beroepsonderwijs worden twee doelstellingen beoogd. Doordat leerlingen de keuze hebben uit meer opleidingen binnen dezelfde school, kunnen in de eerste plaats verkeerde opleidingskeuzen gemakkelijker worden voorkómen en in de tweede plaats kunnen ze sneller en gemakkelijker worden gecorrigeerd. In beide gevallen zou dit voor de school minder uitval tot gevolg hebben.

In het kader van de reeds lopende fusies van scholen is onderzocht in hoeverre een breed opleidingsaanbod van invloed is op het schoolloopbaansucces van MBO-leerlingen. In de beschikbare data was het niet mogelijk onderscheid te maken tussen het eerste element (het voorkómen van uitval) en het tweede (de verbeterde opvang). Er zullen dus alleen globale uitspraken gedaan kunnen worden over de relatie tussen de breedte van het opleidingsaanbod en het schoolloopbaansucces van leerlingen.

Het voorgestane beleid, dat geoperationaliseerd is in de verhoging van het minimum aantal leerlingen per school om voor financiering in aanmerking te komen, werd deels doorkruist, doordat het in de grotere gemeenten vanwege het grotere leerlingenaanbod mogelijk bleek langer smalle, mono-sectorale scholen in stand te houden. Daarnaast echter kennen grote steden over het geheel genomen vaak een breder onderwijsaanbod dan kleinere gemeenten. De omvang van de gemeente is daarom een belangrijke storende factor in het beleid, dat gericht is op de vorming van brede scholengemeenschappen en in onderzoek naar de relatie tussen de breedte van de school en de kans op voortijdig schoolverlaten.

Uit onderzoek is bekend, dat in het MBO het schoolloopbaansucces vooral gerelateerd is aan de vooropleiding (Van Batenburg & Lokman, 1991). Wanneer het effect van andere variabelen aan de orde is, dient hiermee rekening te worden gehouden.

In dit artikel zal getracht worden de volgende vragen te beantwoorden:

1. In hoeverre wordt het schoolloopbaansucces van leerlingen beïnvloed door de breedte van het opleidingsaanbod van de MBO-school, rekening houdend met verschillen in vooropleiding bij instroom?
2. Welke rol speelt de gemeentegrootte daarbij?

ONDERZOEKSOPZET

Voor de beantwoording van de onderzoeksvragen kon gebruik gemaakt worden van het databestand van het zogenaamde PM-project (SVO-1076), een cohortstudie onder MBO-leerlingen, die in 1987 of (voor sommige opleidingen) 1988 aan hun MBO-opleiding begonnen. De onderzoeksgroep bestaat uit 6008 leerlingen van 60 scholen. De samenstelling van de onderzoeksgroep naar sector staat vermeld in tabel 1. Bij de trekking van de steekproef is gestreefd naar een representatieve steekproef van scholen en opleidingsrichtingen. Omdat de MEAO-scholen gemiddeld groter waren dan de MTO- en MDGO-scholen, zijn relatief meer MEAO-leerlingen in de steekproef vertegenwoordigd.

De variabelen in het onderzoeksmodel zijn in de vraagstelling aangegeven: het schoolloopbaansucces is de afhankelijke variabele, vooropleiding, gemeentegrootte en breedte van de instelling de onafhankelijke variabelen. Deze variabelen zijn van verschillend meetniveau: voor-

Tabel 1. Aantallen leerlingen en scholen per sector.

	MTO	MEAO	MDGO	Totaal
Scholen	22	19	19	60
Leerlingen	931	4374	703	6008

schoolniveau

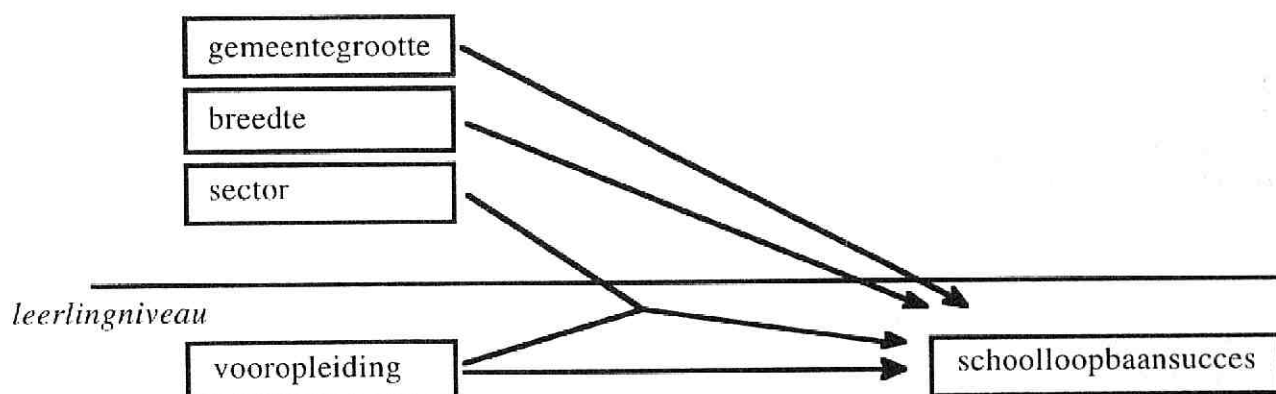


Fig. 1. Onderzoeksmodel.

opleiding en schoolloopbaansucces zijn variabelen op het niveau van de leerling, de breedte van de instellingen is een variabele op het niveau van de school evenals de gemeentegrootte, gezien het feit, dat per gemeente niet meer dan een school in het onderzoek is opgenomen. In figuur 1 wordt de onderzoeksvraag schematisch weergegeven, waarbij de verschillende meetniveaus zijn aangegeven.

Onder voortijdig schoolverlaten (vsv) wordt het verlaten van een opleiding verstaan, voordat deze met succes is afgerond. Sommige leerlingen verlaten de opleiding waaraan zij begonnen voortijdig, bijvoorbeeld vanwege een verhuizing, maar zetten die elders voort. Deze groep leerlingen wordt niet als vsv-er beschouwd, evenmin als de leerlingen die op het moment van de laatste bevraging nog met hun opleiding bezig zijn. Deze beide groepen worden in de analyses buiten beschouwing gelaten. Om het interne rendement van scholen, opleidingen, sectoren en het gehele MBO vast te kunnen stellen, wordt het schoolloopbaansucces van individuele leerlingen vastgesteld, uitgedrukt in vijf categorieën. In tabel 2 worden deze categorieën en hun frequentieverdeling voor de onderzoeksgroep weergegeven.

Uit de gegevens blijkt dat de meeste leerlingen het diploma onvertraagd halen en een deel met vertraging. Het schoolloopbaansucces van het MBO is met 54% geslaagde leerlingen laag te noemen.

De variabele schoolloopbaansucces kan worden opgevat als een rangorde variabele. Wij menen dat voortijdig schoolverlaten in leerjaar 1 'slechter' is dan in leerjaar 2: hoe later leerlingen uitvallen des te groter de kans is dat ze meer hebben geleerd. Ook menen we dat het diploma halen zonder vertraging beter is dan met vertraging. Het is evenwel ook mogelijk een andere volgorde te hanteren: zo zou men vanuit kosten-optiek redenerend kunnen stellen dat juist voortijdig schoolverlaten op een later tijdstip nadelig is. Deze kostenoptiek is echter niet gehanteerd.

Tabel 2. Schoolloopbaansucces.

	n	%
1. vsv in het eerste jaar	1291	23,3
2. vsv in het tweede jaar	640	11,6
3. vsv in het derde jaar	633	11,4
4. gediplomeerd na vertraging	851	15,4
5. gediplomeerd zonder vertraging	2122	38,3

Tabel 3. Vooropleiding.

	n	%
LBO	1013	18,3
MAVO	3824	69,1
HAVO/VWO zonder diploma	179	3,2
HAVO/VWO met diploma	521	9,4

De vooropleiding is beschreven in vier categoriën. De categorie anders/onbekend is daarbij buiten beschouwing gelaten ($n=5537$). Ze zijn weergegeven in tabel 3, evenals de frequentieverdeling in de onderzoeksgroep.

De gemeentegrootte is gedefinieerd als het inwonertal in 1987 van de gemeente waar de school staat. De inwonertallen liepen uiteen van 8.897 tot 691.738, met een gemiddelde van 184.756 en een standaarddeviatie van 191.651. In tabel 4 wordt de frequentieverdeling van de gemeentegrootte weergegeven, die uit overwegingen van presentatie in 5 categorieën is ingedeeld. In de analyses zijn de inwonertallen zelf gebruikt. Uit overwegingen van hanteerbaarheid zijn de aantallen gedeeld door 100. Deze operatie is niet van invloed op de analyses.

De variabele breedte van de instelling is uitgedrukt in het aantal opleidingsrichtingen dat een instelling aanbiedt. Onder opleidingsrichting worden de kernafdelingen en zelfstandige afdelingen verstaan, zoals: MTS-Elektrotechniek, KMBO-Techniek, MEAO-Bedrijfsadministratie, MEAO-Toerisme, MDGO-Agogisch Werk, INTAS, VHBO. De informatie over het aantal door de instelling aangeboden opleidingsrichtingen is betrokken uit de publikatie van gegevens over 'dagscholen voor MBO' van 1 september 1987 (Ministerie van O&W, 1988). De frequentieverdeling staat vermeld in tabel 5.

Vanwege het feit, dat gebruik wordt gemaakt van verschillende meetniveaus, is het gebruik van multi-niveau analyse geïndiceerd. Daarvoor is het programma VARCL gebruikt (Longford, 1986). Omdat VARCL geen ontbrekende waarden toestaat en de categorie 'anders/onbekend'

Tabel 4. Gemeentegrootte.

	n	%
Tot 50.000	15	25
50-100.000	12	20
100-200.000	20	33
200-250.000	4	7
Meer dan 400.000	9	15

Tabel 5. Breedte van het opleidingsaanbod.

Aantal aangeboden opleidingen	n	%
3	7	11,7
4	19	31,7
5	16	26,7
6	7	11,7
7	8	13,3
8	2	3,3
13	1	1,7

van de variabele vooropleiding buiten beschouwing is gelaten, zijn de analyses uitgevoerd over de 5537 leerlingen, waarvoor alle gegevens beschikbaar waren.

Bij multi-niveau analyse wordt de totale variantie opgesplitst in variantie tussen scholen en variantie tussen leerlingen. Om inzicht te verkrijgen in de hoeveel variantie die op elk niveau verklaard kan worden, is een stapsgewijze modellering gehanteerd.

1. In de eerste stap wordt een model getoetst waarin geen voorspellers zijn opgenomen: het zogenaamde lege model. Hiermee worden de percentages te verklaren variantie op de twee niveaus geschat.
2. In de tweede stap wordt het effect van de vooropleiding in het MBO geschat.
3. In de derde stap worden de interactie-effecten tussen vooropleiding en sector samen met de schoolvariabelen sector, gemeentegrootte en breedte van de opleiding gemodelleerd.

In de derde stap wordt niet de vooropleiding, maar de interacties van vooropleiding en sector als voorspellers gebruikt, omdat in eerder onderzoek is aangetoond, dat er een interactieve relatie bestaat tussen de sector en de vooropleiding en het schoolloopbaansucces in de eerste twee jaar van het MBO (Van Batenburg & Lokman, 1992). Van de vooropleiding kan daarom ook hier verwacht worden dat de relatie (regressie) tussen de vooropleiding en de afhankelijke maat (schoolloopbaansucces) per sector verschillend is.

Het in de derde stap getoetste model geeft inzicht in de effecten van de variabelen op leerlingniveau (vooropleiding) en de voor dit onderzoek van belang zijnde variabelen op schoolniveau (gemeentegrootte en breedte).

RESULTATEN

In tabel 6 worden de gegevens van de stapsgewijze modeltoetsing gepresenteerd. Uit model 1 blijkt dat van de totale variantie in de afhankelijke variabele schoolloopbaansucces 94 procent ligt op leerlingniveau en 6 procent op schoolniveau. Verschillen tussen scholen verklaren dus 6 procent van de variantie in schoolloopbaansucces. In model 2 wordt het effect van vooropleiding geschat. Uit de verhouding tussen de parameterschatting en de standaardfout (effect/standaardfout = T-waarde; de kritische T-waarde, bij een significantieniveau van 5% tweezijdig getoetst, is 1.96) blijkt dat er sprake is van een significant effect: er is een duidelijke relatie tussen vooropleiding en schoolloopbaansucces in het MBO. Door de varianties in model 2 te vergelijken met die in model 1 kan geconcludeerd worden dat de vooropleiding 6 procent ($1 - 2.29/2.44$) van de leerlingvariantie in schoolloopbaansucces verklaart. Eveneens kan geconcludeerd worden dat door invoering van vooropleiding in het model de tussenscholenvariantie is toegenomen (tot 7%).

De mate waarin een model met beter bij de data past dan een ander model kan worden nagegaan door de 'deviances' van de verschillende modellen met elkaar te vergelijken. Deze moet voldoende in waarde dalen. Dit verschil is chi-kwadraat verdeeld, waarbij het aantal vrijheidsgraden overeenkomt met het verschil in het aantal parameterschattingen. Uit de vergelijking van de deviances van model 1 en model 2 kan worden geconcludeerd dat model 2 (veel) beter bij de data past dan model 1.

Uit de parameterschattingen van model 3 kan het volgende worden afgeleid: er bestaan, zoals verwacht duidelijke verschillen tussen de sectoren in de sterkte van het verband van vooropleiding met het schoolloopbaansucces. In het MEAO is de relatie tussen vooropleiding en schoolloopbaansucces het sterkst (.63), in het MTO is het aanzienlijk minder sterk (.36) en in het MDGO nog iets minder (.24).

Uit de parameterschattingen van de variabele sector blijkt dat MEAO'ers vergeleken met MTO'ers en MDGO'ers het minst goede schoolloopbaansucces halen, de MDGO'ers het beste.

Uit de parameterschattingen voor gemeentegrootte en breedte van de opleiding blijkt dat deze effecten niet significant zijn: ten aanzien van schoolloopbaansucces maakt het niets uit of men een school bezoekt die in een grote of een kleine gemeente staat en het maakt niets uit of dit een brede of smalle school is. Met de interactievariabelen wordt 7 procent van de leerlingvarian-

Tabel 6. Modeltoetsing schoolloopbaansucces.

Model 1: leeg model			
Variantie leerlingniveau	2.44	~94%	
Variantie schoolniveau	.16	~ 6%	
Deviance = 20756			
Model 2: effect vooropleiding		Effect	Standaardfout
vooropleiding		.54	.03
Variantie leerlingniveau	2.29	~93%	
Variantie schoolniveau	.17	~ 7%	
Deviance = 20397			
Model 3: volledig model		Effect	Standaardfout
leerlingniveau	vooropleiding * MTO	.36	.07
	vooropleiding * MEAO	.63	.03
	vooropleiding * MDGO	.24	.09
sector	MTO	.00	.00
	MEAO	-.84	.21
	MDGO	.40	.24
schoolniveau	Gemeentegrootte	-.04	.03
	Breedte	.02	.04
Variantie leerlingniveau	2.28	~ 95%	
Variantie schoolniveau	.12	~ 5%	
Deviance = 20358			

tie verklaard. Met de variabele sector wordt 27 procent van de tussenscholenvariantie verklaard, dit is 2 procent van de totale variantie in schoolloopbaansucces.

CONCLUSIES

Tussen vooropleiding en schoolloopbaansucces, dat hier gebruikt is als operationalisering van intern rendement, bestaat een duidelijk verband. Daarbij moet worden opgemerkt dat de scholen in 1987 – het jaar waarin de leerlingen uit het onderzochte cohort instroomden in het MBO – gemiddeld een minder breed opleidingsaanbod hebben dan nu, gezien de fusies in het MBO. Ook de gemeentegrootte blijkt geen invloed op de studieresultaten te hebben.

De hier beschreven data hebben betrekking op de periode waarin het ROC- en WEB-beleid werd geformuleerd. Dat beleid was gericht op fusie. Daaraan lag onder andere de veronderstelling ten grondslag, dat daardoor het onderwijsaanbod verbreed zou worden, ten gevolge waarvan het interne rendement van scholen zou stijgen. Hier is aangetoond, dat aan de situatie op de scholen ten tijde van de formulering van dit beleid, geen beleidsondersteunend rendementsargument ontleend kan worden.

Daaruit kan overigens niet direct worden geconcludeerd, dat het rendementsargument daarmee geen geldigheid zou kunnen bezitten, omdat ten tijde van de formulering van het beleid nog nauwelijks echt brede scholengemeenschappen bestonden. Vermoedelijk was de spreiding in schoolbreedte overigens toen nog groter dan nu het geval zal zijn, na de fusie-operaties. Ook op basis van cohort-onderzoek op data betreffende de huidige situatie zouden dergelijke uitspraken niet kunnen worden gedaan, omdat nu vrijwel alle scholen voor beroepsonderwijs uit brede

scholengemeenschappen bestaan. Correlationeel onderzoek kan voor de beantwoording van de vraag naar de geldigheid van het rendementsargument dus geen soulaas (meer) bieden. In het gunstigste geval kan een vergelijking van de rendementscijfers toen en nu (als alle overige omstandigheden gelijk gebleven zijn) enig licht op deze vraag kunnen werpen.

Tussen vooropleiding en schoolloopbaansucces bestaat een duidelijk verband. De vooropleiding verklaart ongeveer 6 procent van de leerlingvariantie in schoolloopbaansucces. Deze samenhang verschilt per sector: het sterkst en eenduidigst is het in het MEAO, waar een hogere vooropleiding een hoger rendement impliceert. In het MTO en MDGO zijn de verbanden minder sterk. Als rekening gehouden wordt met de verschillen tussen de sectoren in de relatie tussen vooropleiding en schoolloopbaansucces wordt 7 procent variantie verklaard. In eerder onderzoek (Van Batenburg & Lokman, 1991) werd aangetoond dat LBO'ers in de eerste fase van de schoolloopbaan op het MTO niet vaker blijven zitten of voortijdig de school verlaten dan MAVO'ers. Deze bevinding kan hier, voor het schoolloopbaansucces aan het eind van de MBO-opleiding, niet worden gerepliceerd. Uit de analyse blijkt voor alle sectoren een significante relatie tussen vooropleiding en schoolloopbaansucces te bestaan, zij het niet in alle sectoren even sterk. Kennelijk lopen LBO'ers in de latere schooljaren van het MTO (het derde en vierde leerjaar) meer vertraging op dan hun medeleerlingen met een MAVO-achtergrond.

DANKBETUIGING

Dit onderzoek is mogelijk gemaakt door een subsidie van het ministerie van onderwijs.

NOOT

1. Alleen in het MDGO doet een belangrijk deel van het voortijdig schoolverlaten zich ook in het laatste jaar voor.

LITERATUUR

- Batenburg, Th.A. van, Lokman, A.J. & Nieuwenhuis, A.F.M. (1990). *School- en beroepsloopbanen van MBO-leerlingen*. Eerste rapport: de aanvang van het onderzoek. SVO-project 6059. Groningen: RION.
- Batenburg, Th.A. van & Lokman, A.H. (1991). *Vooropleiding en rendement in het MBO*. Groningen: RION.
- Batenburg, Th.A. van & Lokman, A.J. (1992). *Leerlingen, scholen en schoolloopbanen in het MBO*. De eerste twee jaar. SVO-project 9075. Groningen: RION.
- Goedhart, P.J., e.a. (1990). *MBO-Almanak 1991*. Met gegevens over het leerlingenwezen. Alphen a.d. Rijn: Samson H.D. Tjeenk Willink.
- Longford, N. (1986). *Variance component analysis: manual*. University of Lancaster.

Manuscript ontvangen 5-12-1994

Definitieve versie ontvangen 11-4-1996

Het intern rendement van het leerlingwezen

G.J. Harms en P.R. den Boer

*GION: Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling,
Rijksuniversiteit Groningen*

ABSTRACT

In Dutch vocational education there is a strong tendency nowadays to incorporate a substantial element of practical learning on the job in all courses. The apprenticeship system serves in this respect as an example for the whole field of vocational education. It is estimated, however, that in this system approximately 50% of the students do not successfully complete their courses. This article inquires how the dual character of this system possibly affects this phenomenon. This is done by reporting research on the volume and possible causes of student drop out in the apprenticeship system. The research covers a study of literature, a secondary analysis of existing data and interviews with key persons. Main factors causing drop out are the student's perception of jobs and education, the characteristics of the learning site on the job and the degree of difficulty of content matter. In particular the characteristics of the learning site on the job are discussed to see which threats a dual system in vocational education should avoid.

INLEIDING

Bij het streven van de overheid naar verdergaande dualisering van het beroepsonderwijs in Nederland dient het leerlingwezen steeds als voorbeeld van een reeds lang bestaand duaal onderwijssysteem. Tegenover de hooggespannen verwachtingen van duale systemen, die de praktijk en de theorie in de opleiding dichter bij elkaar zouden moeten brengen, en tot een betere aansluiting tussen onderwijs en arbeid moeten leiden, staan de alarmerende cijfers over uitval in het leerlingwezen. Een laag intern rendement van een duaal systeem stelt uiteraard grenzen aan het mogelijke externe rendement van zo'n systeem, en vormt derhalve een serieus probleem. In dat verband is het zinvol om na te gaan in hoeverre juist het duale karakter van het leerlingwezen bijdraagt aan het geconstateerde lage rendement. In dit artikel wordt daartoe een poging gedaan. Basis daarvoor vormt een in 1992 uitgevoerd onderzoek naar omvang en achtergronden van uitval in het leerlingwezen (Den Boer en Meesterberends-Harms, 1992). Over dit onderzoek wordt gerapporteerd. Aansluitend daaraan wordt een aantal veronderstellingen geponeerd over de relatie tussen dualisering en rendement.

Reeds lang is bekend, dat het leerlingwezen te kampen heeft met hoge uitval. Zo komen Van der Velden en Groothuis (1984) voor de periode 1964-1980 tot een schatting van gemiddeld 50% uitval. Lange tijd heeft het evenwel ontbroken aan een systematisch overzicht van het rendementsvraagstuk in het leerlingwezen.

Onder invloed van de activiteiten van de Commissie Wagner in 1982/1983 en de Commissie Rauwenhoff in 1990 heeft het leerlingwezen een belangrijke impuls gekregen. Een van de tastbare resultaten van het werk van de Commissie Wagner is een geslaagde verdubbelingsoperatie geweest. Door overheidssubsidies (de Bijdrageregeling Vakopleiding Leerlingwezen) en door arrangementen die de beperkingen van het individuele bedrijf overstijgen (opleidingsfondsen, gemeenschappelijke opleidingsactiviteiten) – naast natuurlijk het aantrekken van de arbeidsmarkt in de tweede helft van de jaren '80 – is de instroom in het leerlingwezen toegenomen van ca. 25000 in 1982 tot ruim 55000 in 1990.¹ Door de Commissie Rauwenhoff is de duale opleidingsweg van het leerlingwezen ten voorbeeld gesteld aan alle vormen van beroepsvoorbereiding in het Nederlandse onderwijs². Daarnaast is het kwalificatieniveau van de primaire

opleidingen van het leerlingwezen aangemerkt als het basisniveau, de startkwalificatie waarmee elke schoolverlater minimaal de arbeidsmarkt zou moeten betreden. In dat licht is systematische informatie over de uitvalsproblematiek zeer relevant.

Het hier te rapporteren onderzoek is bedoeld als een vooronderzoek voor een systematisch empirisch onderzoek. Op basis van analyse van bestaand materiaal, aangevuld met interviews met een beperkt aantal sleutelfiguren, is gepoogd een voorlopig antwoord te formuleren op (onder andere) de volgende twee onderzoeksvragen:

- Hoe groot is het rendement van en hoe omvangrijk de voortijdige uitval in het leerlingwezen in het algemeen en in de afzonderlijke leerlingstelsels in het bijzonder?
- Welke factoren zijn van invloed op rendement en uitval in het leerlingwezen?

Bij de discussie over rendement van het onderwijs wordt veelal een onderscheid gemaakt in intern en extern rendement. Intern rendement verwijst naar de mate waarin en eventueel de snelheid waarmee leerlingen de opleidingen afronden, extern rendement verwijst naar de mate waarin leerlingen, na het afsluiten van een opleiding kans zien om een aansluitende baan te vinden. Voor alle duidelijkheid zij erop gewezen dat het hier gaat om het **interne, numerieke rendement** (cf. Janssen, 1990). Het numerieke rendement van het leerlingwezen is te definiëren als het percentage leerlingen van een bepaalde instroomjaargang dat (binnen redelijke termijn) de opleiding met een volledig diploma afsluit. Uitval is daarvan het complement, het percentage leerlingen van die instroomjaargang dat de opleiding verlaat zonder een volledig diploma gehaald te hebben of daarbij zoveel vertraging heeft opgelopen, dat het diploma niet binnen de gestelde termijn is behaald. Deze definities hebben pas vanaf 1989, op basis van werk van het toenmalige Pedagogisch Centrum Bedrijfsleven Beroepsonderwijs (De Bont-Hoek e.a., 1989) en de Arbeidsvoorziening (Van Gaalen-Oordijk en Westerhuis, 1988), ruime ingang gevonden.

De uitval- en rendementsproblematiek van het leerlingwezen heeft twee aspecten, die met het veel gehanteerde onderscheid tussen doelmatigheid en rechtvaardigheid getypeerd kunnen worden. In veel gevallen staat het doelmatigheidsaspect voorop. Uitval wordt dan vooral problematisch gevonden omdat ze tot verspilling van middelen leidt. Daarnaast hebben rendement en uitval ook een rechtvaardigheidsaspect. Als het niveau van de primaire opleidingen als minimale startkwalificatie beschouwd moet worden, dan is uitval problematisch omdat grote groepen schoolverlaters de kans op een goede start op de arbeidsmarkt onthouden wordt.

Met betrekking tot de factoren die van invloed zijn op rendement en uitval bieden de registratieprocedures van het leerlingwezen – die van oudsher voorzien in het bijhouden van de ‘reden van afvoering’ – onvoldoende informatie. In wezen gaat het hierbij namelijk om een administratieve verantwoording van het uitschrijven van een leerling, met categorieën als ‘schoolverzuim’ en ‘verlies van arbeidsplaats’. Over de oorzaken van succes en falen tijdens de opleiding zegt dat weinig.

Inzicht in deze oorzaken, en vooral in hun onderlinge gewicht als factor ter verklaring van uitval en hun onderlinge samenhang en interactie, vergt systematisch onderzoek volgens het stramien van het schooleffectiviteitsonderzoek uit andere delen van ons onderwijssysteem. Daarbij worden kenmerken van de leerling en diens loopbaan in het leerlingwezen gerelateerd aan kenmerken van de schoolcomponent, de praktijkcomponent en de afstemming tussen beide. Een dergelijke opzet gaat het bestek van het hier te rapporteren onderzoek te buiten. Dat bestek laat slechts gebruik van secundaire bronnen, aangevuld met een beperkt aantal interviews, toe. Op basis van eerder gedaan onderzoek is het echter wel mogelijk in dit artikel een overzicht te geven van de beschikbare kennis over de mogelijke oorzaken, het mogelijke onderlinge gewicht en de mogelijke onderlinge samenhang daartussen. Een dergelijk overzicht vervult de functie van onderzoeksmodel voor vervolgonderzoek.

ONTWERP EN METHODEN

Het onderzoek kende drie onderdelen: een literatuurstudie, secundaire analyse van beschikbare data en een beperkt aantal interviews met landelijke organen en koepels.

De literatuurstudie heeft zich geconcentreerd op Nederlandse literatuur. De belangrijkste categorie betrof de meest directe informatie: rapportages vanuit de landelijke organen voor het leerlingwezen of vanuit de bedrijfstak over de rendementsproblematiek in hun eigen sector. De literatuur is geanalyseerd op relevante informatie met betrekking tot de onderzoeksvragen.

Voor de secundaire analyse van rendementsgegevens is in de eerste plaats een beroep gedaan op de landelijke organen voor het leerlingwezen om cohortsgewijze data over in- en uitstroom van leerlingen beschikbaar te stellen. Dit leverde onvoldoende respons op. Daarom is in de tweede plaats gebruik gemaakt van beschikbare in- en uitstroomgegevens, die jaarlijks worden opgeleverd door het CBS (uitstroomgegevens) en in de Schoolverlatersbrief (instroomgegevens, afkomstig van het ministerie van O&W). Er is daarbij evenwel geen sprake van directe samenhang in de data die met elkaar vergeleken worden; in hoeverre de leerlingen die uitgestroomd zijn in jaar X dezelfde zijn die een aantal jaren daarvoor zijn ingestroomd is uit de gegevens niet te herleiden. Er is bij de op deze data toegepaste berekeningswijze daarom eerder sprake van een schatting van het rendement op basis van de beschikbare in- en uitstroomgegevens. De schatting gebeurt op basis van de volgende definitie:

$$\text{geschat rendement instroomjaar X} = \frac{\text{Aantal geslaagden in jaar (X+Y)}}{\text{Aantal ingestroomden in jaar X}}$$

waarbij:

X = een willekeurig jaar,

Y = de reguliere lengte van de opleiding

Aan deze schattingsmethode kleven twee in het oog springende nadelen. Ten eerste is de schatting onnauwkeurig, doordat die gevoelig is voor fluctuaties in de instroom, geen rekening houdt met verschillen in reguliere opleidingsduur en een oneindige marge van vertraging toestaat. Ten tweede zijn de criteria voor rendement zeer strikt, doordat het verbreken van een leerovereenkomst als uitval wordt gezien terwijl dat de facto niet altijd het geval is, en doordat als criterium voor rendement het behalen van een volledig diploma wordt gehanteerd, waardoor deelresultaten als een praktijkgetuigschrift of deelcertificaten niet meetellen voor de rendementsbepaling. Bij gebrek aan nauwkeuriger registratie en longitudinale data is deze methode evenwel de enig uitvoerbare. Het behoeft geen betoog, dat invoering van de nieuwe methode van rendementsbepaling op basis van cohortgegevens hoge prioriteit heeft, zowel vanuit oogpunt van verantwoord onderzoek als vanuit de wenselijkheid van een verantwoorde aanpak van de uitval. Ook over de criteria voor rendement zal meer duidelijkheid moeten komen, zeker als de overheid haar plannen voor rendementsfinanciering wil uitbreiden richting leerlingwezen.

Er zijn negen interviews gehouden met velddeskundigen. Doel van de interviews was ten eerste een controle op de volledigheid en juistheid van de resultaten uit beide voorafgaande onderdelen. Ten tweede behelsde het interview een globale inventarisatie van huidig, toekomstig en gewenst beleid op het terrein van rendement en uitval. Gesprekspartners waren vertegenwoordigers van zes (bedrijfstakgebonden) landelijke organen voor het leerlingwezen, twee landelijke organisaties van scholen in het beroepsonderwijs en de landelijke organisatie van regionale organen voor het leerlingwezen.

RESULTATEN

Uitval en rendement: de cijfers

Op basis van de zojuist geschetste methode is voor de instroomjaren 1986, 1987 en 1988 een gemiddeld rendementspercentage geschat over alle primaire opleidingen van ca. 50%. De op dezelfde wijze geschatte rendementscijfers per landelijk orgaan zijn weergegeven in tabel 1.

Tabel 1. Rendement primaire opleidingen¹, per landelijk orgaan, instroomjaar 1986 t/m 1988; bron: Schoolverlatersbrief (instroomgegevens), CBS (uitstroomgegevens).

Landelijk Orgaan	1986	1987	1988	gemiddeld
Kappers (SVK)	.69	.70	.88	.76
Grafische techniek (GOC)	.75	.72	.71	.73
Wegvervoer (SLW)	.66	.55	.68	.64
Verzorging (OVDB)	.65	.60	.61	.62
Schilderen (SVS)	.59	.61	.59	.60
Motorvoertuigentechniek (INNOVAM)	.54	.56	.69	.59
Textielindustrie (BETEX)	.60	.36	.77	.59
Consumptieve techniek (SVEB, VBOB, SVH, LSBL, SVO)	.56	.51	.58	.55
Bouw (SVB)	.49	.57	.54	.53
Metaaltechniek (SOM)	.48	.57	.50	.52
Houtbew. (SH&M)	.53	.53	.51	.52
Wegenbouw (SBW)	.51	.43	.55	.50
Procestechniek (Vapro)	.64	.37	.45	.48
Carrosseriebedrijf (VOC-car)	.54	.48	.34	.45
Confectie (VOC-con)	.53	.34	.41	.42
Installatie (SOI)	.39	.49	.39	.42
Elektrotechniek (VEV)	.51	.35	.20	.34
Detailhandel (OVD)	.31	.30	.30	.30

¹ Omdat voor de voortgezette opleidingen niet beschikt kon worden over aantallen nieuw afgesloten leerovereenkomsten per jaar, en gebruikmaking van totale aantallen afgesloten leerovereenkomsten aan het eind van het jaar tot grote onnauwkeurigheden in de schattingen zou leiden, is afgezien van het bepalen van rendementscijfers voor de voortgezette opleidingen.

Er komen duidelijke verschillen in de schattingen van het rendement naar voren, zowel tussen de landelijke organen onderling als per landelijk orgaan tussen de verschillende instroomjaren. De opleidingen met het hoogste geschatte rendement zijn die in de kappersbranche en de grafische techniek, met respectievelijk 76% en 73% rendement. Daarmee is de geschatte uitval uit deze opleidingen overigens nog steeds 24% en 27%, resp. 635 en 349 leerlingen gemiddeld per instroomcohort. De opleidingen met het laagste geschatte rendement zijn die in de elektrotechniek en detailhandel, met respectievelijk 34% en 30%, en een bijbehorende geschatte uitval van resp 66% en 70%.

Uitval en rendement: de oorzaken

Literatuurstudie en interviews hebben een groot aantal factoren opgeleverd die van invloed kunnen zijn op voortijdig schoolverlaten. Ze zijn onder te brengen in drie hoofdcategorieën: factoren op het niveau van het beroep, de bedrijfstak of het bedrijf, het niveau van de leerling en het niveau van de opleiding. We vatten ze hieronder samen.

Centrale factoren op het niveau van de leerling zijn de motivatie en keuze voor het beroep en het beeld dat de leerling heeft van het beroep en van de opleiding. Leerlingen blijken vaak een onduidelijk, een onjuist of slechts een partieel beeld te hebben van wat het beroep inhoudt, onder andere ten gevolge van soms misleidende P.R.-campagnes van bedrijfstakken en voorlichting van slecht op de hoogte zijnde decanen uit de toeleverende opleidingen. Hetzelfde geldt voor de opleiding: leerlingen realiseren zich vaak niet, dat de praktijksituatie ook onderdeel van de opleiding uitmaakt en hebben een beperkt of onjuist beeld van het niveau, de inhoud, de vereiste specifieke vaardigheden en de vereiste tijdsinvestering voor de theorie-opleiding. Ook de keuze voor de opleiding gebeurt vaak op niet inhoudelijke gronden: uit negatieve overwegingen (als je niets anders meer weet ..), uit financiële overwegingen, als verplichting van de kant van de werkgever, etc. Op de achtergrond spelen factoren mee als de intellectuele capaciteit, de

taalvaardigheid, de steun van thuis, randvoorwaarden zoals kinderopvang en de combinatie van taken in werk, opleiding en gezin.

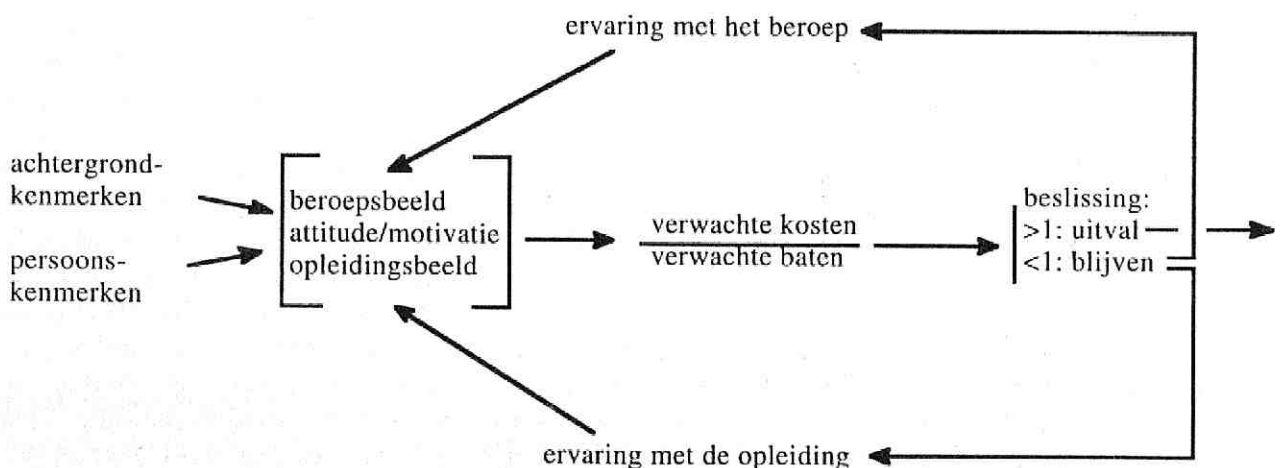
Centrale factor op het niveau van het beroep, de bedrijfstak en het bedrijf is vooral de vormgeving van de praktijkplaats: het niveau, de inhoud, de afwisseling en de leerzaamheid van het werk in de praktijkcomponent van de opleiding. Op de praktijkcomponent heeft de werkgever invloed door een stimulerende houding (loopbaanperspectief, stimuleren afronden opleiding) en de leermeester door het geven van voldoende goede ondersteuning. De consulent van het landelijk orgaan heeft een vergelijkbare invloed. Ook kenmerken van het beroep zelf spelen een rol (de aard van het werk en de specifieke eisen die dat stelt, de arbeidsomstandigheden, het loon) evenals kenmerken van de bedrijfstak (grootte van de bedrijven, met als gevolg meer of minder contact en ruimte voor begeleiding) en de werkgevers als collectief (opleidingstraditie in de bedrijfstak, de rol van werknemers en (het ontbreken van) beloning voor werkgevers voor het afronden van opleidingen).

Op het niveau van de opleiding is de centrale factor: de moeilijkheidsgraad van de theorie. Die is vaak niet in overeenstemming met de mogelijkheden van de leerlingen. Het niveau van de instromende leerlingen is, onder andere onder invloed van de verdubbelingsoperatie en teruglopende leerlingaantallen, gedaald. Tegelijkertijd hebben werkgevers vooral in de technische richtingen soms de neiging op theoretisch niveau te overvragen. Daarnaast is de aansluiting van de theorie op de praktijk een belangrijke factor. Deze problemen stellen extra hoge eisen aan de didactische kwaliteiten van de leerkracht, waaraan het, zeker gezien de heterogeniteit van de instroom nogal eens schort, aan de kwaliteit van het lesmateriaal en aan de kwaliteit van de begeleiding.

Op basis van de gevonden factoren voor uitval en met behulp van literatuur over een viertal theoretische modellen ter verklaring van keuzegedrag (Pryor, 1990; Wong en Siegerist, 1987; Rubenson, 1980; Lokman & Nieuwenhuis, 1990), is een model geconstrueerd, dat de invloed van de diverse factoren weergeeft op de beslissing van een leerling om een opleiding voort te zetten of af te breken. Het model kent drie niveau's, zoals ook boven reeds onderscheiden: het niveau van de leerling, het niveau van het beroep (en alles wat daarbij hoort) en het niveau van de opleiding. Centraal in het model staat het leerlingniveau, met daarin als centraal proces de kosten-baten afweging die de leerling maakt en die leidt tot de beslissing om de opleiding al dan niet te blijven volgen. Schema 1 geeft dit niveau weer.

Leidend naar de kosten-baten afweging zijn op het leerlingniveau vooral van belang: de attitude van de leerling ten opzichte van opleiding en beroep en zijn/haar verwachtingen daaromtrent (aangeduid als beroepsbeeld en opleidingsbeeld).

Leerlingniveau



Schema 1. Beslissmodel op leerlingniveau: uitval of blijven.

In termen van een kosten-baten afweging bepaalt de attitude (de aard van de keuze voor beroep en opleiding) wat kosten zijn en wat baten. We adstrueren dat aan de hand van een aantal voorbeelden. Voor de leerling die kiest voor een opleiding, omdat het daarmee te verwerven beroep hem/haar trekt, zijn de verwachtingen omtrent de opleiding te beschouwen als verwachte kosten. Voor deze leerling zijn de verwachtingen omtrent het beroep op te vatten als de verwachte baten (inhoud van het werk, status, arbeidsomstandigheden, loon, verworven arbeidsmarktpositie). Voor de leerling die een opleiding volgt, omdat de werkgever hem/haar daartoe verplicht bestaan alleen kosten: alle opleidingsactiviteiten zijn kosten, waar voor deze leerling geen baten in termen van kenmerken van het beroep tegenover staan. De enige (verwachte) baten voor deze leerling is het behoud van zijn/haar baan. Voor de leerling die kiest voor een opleiding omdat hij/zij de opleiding interessant vindt (leergerichte attitude), zijn de opleidingsactiviteiten tegelijkertijd kosten en baten: de te leveren inspanningen zijn op te vatten als kosten, de inhoud van de opleiding als baten. Aan de kant van het beroep liggen voor deze leerling dus geen baten.

Op basis van het beslissingsmodel kan de keuze van de leerling voor het al dan niet volgen van een opleiding aldus worden beschreven als afweging van verwachte kosten en verwachte baten. Tijdens de opleiding echter wordt de leerling geconfronteerd met de feitelijke kosten en baten: de feitelijke kenmerken van opleiding en beroep. (Daarmee worden de beide andere niveau's in het model reeds in de overwegingen betrokken; we komen daar dadelijk op terug). Door de ervaring die de leerling heeft met opleiding en beroep worden het beeld van opleiding en beroep en mogelijk ook de attitude beïnvloed, hetgeen leidt tot een bijstelling van de afweging van kosten en baten. De feitelijke kosten en baten, waarmee de leerling tijdens de opleiding geconfronteerd wordt, vormen de terugkoppelingsvariabelen in het beslissingsmodel: op basis van door de leerling verkregen informatie over de feitelijke kosten en baten, worden de verwachtingen omtrent kosten en baten bijgesteld.

Op leerlingniveau worden de attitude van de leerlingen en hun verwachtingen rond opleiding en beroep uiteraard beïnvloed door een aantal factoren, te weten:

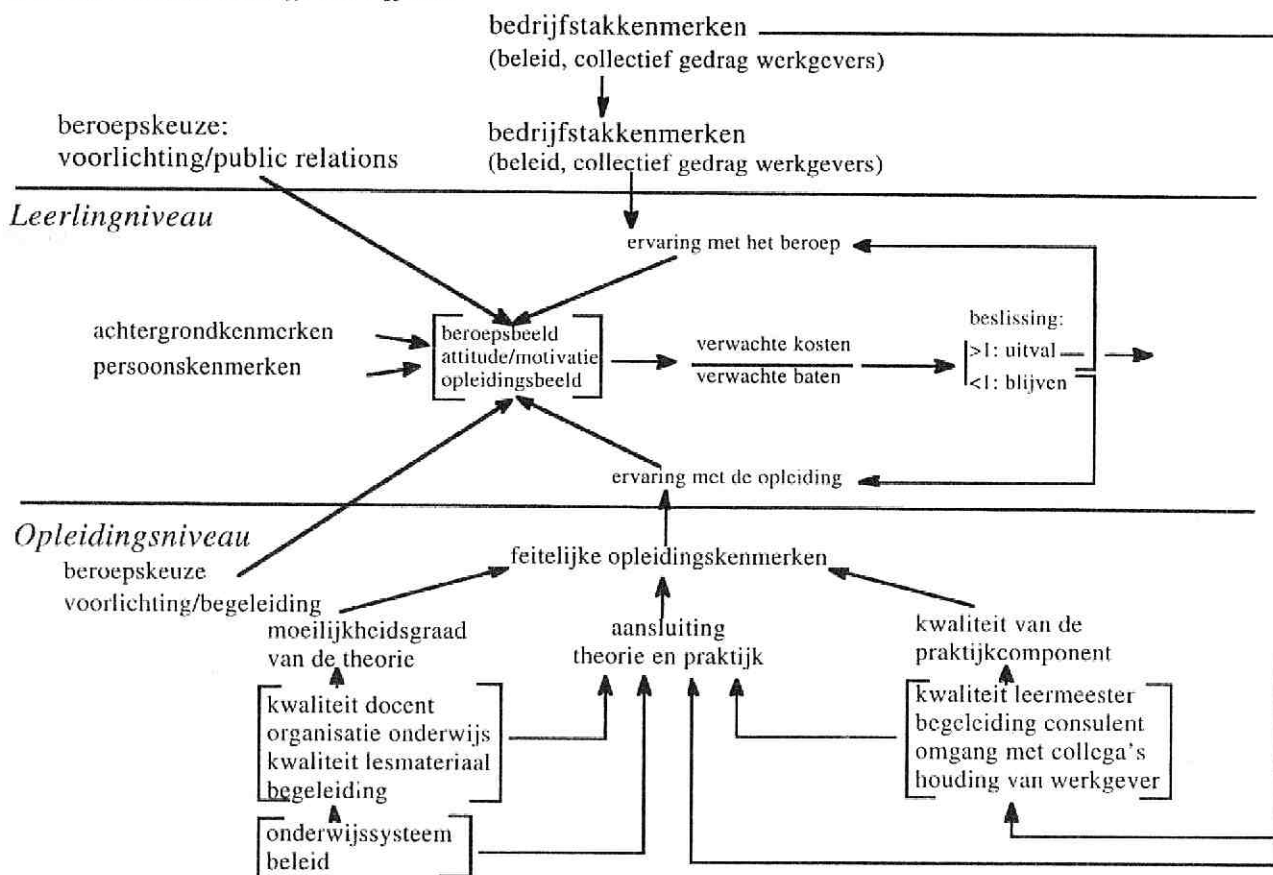
- persoonskenmerken van henzelf: intellectuele capaciteiten, specifieke aanleg, taalvaardigheid en het niveau van de vooropleiding beïnvloeden de (verwachte) inspanning die de opleiding vraagt;
- kenmerken van de achtergrond van de leerling: stimulans van thuis voor het volgen van een opleiding beïnvloedt de attitude en mogelijk ook de inspanning die de opleiding vraagt; de huiselijke situatie (gecombineerde taken in gezin werk en opleiding) hebben invloed op de te leveren inspanning, evenals de randvoorwaarden, zoals kinderopvang.

Daarmee is het model op leerlingniveau beschreven. En passant is de invloed van de beide andere niveau's reeds genoemd: zij vormen de informatiebronnen op basis waarvan de leerling de feitelijke kosten en baten vaststelt, het beeld van opleiding en beroep corrigeert en de voor de toekomst te verwachten kosten en baten heroverweegt. Daarnaast worden de verwachtingen omtrent opleiding en beroep beïnvloed door de voorlichtingsactiviteiten van scholen, regionaal consulenten, decanen van toeleverende scholen, bedrijfstakken en bedrijven. In schema 2 is het volledige model weergegeven.

De feitelijke kenmerken van de opleiding betreffen zowel de theorie als de praktijk, maar ook de samenhang daartussen. Deze kenmerken zijn boven besproken.

De feitelijke kenmerken van het beroep zijn ook boven besproken: de status van het beroep, de aard van het werk, het loon, de arbeidsomstandigheden en de conjunctuur in de bedrijfstak c.q. de spanning op de arbeidsmarkt. Deze factoren bepalen de feitelijke baten voor de leerling die een opleiding volgt om een (betere) baan of arbeidsmarktpositie te verwerven. De feitelijke kenmerken van het beroep worden op hun beurt weer beïnvloed door de kenmerken van de bedrijfstak en door het beleid van de werkgevers, of anders gezegd: het collectieve gedrag van werkgevers.

De ervaring die de leerling in de praktijk opdoet met de feitelijke kenmerken van het beroep wordt echter niet alleen bepaald door deze feitelijke kenmerken, maar tevens door de vormgeving

Niveau beroep/bedrijf/bedrijfstak

Schema 2. Volledig beslismodel.

van de praktijkcomponent. Bij een beperkt ingevulde praktijkcomponent en het ontbreken van mogelijkheden voor de leerling om via collega's kennis te maken met de feitelijke kenmerken van het beroep, zal de leerling aanzienlijk minder ervaring met het feitelijke beroep opdoen, dan bij een hoog realiteitsgehalte van de praktijkcomponent en gemakkelijke omgang met collega's.

DISCUSSIE: DUALISERING EN UITVAL

Over het begrip dualisering bestaat, zoals het geval is met de meeste termen die meer of minder recent in het beroepsonderwijs geïntroduceerd zijn, de nodige onduidelijkheid (Hövels, Frietman en Van Geest, 1992). Minimaal wordt onder dual onderwijs een vorm verstaan waarin naast leren op school, leren in de praktijk een wezenlijk onderdeel uitmaakt. Of dat leren in de praktijk plaats moet vinden in het kader van een al of niet reguliere en dus betaalde arbeidsplaats, staat ter discussie. Voorstelbaar is ook, dat de praktijkcomponent vorm krijgt binnen de kaders van een bedrijfsschool of zelfs in de vorm van een door de school georganiseerde praktijksimulatie. De tendens in de discussie van de afgelopen jaren was om de vormgeving via een gecombineerde leer-arbeidsovereenkomst, zoals gebruikelijk in het leerlingwezen, in te voeren in het gehele beroepsonderwijs. De Commissie van Veen, die in 1993 advies uitbracht over dualisering, laat daarentegen diverse mogelijkheden toe en bepleit beroepsvorming die op verschillende wijzen gestalte kan krijgen en op diverse wijzen in een overeenkomst kan worden vastgelegd. Beroepsvorming heeft per definitie een buitenschools karakter, en maakt naar de mening van de Commissie een substantieel deel van elke beroepsopleiding uit, ter grootte van minimaal circa 20%. (Beroepsvorming, 1993).

In het leerlingwezen, waar het gerapporteerde onderzoek zich afspeelt, is overwegend sprake van leerlingen met een, al dan niet betaalde, leer/arbeidsplaats in een bedrijf, hoewel er ook leerlingen zijn, die (nog) geen leerplaats hebben weten te bemachtigen (Ganga, 1992). De kanttekeningen, die vanuit de onderzoeksresultaten in deze paragraaf bij dualisering geplaatst zullen worden, zijn dan ook gebaseerd op deze situatie in het leerlingwezen.

Zoals eerder vermeld, kunnen diverse factoren van invloed zijn op uitval in het leerlingwezen. We zullen de factoren die specifiek te maken hebben met het duale karakter van het leerlingwezen nader beschouwen. De meeste aandacht gaat daarbij uit naar de factoren op het micro-niveau van de primaire processen in het onderwijs; deze zullen we ook het eerst bespreken.

Kenmerkend voor een duaal systeem is de aanwezigheid van een praktijkcomponent naast de schoolcomponent; cruciaal is de afstemming tussen de praktijk en de theorie. Welke aspecten in de praktijkcomponent dragen bij aan uitval?

Vormgeving van de praktijkcomponent

Arbeidsomstandigheden in het praktijkdeel van de opleiding kunnen aanleiding zijn tot uitval. Uit de literatuurstudie en de interviews bleken de volgende elementen van belang te zijn: werktempo, werkdruk, onregelmatige werktijden, beloning en omgang met collega's. Daarnaast is de inhoud van het werk van belang. Simpele werkzaamheden, te weinig afwisselend en te eenvoudig werk, het uitvoeren van niet leerzame klusjes en een te beperkte inhoud van het werk leiden tot verlies van interesse voor het vak. Ook de begeleiding op de werkplek kan aanleiding zijn voor uitval: waar die tekort schiet of eerder kritiserend dan ondersteunend is, hebben leerlingen de neiging om af te haken. Om de uitval op dit onderdeel tegen te gaan zal men dualisering met de nodige waarborgen moeten omgeven. Deze moeten allereerst worden gezocht in een systeem van kwaliteitsbewaking voor leerbedrijven. De arbeidsomstandigheden moeten in overeenstemming zijn met de primaire functie die de werkplek voor de leerling heeft: opleiden. Datzelfde geldt voor de inhoud van het werk. Tevens moet het bedrijf in staat zijn tot een goede begeleiding van de leerling. Landelijke organen zouden op deze punten criteria moeten ontwikkelen voor de kwaliteit van leerbedrijven. Leerbedrijven die aan de criteria voldoen zouden dan gecertificeerd moeten worden.

Geconstateerd kan worden, dat de aanwezigheid van een praktijkcomponent in een opleiding een aantal extra risico's voor uitval met zich meebrengt. In een duale opleiding wordt de leerling, in vergelijking met de situatie in een voltijd opleiding, eerder, in grotere omvang en in bredere zin, geconfronteerd met alle aspecten die te maken hebben met beroep en bedrijf. Deze vormen dan ook al tijdens de opleiding een bron van uitval, waar dat anders pas na het afronden van de opleiding een struikelblok vormt. In die zin kan de stelling verdedigd worden, dat een duaal systeem een efficiëntere route biedt dan full-time onderwijs: een onjuiste beroepskeuze wordt eerder gecorrigeerd. Om deze stelling nader te onderbouwen zou nagegaan moeten worden wat er met de uitvallers gebeurt die afhaken op basis van beroeps- of bedrijfsgebonden factoren. Vinden zij een opleidingsplaats in een andere branche of zijn zij definitief voor het onderwijs verloren? In het laatste geval wordt het belang van een goede begeleiding bij de beroepskeuze, nog sterker dan in het eerste geval, een zeer belangrijk middel om te voorkomen dat mensen zonder startkwalificatie uit het onderwijs verdwijnen. In een niet gedualiseerd systeem kan het later blijken van een verkeerde beroepskeuze in ieder geval nog het voordeel hebben dat men niet ongekwalficeerd het onderwijs verlaat. Men kan dan eventueel gemakkelijker uitwijken naar een ander bedrijf (bedrijfsgebonden factoren) of eventueel naar een ander, verwant, beroep.

De afstemming tussen theorie en praktijk

Uit de literatuur en de interviews bleken drie elementen die uitval bevorderen hier mogelijk een rol te spelen. Ten eerste wordt gesignaleerd, dat er onvoldoende integratie is van algemeen vormende vakken met zowel vaktheorie als vakpraktijk. Tevens is de inhoud van de algemeen vormende vakken onvoldoende toegespitst op de situatie van de leerling in de praktijk. Deze

klachten komen deels ook van leerlingen zelf. Het lijkt zinvol om in een duaal systeem dergelijke informatie van leerlingen serieus te nemen en zo mogelijk zelfs te systematiseren. Een en ander kan bijdragen aan een soepele werking van een dergelijk systeem. Ten tweede is de aansluiting tussen de theorie en de technieken die in de praktijk gebruikt worden, te globaal. Zo gebruikt men in het werk vaak meer gespecialiseerde technieken dan op school aangeleerd worden. Deze constatering maakt het wenselijk, dat ook de praktijkplaats meer gezien wordt als opleidingsplek. Als een gespecialiseerde techniek niet in de school aan de orde kan komen, dan kan hij wellicht beter in de praktijk ingeoeft worden. Ook loopt de theorie vaak achter bij de laatste ontwikkelingen in de beroepspraktijk. Er is permanente afstemming via beroepsprofielen nodig. Bij duale systemen signaleren leerlingen dit soort discrepanties snel en zij plaatsen van daaruit (terechte) vraagtekens bij het nut van het theoretisch deel van de opleiding. Ten derde ontbreekt een wisselwerking tussen school en praktijk. De klachten in dat verband hebben betrekking op gebrekkig lesmateriaal dat voor de verbinding zou moeten zorgen en op onvoldoende begeleiding en communicatie met begeleiders i.c. de consulent. Ook zouden didactische kwaliteiten, die leraren in het beroepsonderwijs in hun opleiding verwerven, onvoldoende gericht zijn op het integreren van leerprocessen in school en bedrijf. Op basis van deze constatering, die gebaseerd zijn op literatuurstudie en uitspraken van sleutelfiguren, dringt de gedachte zich op, dat de leerling de duale opleiding vaak niet als een eenheid ziet. Sterker nog: de leerling zelf lijkt in veel gevallen de enige verbindende schakel tussen school en praktijk te zijn en moet zo zelf de aansluiting tussen theorie en praktijk verzorgen. Daarmee wordt hij of zij ongetwijfeld te zwaar belast.

Op meso-niveau herhaalt zich het probleem: de gewenste samenwerking tussen scholen en bedrijven ontbreekt vaak. In dit verband wordt onder andere gepleit voor het instellen van regionale contactcommissies per school en per bedrijfstak.

Over het macro-niveau (van landelijke organen bedrijfstakken, organisaties van sociale partners) zijn geen duidelijke uitspraken van respondenten voorhanden. Er zijn evenwel wel enkele opmerkingen die in dezelfde richting wijzen, zoals de geuite noodzaak tot het permanent bijstellen van beroeps- en opleidingsprofielen. Hoewel er op dit terrein de laatste jaren veel gebeurt, is het effect in lang niet alle bedrijfstakken op microniveau aanwijsbaar. Door de landelijke organen opgestelde totaalprogramma's waarin de school- en de praktijkcomponent op leerplanniveau beschreven worden, en die de mogelijkheid bieden tot onderlinge integratie van beide componenten, blijven soms vele jaren ongewijzigd, hetgeen twijfels oproept over de relevantie van de leerplannen voor de te bedienen beroepen en functies.

Voorstelbaar is, dat er tussen opleidingen of tussen landelijke organen verschillen zijn in de mate waarin andere elementen dan alleen de leerling een verbindende schakel tussen school en praktijk vormen. Te denken valt aan lesmateriaal en aan overleg tussen personen op microniveau (verantwoordelijk voor het sturen van de primaire processen, te weten leermeesters en leraren), op mesoniveau (in regionale contactcommissies) en op macroniveau (ontwikkeling van beroeps- en opleidingsprofielen door landelijke organen en andere betrokken instanties). Het zou interessant zijn om nader te onderzoeken, of de verschillen in rendement tussen opleidingen onderling en/of tussen bedrijfstakken c.q. landelijke organen zoals die hiervoor vermeld zijn, te verklaren zijn uit de mate waarin op de diverse niveaus deze verbindende schakels voorkomen.

DANKBETUIGING

Dit onderzoek is mogelijk gemaakt door subsidie van SVO (SVO-project 1604).

NOTEN

1. Inmiddels is deze stijgende tendens veranderd in een dalende. In de periode 1992-1994 is, onder invloed van de economische recessie, het aantal leer/arbeidsplaatsen in het leerlingwezen met 20% afgenomen. (Peil, 1994)

2. De Commissie Van Veen, die in 1993 advies uitbracht over dualisering, neemt een genuanceerder standpunt in. Zie de paragraaf discussie in dit artikel.
3. Omdat voor de voortgezette opleidingen niet beschikt kon worden over aantallen nieuw afgesloten leerovereenkomsten per jaar, en gebruikmaking van totale aantallen afgesloten leerovereenkomsten aan het eind van het jaar tot grote onnauwkeurigheden in de schattingen zou leiden, is afgezien van het bepalen van rendementscijfers voor de voortgezette opleidingen.

LITERATUUR

- Bont-Hock, J.G.M. de, Haas, Th.H.C. de & Kraft, P. (1989). *Aanpak van de uitval. Resultaten van een vooronderzoek naar de uitval in het leerlingwezen*. Den Bosch: Pedagogisch Centrum Beroepsonderwijs Bedrijfsleven.
- Beroepsvorming langs vele wegen*. Rapport Commissie Dualisering. (1993). Zoetermeer, Ministerie van Onderwijs en Wetenschappen.
- Boer, P.R. den en G.J. Meesterberends-Harms (1992). *Het intern rendement van het leerlingwezen, een vooronderzoek*. Groningen: RION.
- Gaalen-Oordijk, C.E. van en A.F. Westerhuis (1988). *Naar een uniforme rendementsmeting en uitvalregistratie in het leerlingwezen*. Rijswijk, Ministerie van Sociale Zaken en Werkgelegenheid.
- Ganga, V. (1992). *Deelname, uitval en rendement van het leerlingwezen*. Den Haag: Ministerie van Onderwijs en wetenschappen.
- Hövels, B., Frietman, J., & Geest, F. van (1992). *Branches over dualisering*. Uitkomsten van een enquête onder werkgevers- en werknemersorganisaties over dualisering in het beroepsonderwijs. Nijmegen, ITS.
- Janssen, A.T.H. (1990). *Het interne rendement van het middelbaar beroepsonderwijs: een literatuurstudie*. Literatuurrapport O en W nr. 12. Zoetermeer: Ministerie van Onderwijs en Wetenschappen.
- Lokman, I. & Nieuwenhuis, L. (1990). Zittenblijven of schoolverlaten; kosten-batenanalyses van MBO-leerlingen omtrent hun school- en beroepsloopbaan. In C.A.C. Klaassen & P.L.M. Jungbluth. *Onderwijs en Samenleving*. Bundel Onderwijsresearch Dagen 1990. Nijmegen: ITS.
- Peil (1994). *Leerlingwezen in meervoud. Verkenning van mogelijkheden en grenzen van groei*. Startrapport. Rijswijk: Peil (Project extra impuls leerlingwezen)/CBA.
- Pryor, B.W. (1990). Predicting and explaining intentions to participate in continuing education: An application of the theory of reasoned action. *Adult Education Quarterly*, 40, 146-157.
- Rubenson, K. (1987). Participation in recurrent education: A research review. In H.G. Schütze & D. Istance (eds.). *Recurrent education revisited* (pp. 39-67). Stockholm: Almqvist & Wiksel.
- Velden, R.K.W. van der en H. Groothuis (1984). Werving en selectie binnen het leerlingwezen. In J. Geurts et al. (red). *Onderwijs en Arbeid* (pp. 51-64). Lisse: Swets en Zeitlinger.
- Wong, R. & Siegerist E. (1989). *Werk hebben, werk houden*. OSA-voorstudie V-34. Den Haag: OSA.

Manuscript ontvangen 5-12-1994

Effecten van de aanwezigheid en inrichting van de praktijkcomponent bij scholing van laag opgeleide volwassenen

P.R. den Boer

*GION: Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling,
Rijksuniversiteit Groningen*

ABSTRACT

The Dutch government stresses the importance of a practical component in vocational education. The presence of a practical component is expected to improve study-success and to provide students with some work experience. For adults who have had little formal education and (therefore) fewer chances on the labour market, work experience is considered to be one of the keys to improve their positions. This article investigates the effects of the presence of a practical component and the way practical instruction was organized. Training results and job careers are used as the dependent variables. Results show that training results are only related to duration of the course and that job careers are related to both duration of the course and the presence of a practical component. The way practical instruction is organized has no effect on either training results or job careers. From these findings it is concluded that the effect of training lies mainly in offering employers a better opportunity for selection.

INLEIDING

Voor het beroepsonderwijs in het algemeen wordt door het beleid dualisering als middel gepropageerd om het rendement van deze opleidingen te vergroten. Van een nauwere aansluiting tussen theorie en praktijk, die van dualisering het gevolg dient te zijn, verwacht men in de eerste plaats een verhoging van het interne rendement in het MBO, ondanks de alarmerend lage interne rendementscijfers voor het leerlingwezen. Daarnaast, of door middel daarvan, dient de verdergaande dualisering het extern effect¹, ook in termen van inhoudelijke aansluiting tussen onderwijs en arbeidsmarkt, te vergroten. De processen waarlangs dualisering naar een hoger rendement en gunstiger externe effecten zou moeten leiden hebben meestal te maken met het 'opdoen van relevante praktijkervaring' en met socialiseringsprocessen. Voor het beroepsonderwijs in het algemeen geldt dat dergelijke invloeden niet of nauwelijks zijn aangetoond (De Vries, 1988 en Nieuwenhuis, 1990, Frietman, 1990).

Ook voor scholing van laagopgeleide volwassenen wordt veel verwacht van het opdoen van relevante praktijkervaring, veelal vertaald in: 'opdoen van werkervaring' en in 'trainen van werkcritme' (zie bijvoorbeeld De Grip, 1988). Deze vertaling is gebaseerd op de veronderstelling, dat de meeste deelnemers aan dergelijke scholing weinig werkervaring zouden hebben en/of al langdurig werkloos zijn en daardoor werkcritme ontberen, waardoor zij voor werkgevers minder aantrekkelijk zouden zijn.

In diverse onderzoeken echter is aangetoond, dat allerlei andere factoren dan werkervaring een rol spelen bij de selectie van personeel. Van Beek (1993) toont aan, dat niet de opleiding of de werkervaring, maar de sekse, de etnische herkomst, de gezondheid en de leeftijd van de sollicitant de kans op een baan bepalen². Onderzoek van Oosterhuis en Glebbeek (1988) wees

erop, dat het opleidingsniveau bij personeelsselectie alleen in zoverre een rol speelt, dat werkgevers veronderstellen, dat een bepaald opleidingsniveau gecorreleerd is aan bepaalde kenmerken van de sollicitant. Ook uit onderzoek naar de effecten van werkervaringsprojecten, blijkt weinig reden voor het veronderstellen van rechtstreekse relaties (De Koning, Koss & Verkaik, 1988), maar eerder van 'koudwatervrees' ten opzichte van doelgroepen, die alleen met een grichte benadering en veel begeleiding te overwinnen is (Van der Berg & Van der Veer, 1990, Kraayvanger & Van Onna, 1990, De Koning & Van Nes, 1989).

Samengevat lijkt het juister te veronderstellen, dat scholing een signaalfunctie vervult waarmee werkgevers wordt getoond, dat de betrokkene inzet en intelligentie bezit (De Koning, 1992). Uitgaande van deze wetenschap, valt te verwachten, dat de aanwezigheid en de inrichting van een praktijkdeel bij scholing aan de effecten daarvan weinig of niets meer kan toevoegen, omdat de variabelen op basis waarvan de personeelsselectie plaatsvindt met hetgeen op de praktijkplaats gebeurt niets van doen hebben.

In dit artikel zal deze verwachting worden onderzocht voor deelnemers aan scholing in het kader van de Primaire Beroepsgerichte Volwassenen Educatie (PBVE). De Primaire Beroepsgerichte Volwassenen Educatie (PBVE) was een scholingsmaatregel van de gezamenlijke ministeries van Onderwijs en Wetenschappen (O&W) en Sociale Zaken en Werkgelegenheid (SZW). De PBVE-maatregel had tot doel op het niveau van de regio een onderwijsinfrastructuur te creëren waarmee gestalte kon worden gegeven aan arbeidsmarktrelevante scholing voor laag opgeleide volwassenen met een zwakke positie op de arbeidsmarkt. Het scholingsaanbod moest bestaan uit een samenhangend pakket van oriënterende, schakelende en beroepskwalificerende opleidingsprogramma's, die gerealiseerd konden worden bij bestaande scholingsinstellingen. Het streefniveau van de opleidingen was het kwalificatieniveau van het Primair Leerlingwezen (niveau beginnend beroepsbeoefenaar). De aansluiting met de beroepspraktijk werd in de PBVE, in ieder geval voor het beroepsopleidende cursusaanbod, nagestreefd door middel van de inrichting van een praktijkdeel (Beroepseducatie volwassenen, 1987).

In dit artikel worden twee vragen beantwoord³:

1. draagt de aanwezigheid van een praktijkdeel in de opleiding bij tot het intern rendement en de externe effecten van scholing, en;
2. welke elementen van dat praktijkdeel zijn daarin vooral van belang.

DATA, MODEL EN METHODE

Voor de beantwoording van de eerste onderzoeksvraag zijn data gebruikt die zijn verzameld bij oud-PBVE-cursisten, die tussen mei en juli 1989 hun PBVE-cursus verlieten, al dan niet met een diploma. Een steekproef van 1056 oud-cursisten is in juni 1990, een jaar na het verlaten van de cursus, telefonisch benaderd. Van 542 van hen (51%) werd zinvol te verwerken respons verkregen. De steekproef is zodanig getrokken, dat oud-cursisten uit technische en verzorgende opleidingsrichtingen oververtegenwoordigd waren en cursisten uit Oriënterende en Schakelende (O&S) programma's en uit andere dan technische, verzorgende of administratieve beroepsopleidingen ondervertegenwoordigd. De reden daarvoor was, dat grotere variatie in inrichting van de praktijk verwacht werd in deze opleidingsrichtingen. Omdat onder allochtonen een hogere non-respons werd verwacht, zijn extra allochtonen benaderd voor deelname aan het onderzoek.

Voor de beantwoording van de tweede onderzoeksvraag zijn, naast de beschreven steekproef van cursisten, ook hun praktijkbegeleiders bevroegd, vooral betreffende de inrichting van de praktijkplaats. De praktijkbegeleiders zijn benaderd met behulp van informatie van de cursisten. Deze hadden gezamenlijk 369 praktijkplaatsen bezocht, waarbij sommige cursisten meer dan één praktijkplaats bezochten. Van deze praktijkplaatsen, waren uiteindelijk 204 begeleiders traceerbaar (55%). De overige waren niet meer te achterhalen vanwege onvoldoende, ontbrekende of onjuiste gegevens. Van 138 praktijkbegeleiders (68%) werd respons verkregen. Zij beschrijven de praktijkplaatsen van 132 cursisten.

De steekproeven van oud-cursisten en van praktijkbegeleiders zijn niet representatief voor hun populaties. Dat is voor de beantwoording van de hier gestelde vragen ook minder van belang, omdat het hier gaat om het opsporen van samenhang tussen variabelen en niet om een representatieve beschrijving.

In het algemeen mag verondersteld worden, dat op het optreden van effecten vier actoren invloed uitoefenen: 1) de cursist zelf, 2) de opleiding, 3) de arbeidsmarkt en 4) de intermediair(s) tussen cursist en opleiding en tussen cursist en arbeidsmarkt. Het niveau van de intermediairs wordt in dit artikel buiten beschouwing gelaten, het is voor de beantwoording van de hier gestelde vragen niet relevant (zie daarvoor Den Boer, 1995). De variabelen van de opleiding zijn als voorspellers van effect beschouwd, voor zover deze de aanwezigheid en inrichting van het praktijkdeel betreffen. De overige variabelen, die van de cursist, de arbeidsmarkt en van de opleiding voor zover ze de inrichting van het theoretische deel van de opleiding betreffen, zijn als storende variabelen beschouwd.

Aan de cursist zijn de volgende aspecten onderscheiden:

1. ascriptive kenmerken: sekse, herintreedsters, etniciteit en leeftijd,
2. kenmerken van de loopbaan tot dusver: lengte van de periode van werkloosheid, werkervaring, vooropleiding, positie bij aanvang aan de cursus (werkend in dezelfde of een andere branche als de opleiding of werkloos),
3. de motivatie om aan de cursus te beginnen: 'gedwongen' door arbeidsbureau/sociale dienst of ongemotiveerd en: motivatie voor deelname gericht op leren of het vinden van werk.

Op de arbeidsmarkt zijn, zoals door Onstenk & Wilbrink (1992) aangetoond (overigens op grond van hetzelfde datamateriaal als hier gebruikt), vooral de krapte en gestructureerdheid van de arbeidsmarkt relevant: technische en verzorgende opleidingen leiden naar een krappe arbeidsmarkt en resulteren in relatief gunstiger arbeidsmarktposities, administratieve opleidingen leiden naar een niet-krappe, weinig gestructureerde arbeidsmarkt en resulteren in relatief ongunstiger arbeidsmarktposities. In de analyses is daarom voor opleidingsrichting gecontroleerd. Daarin zijn cursussen onderscheiden in de volgende richtingen: 1) O&S, 2) administratie, 3) techniek, 4) verzorging en 5) andere beroepsopleidingen.

Aan het theoretisch gedeelte van de opleiding is in de eerste plaats relevant te onderscheiden van welk niveau de opleiding is (zie inleiding). Zoals ook door de toenmalige intermediair PBVE al geconstateerd (Berenschot, 1989, p.15), bleek het onmogelijk een niveau-indeling te maken op basis van de beschikbare informatie over de cursussen. De lengte van de gevolgde opleiding is daarom als indicator voor het opleidingsniveau genomen. Dat is niet helemaal correct: voor cursisten met hogere vooropleidingen zullen kortere cursussen wellicht van hoger niveau zijn geweest dan langere cursussen voor cursisten met lagere vooropleidingen. Over het geheel genomen mag echter verondersteld worden, dat tussen niveau en lengte een lineair verband bestaat: hoe langer de cursusduur, hoe hoger het niveau. Daarnaast zijn aan het theorie-deel onderscheiden: het aantal dagen theorie per week (dat is: de mogelijkheid tot het realiseren van een praktijkdeel), de instelling die het theoriedeel verzorgde, de aan de cursus gekoppelde garantie op een baan en het door de cursist doorlopen traject (met of zonder voorafgaande O&S of andere cursus).

Naast de aanwezigheid van een praktijkdeel (wel/niet) zijn variabelen onderscheiden betreffende de inrichting van de praktijkplaats en betreffende de organisatie en de randvoorwaarden bij de praktijk.

Met behulp van factor-analyse, konden de 24 variabelen van de inrichting van de praktijkplaats gereduceerd worden tot vijf, te weten:

- de gerichtheid op werkcritmetraining, waaronder wordt verstaan: confrontatie met verschillende arbeidsomstandigheden (zoals prikklokken, overwerk en werken op onregelmatige tijden), verschillende soorten werkzaamheden (zoals omgaan met cliënten en overleg met collega's) en een duidelijke opbouw van werkzaamheden van gemakkelijk naar moeilijk, eenvoudig naar complex;

- de gerichtheid op ervaringsleren, waaronder wordt verstaan: meedraaien met de normale beroepspraktijk waarbij de cursist wordt ingezet voor de produktie, aanwezigheid van vaklui in de buurt van de cursist en de mogelijkheid om veel verschillende werkervaringen op te doen;
- het theoretisch niveau, waaronder wordt verstaan: de aansluiting bij en het gebruik van het op school geleerde, de erkenning door het landelijk orgaan van het leerlingwezen, het niveau van de doelstellingen voor het praktijkgedeelte op een schaal van het opdoen van werkervaring tot het leren omgaan met nieuwe beroepsinhouden en -technieken;
- de begeleiding, bestaande uit: aanwezigheid van de begeleider op de praktijkplaats, veel begeleidingstijd en goede kwalificaties van de begeleiding;
- de concurrentie op de praktijkplek: de cursist heeft te maken met concurrentie voor het behouden van de praktijkplaats en voor het verwerven van een baan na afloop.

Aan de organisatie en de randvoorwaarden op de praktijkplaats zijn de volgende vijf variabelen onderscheiden:

- of de praktijk plaats vond in een bedrijf of een andere instelling (op school, in een GOA⁴, in een Centrum Vakopleiding (CV) of Centrum voor Beroepsoriëntatie en Beroepsoefening (CBB), eerst op school en daarna in een bedrijf of in een andere instelling),
- de aard van de aanstelling tijdens de praktijkperiode (geen, tijdelijk of vast),
- het aantal praktijkplaatsen, dat de cursist tijdens de scholing bezocht,
- de totale lengte van de praktijkperiode(s),
- de betrokkenheid van school, het landelijk orgaan en de PBVE bij de begeleiding op de praktijkplaats.

Het intern rendement van een opleiding is de proportie leerlingen van een instroomcohort, dat, binnen een bepaalde tijd, de opleiding gediplomeerd verlaat (Van der Aa, 1988; Van Galen-Oordijk, 1987; De Bont-Hoek, 1987; Den Boer & Meesterberends-Harms, 1992). Deze definitie, hoe geldig ook, kon in dit geval niet worden toegepast. Bij de dataverzameling is namelijk prioriteit gegeven aan de vaststelling van het extern effect. Daarbij was van belang te zorgen voor een zo gering mogelijke spreiding in het moment van intrede op de arbeidsmarkt, om storingen ten gevolge van fluctuaties op de arbeidsmarkt zo veel mogelijk te kunnen uitsluiten. Het intern rendement, gebaseerd op de hier gebruikte gegevens van de bevraagde groep PBVE-cursisten die in de periode mei t/m juli 1989 al dan niet gediplomeerd uitstroomde, moet dus beschouwd worden als een schatting van het intern rendement van de PBVE-cursussen. In de analyses, die op individu-niveau zijn uitgevoerd, zijn bovendien feitelijk de slaagkansen van de individuele cursisten gebruikt voor een schatting van het intern rendement van het PBVE-scholingssysteem. In hoeverre daardoor een over- of onderschatting van het feitelijk intern rendement is opgetreden is niet duidelijk. Voor het bepalen van de invloed van de aanwezigheid van een praktijkdeel op het intern rendement is dat echter niet direct van belang.

Vanwege het gedifferentieerde cursusaanbod in het kader van de PBVE, bestaande uit oriënterende, schakelende en beroepsopleidende cursussen van verschillend niveau, en vanwege het feit dat het niveau van het beroepsopleidende cursusaanbod in het kader van de PBVE geheel regionaal bepaald is en daardoor moeilijk te overzien (zie Onstenk & Wilbrink, 1992 en Den Boer, 1991a en b) is een gedifferentieerd slaagcriterium gebruikt. Voor de oriënterende en schakelende (O&S-) cursussen is als criterium gehanteerd dat de cursist alleen de cursus moet hebben afgemaakt, omdat diploma's in dit soort cursussen lang niet altijd worden uitgereikt. Voor de beroepsopleidingen is het behaald hebben van een diploma als criterium gehanteerd, ongeacht of het een diploma voor het theorie- of het praktijkdeel of voor beide betreft.

Als beoogd extern effect van scholing voor de deelnemers werd door het beleid genoemd: 'kwalificatie ten behoeve van het verwerven van een plaats in het arbeidsbestel' (Beroepseducatie volwassenen, 1987, p.9). Naast het vinden van een baan gaat het daarbij om de aard van de verworven baan (tijdelijk of vast), de promotiemogelijkheden in het bedrijf (geen, weinig of

onbekend, veel) en het niveau van de baan (vastgesteld op basis van de benodigde kennis en inwerktijd). Deze vier aspecten zijn tot één afhankelijke variabele gecombineerd door: het vinden van een baan (0=niet, 1=wel), de soort baan (1=tijdelijk, 2=vast) en de promotiemogelijkheden (1=geen, 2=weinig/onbekend, 3=veel) met elkaar te vermenigvuldigen en dit resultaat te vermeerderen met het beroepsniveau (1 tot 5). De aldus gecreëerde maat voor extern effect kent 11 categorieën.

O&S-cursussen zijn niet in eerste instantie gericht op het vinden van werk, maar op toeleiding naar beroepsopleidende cursussen, al bleek in de praktijk ook een arbeidsmarkteffect van deze cursussen uit te gaan (Den Boer, 1991a en b). Gezien de diversiteit van het scholingsaanbod wordt instroom in een vervolgopleiding veelal eveneens als een resultaat van scholing beschouwd. Uitgaande van de 'scholingstraject'-gedachte is dat niet juist. Onder een scholings-traject wordt een samenhangend stelsel van scholing verstaan, dat voor de deelnemer de passende route naar de arbeidsmarkt vormt. Zolang een deelnemer zich nog in dit traject bevindt, is in feite nog geen sprake van uitstroom en kunnen dus ook nog geen externe effecten worden vastgesteld. Bij de berekeningen met extern effect als afhankelijke variabelen zijn de cursisten die bij de bevraging in juni 1990 bezig waren met het doen van een (vervolg)-opleiding en geen werk hadden, dan ook buiten beschouwing gelaten. Bij deze berekeningen geldt dan ook een afwijkende steekproefomvang ($n=464$).

Met uitzondering van de variabelen van de inrichting van de praktijkplaats (werkritmetraining, etc.) zijn alle variabelen geoperationaliseerd in vragen aan de oud-cursisten. De gegevens over de inrichting van de praktijkplaats zijn verzameld bij de praktijkbegeleiders.

De analyses zijn uitgevoerd met behulp van regressie-analyse. Voor de analyses met het intern rendement als afhankelijke variabele was daarvoor een logaritmische transformatie nodig, omdat het intern rendement een dichotome (0-1) variabele is. Met behulp van LOGIT zijn, na deze logaritmische transformatie, regressie-coëfficiënten geschat. Bij alle analyses is een significantie-niveau van 5% gehanteerd, eenzijdig getoetst (kritische T-waarde: ± 1.64).

Voor de beantwoording van de eerste onderzoeksvraag is eerst een goede set controlevariabelen gezocht. Daartoe zijn alle variabelen van de cursist en de opleidingsrichting (geforceerd) in de analyse ingevoerd. Vervolgens zijn stapsgewijs de variabelen met de laagste (geschatte) regressiecoëfficiënten uit de analyse verwijderd.

Deze variabelen zijn vervolgens samen met de voorspeller 'aanwezigheid van een praktijkdeel' in een analyse ingevoerd. Daarmee kon de mate waarin de aanwezigheid van het praktijkdeel significant aan de effectmaten bijdraagt worden bepaald.

Door nu vervolgens die variabelen van de inrichting van het theoriedeel, waarvoor (in een afzonderlijke analyse) significante invloeden op de effectmaten zijn gevonden, in de analyse in te voeren, kon tevens worden nagegaan of en in welke mate deze de invloed van de aanwezigheid van een praktijkdeel op de effecten wegnemen. De voor het intern rendement relevante variabelen van het theoriedeel zijn opgespoord in een afzonderlijk uitgevoerde LOGIT-analyse. Voor het opsporen van de relevante variabelen van het theoriedeel voor het extern effect, is gebruik gemaakt van regressie-analyse met forward inclusion. Daarbij worden die variabelen in de analyse opgenomen, die een significante relatie met de effectmaat vertonen, nadat de controlevariabelen en de aanwezigheid van praktijk (geforceerd) in de analyse zijn ingevoerd, zodat voor deze variabelen is gecorrigeerd.

De beantwoording van de tweede onderzoeksvraag is afhankelijk van de beantwoording van de eerste: alleen als blijkt dat de aanwezigheid van een praktijkdeel bijdraagt aan het effect, is het zinvol na te gaan welke aspecten van de organisatie, randvoorwaarden en inrichting van het praktijkdeel daaraan nog iets toevoegen.

Daarvoor is eerst in afzonderlijke analyses nagegaan welke van deze variabelen van significante invloed zijn op de afhankelijke variabelen. Voor de variabelen betreffende organisatie en randvoorwaarden konden die analyses worden uitgevoerd voor de gehele onderzoeksgroep,

uiteraard voor zover zij praktijk bij hun opleiding hadden gehad ($n=300$). Voor de variabelen betreffende de inrichting konden de analyses alleen met een deel van de onderzoeksgroep worden uitgevoerd ($n=132$), van de overigen werden geen gegevens verkregen (zie boven). De bijdrage van deze variabelen aan de effecten is vervolgens vastgesteld, door ze in de regressie-analyses in te voeren nadat de significante controle-variabelen, de significante variabelen van het theoriedeel en de voorspeller 'aanwezigheid van een praktijkdeel' (geforceerd) waren ingevoerd.

RESULTATEN 1: PRAKTIJK EN INTERN RENDEMENT

Van de cursistvariabelen en de opleidingsrichting bleken etniciteit, werkervaring, vooropleiding, O&S-cursussen en opleidingen in de verzorgende beroepen van significante invloed op de slaagkansen van de cursisten. De invloed van etniciteit was negatief, de invloed van de overige variabelen positief (zie tabel 1).

Deze variabelen zijn vervolgens samen met de voorspeller 'aanwezigheid van een praktijkdeel' in een analyse ingevoerd. De aanwezigheid van een praktijkdeel bleek daarin een significante bijdrage aan de slaagkansen te leveren.

In een afzonderlijk analyse bleken vervolgens twee variabelen van de inrichting van het theoriedeel de slaagkansen van de cursisten significant te beïnvloeden, namelijk de lengte van de opleiding (positief) en het aantal dagen theorie per week (negatief). Voor geen van de overige variabelen van het theoriedeel, de instelling waar de theorie gevolgd werd, het voortraject (O&S- of andere cursus) en de bij de opleiding geboden baangarantie, werd een significante invloed op het intern rendement gevonden.

Bij invoer van deze variabelen in de LOGIT-analyse bleken de eerder gevonden invloeden van de aanwezigheid van een praktijkdeel en van de opleidingsrichting verzorging weg te vallen ten gunste van de lengte van de opleiding (zie tabel 2).

Gezien het feit, dat de aanwezigheid van een praktijkdeel geen bijdrage aan het intern rendement leverde, is niet verder nagegaan welke elementen van de praktijk nog extra bij konden dragen aan de slaagkansen van de cursisten en het intern rendement van de opleidingen.

Tabel 1. Invloed cursistvariabelen en opleidingsrichting op de slaagkansen van cursisten.

	geschatte regressie coëfficiënt	standaardfout	regr.coëff./st.ft.
Sekse (vrouw)	.06	.13	.46
Herintreedsters	.21	.13	1.59
Etniciteit (allochtoon)	.36	.13	-2.75**
Leeftijd	.01	.04	.27
Werkervaring	.09	.04	1.92*
Werkloosheidsduur	.06	.04	-1.45
Vooropleiding	.07	.04	1.68*
Aanvangssituatie/branche waarin werkend	.06	.08	-.83
Motivatie (leren/werken)	.01	.07	.10
Verplichting/ongemotiveerdheid	.14	.17	-.81
Opleidingsrichting: O&S	.57	.17	3.38**
Verzorging	.46	.21	2.22*
Administratie	.12	.13	.91
Techniek	.24	.15	1.55
Andere beroepsopleiding	.00	.00	

$N=542$; * = $p \leq .05$, ** = $p \leq .01$ (enzijdig getoetst).

Tabel 2. Invloed aanwezigheid praktijk op slaagkansen, na invoer van de lengte van de opleiding.

	geschatte regressiecoëfficiënt	standaardfout	reg.coëff./st.fl.
Etniciteit (allochtoon)	.34	.13	-2.59**
Werkervaring	.11	.04	2.81**
Vooropleiding	.07	.04	1.91*
Opleidingsrichting: O&S	.64	.15	4.32**
Verzorging	.25	.19	1.30
Aanwezigheid van een praktijkdeel	.09	.10	.92
Lengte van de opleiding	.12	.03	4.03**
Aantal dagen theorie per week	-.07	.04	-1.59

$N=542$; * = $p \leq .05$; ** $p \leq .01$ (eenzijdig getoetst).

RESULTATEN 2: PRAKTIJK EN EXTERN EFFECT

Van de controlevariabelen bleken de volgende een significante relatie met het extern effect te hebben: sekse, etniciteit, werkervaring, werkloosheidsduur, het gevoel van 'verplichting' tot deelname (of de ongemotiveerdheid van de cursist) en de beroepsopleidingen in de andere richtingen dan technische, verzorgende of administratieve opleidingen (zie tabel 3). Voor alle overige voorspellers van de cursist, te weten herintreedsters, leeftijd, vooropleiding, aanvangssituatie, motivatie voor leren of werken, werden geen significante relaties met het extern effect gevonden, evenmin als voor de opleidingsrichtingen O&S, administratie, techniek en verzorging.

De invloed van werkervaring was positief, alle andere negatief. Gezamenlijk konden deze variabelen 16% van de variantie in het extern effect verklaren.

Vervolgens is, onder controle voor de boven genoemde controlevariabelen, de variabele 'aanwezigheid van een praktijkdeel' in de analyse ingevoerd. Deze bleek significant bij te

Tabel 3. Invloed cursistkenmerken en opleidingsrichtingen op extern effect.

	Bèta	T
Sekse (vrouw)	-.24	-5.47**
Etniciteit (allochtoon)	-.12	-2.70**
Werkervaring	.13	2.96**
Lengte van werkloosheid	-.20	-4.50**
Gevoel van 'verplichting', ongemotiveerdheid	-.08	-1.92*
Opleidingsrichting: andere beroepsopleidingen	-.10	-2.38**
Herintreedsters	-.02	-.29
Leeftijd	.00	-.00
Vooropleiding	.03	.58
Aanvangssituatie	.06	1.10
Motivatie (leren/werken)	.00	.08
Opleidingsrichting: O&S	-.07	-1.58
Techniek	-.02	-.41
Verzorging	.07	1.48
Administratie	.03	.62
Totale verklaarde variantie (R^2): 16,0%		

$N=464$; * = $p \leq .05$; ** = $p \leq .01$ (eenzijdig getoetst).

Tabel 4. Invloed aanwezigheid praktijk op extern effect na invoer van de lengte van de opleiding.

	Bèta	T
Sekse (vrouw)	-.21	-4.82**
Etniciteit (allochtoon)	-.10	-2.23*
Werkervaring	.15	3.41**
Lengte van werkloosheid	-.20	-4.68**
Gevoel van 'verplichting', ongemotiveerdheid	-.09	-2.15*
Opleidingsrichting: andere beroepsopleidingen	-.08	-1.80*
Aanwezigheid van een praktijkdeel	.15	3.30**
Lengte van de opleiding	.09	1.90*

Totale verklaarde variantie (R^2): 19,4%

$N=464$; * = $p \leq .05$; ** = $p \leq .01$ (eenzijdig getoetst).

dragen aan het extern effect. Door toevoeging van deze variabele kon 2,8% extra variantie in extern effect worden verklaard.

Van de variabelen van het theoriedeel bleek alleen de lengte van de opleiding een significante extra bijdrage aan het extern effect te kunnen leveren. Het aantal dagen theorie per week, de instelling waar de opleiding gevolgd werd, het doorlopen hebben van een voortraject (bestaande uit een O&S- of andere cursus) noch het hebben van een baangarantie bij het volgen van de cursus konden aan het extern effect van de scholing bijdragen.

Door de variabele opleidingslengte in de analyse in te voeren is nagegaan of het opleidingsniveau van invloed was op het extern effect, ongeacht de vormgeving van het theoriedeel van de opleiding. De invloed van deze variabele op het extern effect bleek significant, nadat de invloed van de genoemde variabelen is uitgepartialiseerd (zie tabel 4). De lengte van de opleiding kon nog 0,8% extra variantie van het extern effect verklaren. De invloed van de aanwezigheid van een praktijkdeel op het extern effect werd enigszins gereduceerd, de β -coëfficiënt liep terug van .168 naar .146, maar de relatie met het effect bleef significant. De variabelen verklaarden nu gezamenlijk 19,4% van de variantie in het extern effect.

Na de vaststelling, dat de aanwezigheid van een praktijkdeel van invloed is op het extern effect, is, ter beantwoording van de tweede onderzoeksvraag, nagegaan of onderscheiden aspecten van de inrichting, organisatie en randvoorwaarden daarop nog aanvullend invloed uitoefenden.

In afzonderlijke analyses is nagegaan welke van deze variabelen van significante invloed zijn op het extern effect. Dat bleken de volgende te zijn: het doorlopen van de praktijkperiode in de categorie instellingen anders dan een bedrijf, de aard van de aanstelling van de cursist op de praktijkplaats en de gerichtheid van de praktijkervaring op werkritmetraining (alle drie positief). Vervolgens zijn deze significante aspecten van het praktijkdeel in regressie-analyses ingevoerd met de genoemde cursistvariabelen, opleidingsrichting, de aanwezigheid van een praktijkdeel en de lengte van de opleiding, om de invloed van elementen van de inrichting op het extern effect vast te stellen.

Voor geen van deze variabelen werd, na controle voor de reeds in de analyse ingevoerde variabelen, een significante relatie met het extern effect gevonden: geen van allen voegt nog iets substantieels aan het extern effect toe.

CONCLUSIES EN DISCUSSIE

Voor de slaagkansen van de cursisten is de aanwezigheid van een praktijkdeel niet van belang. Ook het aantal theoriedagen per week (geïnterpreteerd als een inrichting van het theoriedeel die een duale opleiding (on)mogelijk maakt; hoe meer dagen theorie hoe minder een duale oplei-

ding mogelijk is en omgekeerd) levert geen significante bijdrage aan de slaagkansen van de cursist. De slaagkansen worden beïnvloed door cursistkenmerken (allochtonen vallen veel vaker uit en cursisten met werkervaring en met een hogere vooropleiding vallen minder vaak uit), en de lengte van de opleiding. De O&S-cursussen hebben daarnaast een hoger intern rendement dan beroepsopleidingen. Omdat de slaagkansen van de cursisten de berekeningsbasis vormen voor intern rendement, lijkt dualisering dus niet het aangewezen middel voor de verhoging van het intern rendement van beroepsgerichte volwasseneneducatie.

Daarnaast wordt geconstateerd, dat voor opleidingen in de scholing voor laag opgeleide volwassenen kennelijk geldt: hoe langer de opleiding is, hoe hoger het intern rendement. Dat is tegengesteld aan wat verwacht zou mogen worden. Voor de scholingsdoelgroep, bestaande uit laaggeschoolden, waarvoor aangenomen wordt dat zij slechte ervaringen met opleidingen hebben, is de verwachting juist, dat kortere cursussen een hoger intern rendement opleveren dan langere. Gezien het positieve effect van deelname aan O&S-cursussen op het intern rendement en het tegelijkertijd optredende positieve effect van de lengte van de opleiding, bestaat het vermoeden, dat er een aantal niet al te lange beroepsopleidingen bestaat zonder praktijkdeel en met een ongunstig intern rendement. Bij nadere analyse kon dit vermoeden overigens niet worden bevestigd (zie Den Boer, 1995).

Het extern effect wordt voorspeld door de lengte van de opleiding en door de aanwezigheid van een praktijkdeel, ongeacht de inrichting en organisatie daarvan en ongeacht de randvoorwaarden daarbij.

De significante bijdrage van de lengte van de opleiding aan het extern effect werd verwacht, immers de lengte van de opleiding zegt iets over het niveau van de opleiding en in de inleiding werd uit ander onderzoek reeds aangehaald, dat het niveau van de opleiding van belang is voor het optreden van effect.

De significante bijdrage van de aanwezigheid van een praktijkdeel aan het extern effect werd niet verwacht. In de inleiding werd verondersteld, dat scholing vooral een signaalfunctie zou hebben, waarmee de geschoolde de werkgever kan tonen over inzet en intelligentie te beschikken. Omdat personeelsselectie plaats vindt op grond van andere factoren dan dat wat op de praktijkplaats gebeurt, werd verondersteld, dat noch de aanwezigheid, noch de inrichting van de praktijkplaats van invloed zou zijn op het extern effect. Het eerste is niet het geval, maar het tweede wel: de wijze van vormgeving en organisatie van de praktijk en de randvoorwaarden die daarbij gerealiseerd zijn lijken aan het extern effect niets meer te kunnen toevoegen.

Deze laatste constatering roept vraagtekens op, vooral omtrent het belang van dualisering voor verhoging van het extern effect. Dualisering is immers bedoeld, om een inhoudelijke aansluiting van theorie en praktijk in de beroepsopleiding tot stand te brengen. Ten gevolge van een inhoudelijk betere aansluiting zou de waarde van de scholing op de arbeidsmarkt moeten toenemen. Daarom zou verwacht mogen worden, dat in dit onderzoek een directe invloed gevonden zou worden van het theoretisch niveau van de praktijkplaats en/of de gerichtheid op ervaringsleren op de arbeidsmarktpositie. Dat is niet het geval. Voor de werkgever/selecteur is de inrichting c.q. de kwaliteit van de praktijkplaats kennelijk niet van belang, terwijl de aanwezigheid van een praktijkplaats wel invloed heeft op het extern effect. Dat lijkt erop te wijzen, dat werkgevers/selecteurs de praktijkperiode tijdens de scholing als een soort verlengde proefperiode gebruiken. De gevonden invloeden van cursistkenmerken sekse (vrouwen komen er slechter af), etniciteit (allochtonen komen er slechter af) en werkervaring (positief effect), maar ook van de lengte van de opleiding, hetgeen opgevat kan worden als indicator voor het opleidingsniveau, lijken deze veronderstelling te bevestigen. Werkgevers/selecteurs letten niet op de inrichting van de opleiding, c.q. de praktijkperiode, maar op het opleidingsniveau, hetgeen de verwachte extra investering in bijscholing representeert (Glebbeeck, 1988) of andere 'relevante' kenmerken van de cursist. Daarnaast wordt door de werkgever/selecteur gelet op werkervaring (zie ook Hoeben, 1992) en sekse en etniciteit (zie ook Van Beek, 1993; Van der Aa, 1992).

Deze conclusies zijn gebaseerd op gegevens over scholing voor laagopgeleide volwassenen. Ze zijn dus niet zonder meer toepasbaar op het beroepsonderwijs in het algemeen. Ze hebben echter ook voor deze opleidingen enige zeggingskracht. De mechanismen die spelen op de

arbeidsmarkt zijn niet beperkt tot de doelgroepen waar scholing zich op richt, maar bezitten algemenere geldigheid, zij het dat ze waarschijnlijk een belangrijker rol zullen spelen op een ruimere arbeidsmarkt (scholing), dan voor een krappere (MBO, leerlingwezen). In dat opzicht lijkt een waarschuwing aan het adres van de beleidmakers op zijn plaats waar het de risico's van verdergaande dualisering betreft. Dualisering leidt *niet* automatisch naar een hoger intern rendement of naar gunstiger externe effecten. Dat laat de wenselijkheid van een nauwere aansluiting van theorie en praktijk in het beroepsonderwijs onverlet.

NOTEN

1. In onderwijskundige literatuur is hier de term rendement gangbaar. Omdat daarmee in economische literatuur winst op onderneming wordt verstaan en hier het kostenaspect volledig buiten beschouwing blijft, zal verder van het extern effect gesproken worden.
2. Daarbij moet overigens worden opgemerkt, dat Van Beek scholing alleen in termen van onder- en overscholing heeft meegenomen, gedefinieerd door de selecteur.
3. Het onderzoek is mogelijk gemaakt door subsidie van SVO (SVO-project 9065)
4. Gemeenschappelijk Onderwijs Activiteit, een gezamenlijk initiatief van verschillende bedrijven, waarbij de praktijk in een simulatie en/of afwisselend in verschillende bedrijven is vormgegeven

LITERATUUR

- Aa, R. van der (1992). Tussen wens en werkelijkheid: scholing en werving van doelgroepen op de arbeidsmarkt. In A.F.M. Nieuwenhuis, *Scholing: onderzoek en beleid. Effectief scholingsbeleid voor markt en beroep*. Utrecht: Uitgeverij Lemma BV.
- Beek, K.W.H. van (1993). *To be hired or not to be hired, the employer decides*. Proefschrift Universiteit van Amsterdam. Amsterdam: Universiteit van Amsterdam.
- Berg, H. van den & Veer, K. van der (1990). *Hoezo onbemiddelbaar? Een evaluatie van een nieuwe methode van arbeidsbemiddeling voor moeilijk plaatsbare werklozen*. Amsterdam: Free University Press.
- Berenschot (1989). *Evaluatieverslag 1988 Primaire Beroepsgerichte Volwasseneneducatie*. Utrecht.
- Beroepseducatie volwassenen* (1987). Notitie PBVE. Tweede Kamer, 1986-1987, 18 886, nr.30, Den Haag: SDU.
- Boer, P.R. den (1991a). Vooronderzoek uitstroom PBVE. In *Evaluatie PBVE. Effecten en uitvoeringspraktijk*. Arbeidsvoorzieningsreeks nr 1. Rijswijk: Arbeidsvoorziening, centraal bureau.
- Boer, P.R. den (1991b). *De uitstroom uit de PBVE in kaart gebracht*. Groningen: RION.
- Boer, P.R. den & Hofman, R.H. (1992). *Inrichting en effectiviteit van de praktijkcomponent PBVE*. Groningen: RION
- Boer, P.R. den & Meesterberends-Harms, G.J. (1992). *Het intern rendement van het leerlingwezen. Een vooronderzoek*. Groningen: RION
- Boer, P.R. den (1995). *Scholing van laag opgeleide volwassenen*. Een onderzoek naar de inrichting en effecten van scholing in het kader van de PBVE. Proefschrift Groningen. Groningen: RION.
- Bont-Hoek, J.G.M. de, Haas, Th.H.C. de & Kraft, P. (1989). *Aanpak van de uitval: resultaten van een vooronderzoek naar de uitval in het leerlingwezen*. Den Bosch: PCBB.
- Frietman, J. (1990). *De kwaliteit van de praktijkcomponent in het leerlingwezen*. Proefschrift Nijmegen. Nijmegen: ITS.
- Gaalen-Oordijk, C.E. van & Westerhuis, A.F. (1989). *Naar een uniforme rendementsmeting en uitvalregistratie in het Leerlingwezen*. Rijswijk: Ministerie van Sociale Zaken en Werkgelegenheid.
- Glebbeeck, A.C. (1988). De arbeidsmarktpositie van opleidingen. *Tijdschrift voor arbeidsvraagstukken*, 4, 75-89.
- Grip, A. de (1988). *Scholing en werk. Een analyse van de scholingsproblemen op de arbeidsmarkt en het arbeid*. SMO-boek. Den Haag: Stichting Maatschappij en Onderneming.
- Koning, J. de, Koss, M. & Verkaik, A. (1988). *Tijdelijk werk bij de overheid: een goede start voor een baan in het bedrijfsleven?* Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.
- Koning, J. de & Nes, P.J. van (1989). *Landelijke evaluatie Kaderregeling Scholing*. O.A.V.-rapport 90-11. Rijswijk: Arbeidsvoorziening.

- Koning, J. de (1992). Evaluatie van scholingsmaatregelen voor werklozen: wat is er bekend van de resultaten. In A.F.M. Nieuwenhuis (red.) (1992). *Scholing: onderzoek en beleid. Effectief scholingsbeleid voor markt en beroep*. Utrecht: Uitgeverij Lemma BV.
- Kraayvanger, G. & Onna, B. van (1990). *Scholing en werk voor langdurig werklozen: Op weg naar een meer integrale benadering!?* Nijmegen: ITS
- Nieuwenhuis, A.F.M. (1990). *Complexe leerplaatsen in school en bedrijf*. Proefschrift Groningen. Groningen: RION.
- Onstenk, J.H.A.M. & Wilbrink, B. (1992). *Arbeidsmarktrelevantie van beroepsgerichte volwasseneneducatie: de beoordeling van arbeidsmarktrelevantie in de PBVE-praktijk*. Amsterdam: SCO.
- Oosterhuis, G.J. & Glebbeek, A. (1992). Ras en geslacht bij personeelsselectie. *Mens en maatschappij*, 63, 237-258.
- Vries, B. de (1988). *Het leven en de leer*. Een studie naar de verbinding van leren en werken in de stage. Proefschrift Nijmegen. Nijmegen: ITS.

Manuscript ontvangen 5-12-1994

Definitieve versie ontvangen 11-4-1996

Het duale karakter van beroepsopleidingen en hun waardering door werkgevers

W.Th.J.G. Hoeben

GION: Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling, Rijksuniversiteit Groningen

ABSTRACT

Employers in the building and metal industries, in nursing and welfare, and of clerical staff were asked to provide criteria of personnel selection in order to explain possible differences in job chances for qualified students in vocational training. They were also asked to evaluate four types of vocational training in these sectors of the labor market. The types of training which were evaluated are the primary and the secondary training in the apprentice system and the short and the long vocational training in regular full-time education. Training in the apprentice system is characterized by the dual components of practical work and theoretical education, the practical work being the main component; regular full-time vocational training is mainly theoretical education. The value judgements of the employers correlated highly with the qualification level of the training, and not at all with the amount of practical work during the training, with the job ranking or with the specific job demands. Nevertheless, working experience is a criterion that more often than training qualification tips the scale in personnel selection.

1. INLEIDING

Het beleid verwacht aanzienlijke effecten van dualisering van het beroepsonderwijs, van schaalvergroting, van een betere marktorientatie en van betere kwaliteitszorg van de afzonderlijke opleidingen en van meer overleg en gezamenlijk gedragen verantwoordelijkheid ('co-maker-ship') tussen de opleidingen en het bedrijfsleven. De effecten worden gezocht in de aansluiting tussen beroepsonderwijs en arbeidsmarkt en dus in de kansen van schoolverlaters op de arbeidsmarkt (het extern rendement van de beroepsopleidingen).

Voor het effect van dualisering kunnen we kijken naar de verschillen in arbeidsmarktpositie tussen de meer en de minder gedualiseerde beroepsopleidingen. Uit onderzoek blijkt dat leeftijd, sekse en vooropleiding van schoolverlaters en de waardering door werkgevers van typen van beroepsopleidingen de helft van alle verschillen in arbeidsmarktpositie verklaren. Beroepsopleidingen worden door werkgevers niet gunstiger gewaardeerd naarmate zij meer zijn gedualiseerd. Deze onderzoeksresultaten geven aanleiding de waardering van werkgevers in detail te bekijken. Alvorens dat te doen worden de genoemde onderzoeksresultaten samengevat.

Het betreft resultaten van een onderzoek naar de relatie tussen onderwijs en arbeidsmarkt¹, dat in 1990 en 1991 zowel is uitgevoerd bij schoolverlaters uit Kort Middelbaar Beroepsonderwijs (KMBO), primair en voortgezet Leerlingwezen (LLW) en Middelbaar Beroepsonderwijs (MBO) in de bouw, de metaal, de verzorging en administratieve beroepen, als bij werkgevers uit deze sectoren (Hoeben, 1993a²; 1993b).

In 1990 hebben 2814 schoolverlaters van opleidingen in de genoemde sectoren een vragenlijst ingevuld over hun vooropleiding, hun bezigheden een jaar na het verlaten van de school (werk, vervolgopleiding, militaire dienst en dergelijke), hun functie, hun inkomen, en wat zij vonden van de beroepsopleiding waarvoor zij een diploma hadden gehaald. Gedeeltelijk zijn de

Adres: GION, postbus 1286, 9701 BG Groningen.

Dit is het laatste artikel behorende bij het themanummer.

gegevens overgenomen van de periodieke RUBS- en CIBB-enquêtes (ongeveer de helft van de steekproef); om de steekproef landelijk representatief te maken zijn de gegevens aangevuld met enquêtes onder schoolverlaters uit andere opleidingen en andere regio's.

In 1991 hebben 458 werkgevers, c.q. personeelsfunctionarissen van bedrijven en instellingen in dezelfde sectoren informatie gegeven over eisen die zij aan functies stellen, hun waardering voor het algemene niveau van beroepsopleidingen en voor de aansluiting ervan bij de functie-eisen, en enkele andere gegevens over werving en selectie van personeel.

De gegevens van deze twee steekproeven maakten het mogelijk conclusies te trekken over de relatieve invloed van verschillende factoren op het externe rendement van het beroepsonderwijs, met name op de aansluiting op de arbeidsmarkt. Die factoren kunnen worden gevonden op de niveaus van de schoolverlaters zelf, van de scholen waar zij een diploma hebben gehaald, en van de deelarbeitsmarkten waarvoor zij zijn opgeleid. Om de gegevens van de steekproeven van schoolverlaters en van werkgevers een geneste structuur te geven zijn deelarbeitsmarkten technisch gedefinieerd als het product van opleidingstypen en beroepssectoren: (K)MBO en LLW leiden op voor verschillende beroepen per sector. Invloeden van de deelarbeitsmarkten werden in het bijzonder gevonden in de gemiddelde waarderingen van werkgevers voor de opleidingen. Aan het begin van het onderzoek zijn over effecten van beroepsopleidingen en van deelarbeitsmarkten voorspellingen gedaan op basis van Thurows model van de 'labour queue' en Glebbeeks model van relatieve trainingskosten (Glebbeek, 1988; Thurow, 1975).

Samenvattend kon het volgende worden geconcludeerd. Leeftijd, sekse en vooropleiding van schoolverlaters verklaren meer dan een derde van de verschillen in hun arbeidsmarktposities; het niveau en de aard van de baan verklaren ongeveer 10% van de verschillen.

Meer of minder dualisering is een verschil tussen opleidingstypen: namelijk tussen het primaire en het voortgezette leerlingwezen aan de ene kant en het voltijds middelbaar beroepsonderwijs in korte of lange varianten aan de andere kant. Kansen op de arbeidsmarkt van schoolverlaters en extern rendement van opleidingen hangen niet met dit verschil in dualisering samen. Schoolverlaters van het KMBO hebben de minste en die van het MBO de beste kansen, de schoolverlaters van het primair en het voortgezette leerlingwezen zitten daar in die volgorde tussenin (Hoeben, 1993a).

Waarderingen van de opleidingen door werkgevers volgen in grote lijnen hetzelfde patroon: het KMBO krijgt de laagste waardering, het MBO de hoogste en het LLW zit daar tussenin. Uit de gegevens bleek verder dat de arbeidsmarktposities van schoolverlaters inderdaad beter zijn, als het type opleiding dat ze hebben gevolgd, door werkgevers gunstiger wordt gewaardeerd dan andere typen opleiding. Waarderingen van de opleidingen door werkgevers verklaren vrijwel alle verschillen tussen scholen in extern rendement en de resterende verschillen tussen de deelarbeitsmarkten; kenmerken van scholen zelf voegen daar niets aan toe. De genoemde kenmerken van schoolverlaters en van werkgevers blijken de helft van de verschillen in arbeidsmarktpositie tussen schoolverlaters te verklaren. Extern rendement blijkt niet zozeer een effect van afzonderlijke opleidingen op schoolniveau, maar een arbeidsmarkteffect dat voor alle opleidingen van eenzelfde type en voor eenzelfde beroepssector hetzelfde is, althans voor zover met instroomkenmerken van de leerlingen rekening wordt gehouden (Hoeben, 1993a).

Probleemstelling

De waarderingen van beroepsopleidingen door de werkgevers hebben zoveel invloed op de arbeidsmarktpositie van schoolverlaters, dat zij alle verschillen tussen scholen en tussen deelarbeitsmarkten verklaren. Dit gegeven is een reden om dieper op deze waarderingen in te gaan. Dat doen we aan de hand van de volgende vragen:

1. Hoe waarderen werkgevers de opleidingen voor de afzonderlijke beroepssectoren en functie-niveaus en hoe verhouden de waarderingen zich tot de praktijkcomponent van de beroepsopleidingen?
2. Hoe verhouden de waarderingen van werkgevers voor opleidingen zich tot functie-eisen en tot de feitelijke selectiecriteria en wervingsactiviteiten?
3. Welke beleidsimplicaties kunnen uit de conclusies worden getrokken?

2. METHODE

Steekproef en respons

Uit het Centraal Vestigingsregister is een naar bedrijfsgrootte en bedrijfsklasse gestratificeerde aselechte steekproef getrokken van 1000 bedrijven: in de bouw, de metaal, de verzorging en de administratieve sector elk 250 bedrijven en instellingen. Aan de personeelsafdeling werd een vragenlijst gestuurd. De respons was 46%, waarvan 36% in de bouw, 44% in de metaal, 58% in de verzorging en 45% in de administratieve sector. Bij grote bedrijven of instellingen was de respons hoger dan gemiddeld.

Vragenlijst

Aan de hand van drie globaal omschreven karakteristieke functies op verschillend niveau werd informatie gevraagd. Het eerste niveau betreft eenvoudige uitvoerende functies, in de bouw- en de metaalsector het niveau van hulpmiddelen en aankomende vaklieden. Het tweede niveau betreft geschoolde functies, de (all-round) vaklieden in de bouw- en de metaalsector. Het derde niveau betreft functies in het middenkader of van gespecialiseerde vaklieden, die tevens toezicht houden op anderen. De niveaus en de omschrijvingen zijn ontleend aan Huijgen e.a. (1983); het betreft functies op het tweede, derde en vierde niveau van hun functieniveau-indeling.

Personeelsfunctionarissen werd gevraagd informatie te geven over eisen die zij aan functies van deze niveaus stellen, hun waardering te geven voor het algemene niveau van beroepsopleidingen, hun waardering voor de aansluiting ervan bij de functie-eisen, en enkele andere gegevens over werving en selectie van personeel. Per functieniveau is gevraagd naar eigenschappen en vaardigheden die voor een goede vervulling van die functie van belang worden geacht, zoals vakkennis en -inzicht, kunnen omgaan met machines en apparatuur, nauwkeurigheid, zelfstandigheid, samenwerking, omgangsvormen, het vermogen om nieuwe dingen te leren en het vermogen om leiding te geven.

De waardering van het niveau van de beroepsopleidingen en hun aansluiting aan de eisen, die zij stellen aan functies van verschillend niveau, is geoperationaliseerd door twee vier-punts schalen die per beroepssector en per functieniveau een waardering voor elke opleiding geven (resp. een waardering voor het algemene niveau van de opleiding gelet op het functieniveau en een waardering voor de aansluiting van de opleiding bij de specifieke functie-eisen). Beide vier-punts schalen zijn in somscores gecombineerd. De somscores per sector en functieniveau lopen van 2 (opleiding sluit zeer slecht aan ...) tot 8 (... sluit zeer goed aan bij de functie-eisen).

3. RESULTATEN

Begonnen wordt met een beschrijving van de opleidingswaarderingen zelf (vraag 1), waarna de relaties worden beschreven van de opleidingswaarderingen met de functie-eisen, met wervingsactiviteiten en met criteria voor personeelselectie (vraag 2). De beantwoording van vraag 3 is onderdeel van de discussie.

3.1 De waardering van opleidingen door werkgevers

De waarderingen die de werkgevers (personeelsfunctionarissen) aan de onderscheiden opleidingen geven, worden per beroepssector beschreven.

Waarderingen voor opleidingen in de bouwsector

In figuur 1 zijn de waarderingen voor de opleidingen in de bouwsector gepresenteerd. Er zijn duidelijke verschillen. Het KMBO wordt in alle omstandigheden lager gewaardeerd dan de overige opleidingstypes. Voor de twee laagste functieniveaus van aankomende en allround vaklieden kan het primair LLW en het voortgezet LLW van de Stichting Vakopleiding Bouwbedrijf (SVB) concurreren met het Middelbaar Technisch Onderwijs (MTO). Daarentegen wordt

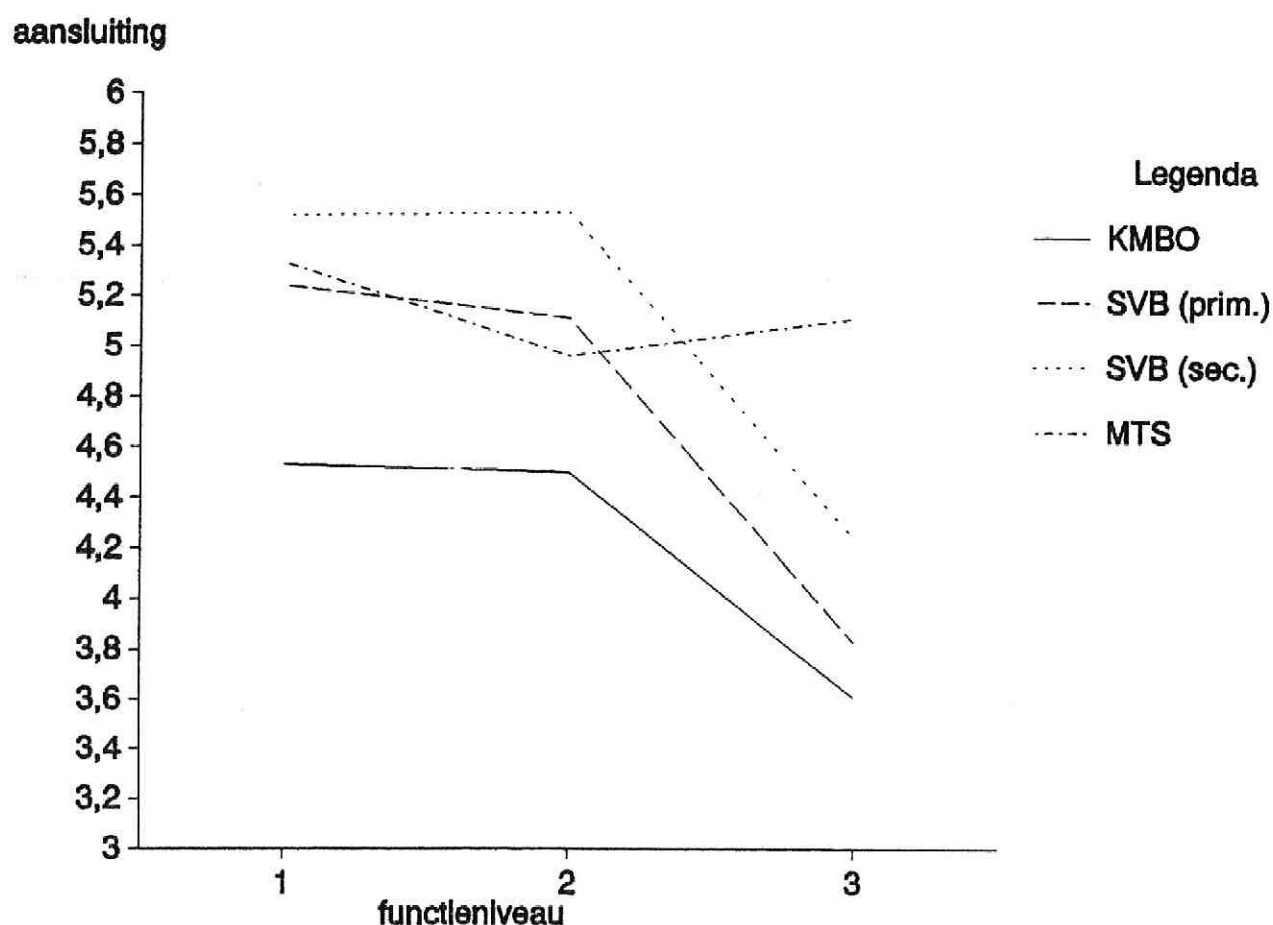


Fig. 1. Gemiddelde waardering van opleidingen voor aansluiting aan de functie-eisen naar functieniveau in de bouwsector (1 = aankomende vaklieden; 2 = vaklieden; 3 = middenkader of gespecialiseerde vaklieden).

voor het functieniveau 3 van gespecialiseerde vaklieden en vaktechnici het MTO aanzienlijk hoger gewaardeerd dan de duale opleidingen.

Waarderingen voor opleidingen in de metaalsector

In figuur 2 zijn de waarderingen voor opleidingen in de metaalsector gepresenteerd. Vergeleken met de bouwsector zijn de waarderingen in het algemeen gunstiger. Ook in de metaalsector krijgt het KMBO systematisch de laagste waardering, kunnen de duale opleidingen in het LLW van de Stichting Opleidingen Metaal (SOM) voor de lagere functieniveaus concurreren met het MTO, terwijl voor het derde functieniveau van gespecialiseerde vaklieden en vaktechnici het MTO aanzienlijk hoger wordt gewaardeerd dan de duale opleidingen.

Waarderingen voor opleidingen in de sector verzorging

In figuur 3 zijn de waarderingen voor de opleidingen in de sector verzorging weergegeven. In de vragenlijst is voor het Middelbaar Dienstverlenend en Gezondheids Onderwijs (MDGO) onderscheid gemaakt tussen specifiek op de sector verzorging gerichte MDGO-V en overig MDGO. Tevens is de opleiding voor assistenten (INTAS) er bij betrokken. In de sector verzorging blijkt een met de sectoren bouw en metaal vergelijkbaar lage waardering voor het KMBO. Duale opleidingstrajecten in de verzorging worden echter systematisch hoger gewaardeerd dan het voltijdsonderwijs: ook het voortgezette LLW van de Stichting Opleiding Verzorgende en Dienstverlenende Beroepen (OVDB) wordt op alle functieniveaus hoger gewaardeerd dan het overige MDGO en alleen op het tweede functieniveau lager dan het MDGO-V. In tegenstelling tot de

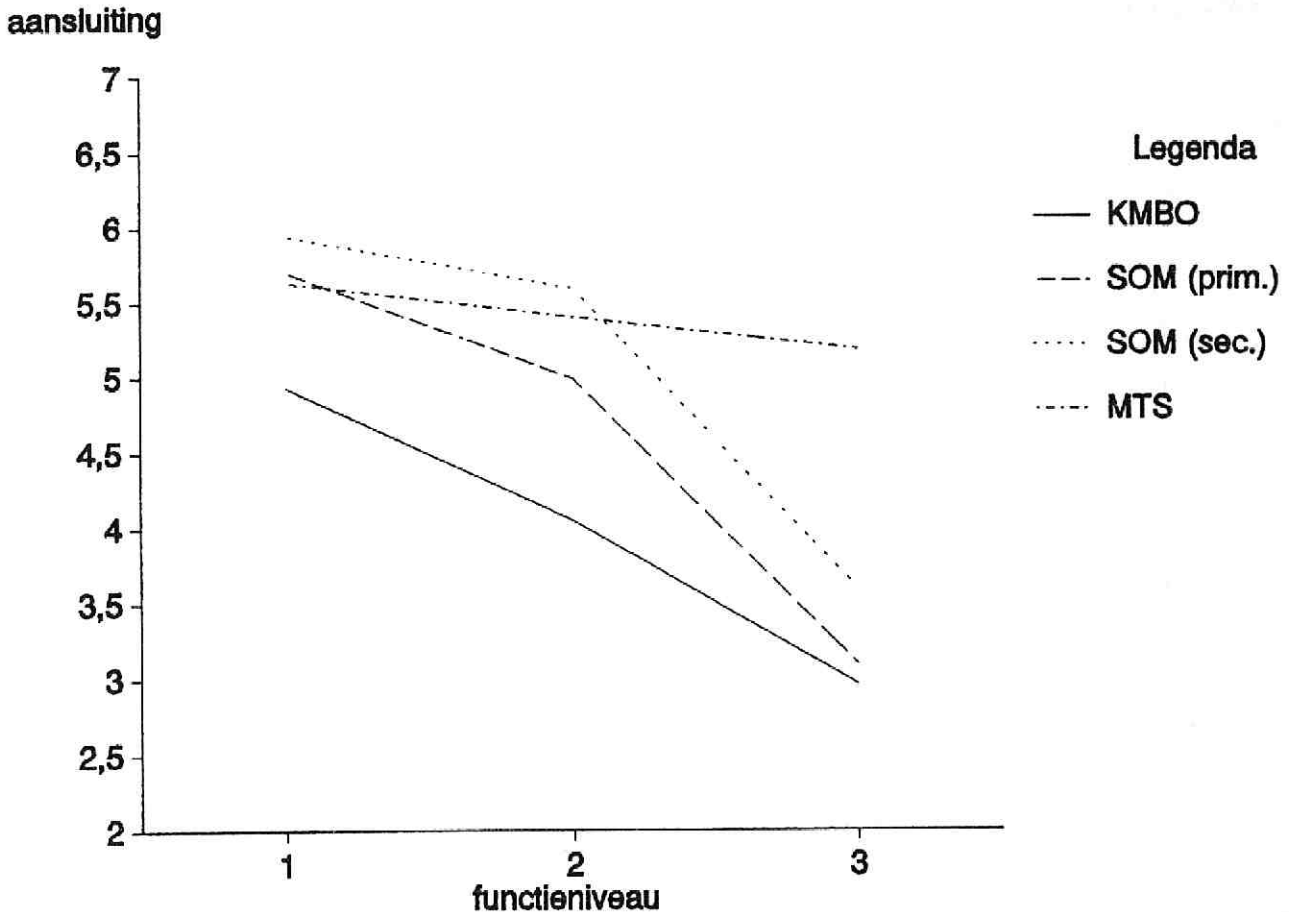


Fig. 2. Gemiddelde waardering van opleidingen voor aansluiting aan de functie-eisen naar functieniveau in de metaalsector (1 = aankomende vaklieden; 2 = vaklieden; 3 = middenkader of gespecialiseerde vaklieden).

bouw en de metaal worden door werkgevers in de sector verzorging de duale opleidingstrajecten in het voortgezette OVDB voor alle functieniveaus op een hoger of een met het MDGO concurrerend niveau gewaardeerd.

Waarderingen voor opleidingen in de administratieve sector

In tabel 4 zijn de waarderingen voor de opleidingen in de administratieve sector gepresenteerd. In de vragenlijst voor de lagere functieniveaus is er op geanticipeerd dat schoolverlaters van administratieve opleidingen op de arbeidsmarkt moeten concurreren met schoolverlaters van het algemeen vormend onderwijs. In de rangordening kan het algemeen vormend onderwijs concurreren met beroepsopleidingen. Zo wordt het MAVO hoger gewaardeerd dan het primaire LLW van de Stichting Landelijk Orgaan ter Bevordering van de Economische en Administratieve Beroepencategorie (ECABO) en kan het concurreren met het KMBO in deze sector. Het HAVO/VWO wordt hoger gewaardeerd dan het voortgezette ECABO en kan concurreren met het Middelbaar Economisch en Administratief Onderwijs (MEAO). In tegenstelling tot de andere beroepssectoren wordt het KMBO op alle functieniveaus hoger gewaardeerd dan het primaire leerlingwezen (ECABO). Ook het MEAO wordt hoger gewaardeerd op alle functieniveaus dan het voortgezette leerlingwezen (ECABO).

Conclusies

Uit de waarderingen van werkgevers in de verschillende beroepssectoren voor de voor hun sector bedoelde opleidingen kan worden afgelezen dat zij de opleidingen niet hoger waarderen

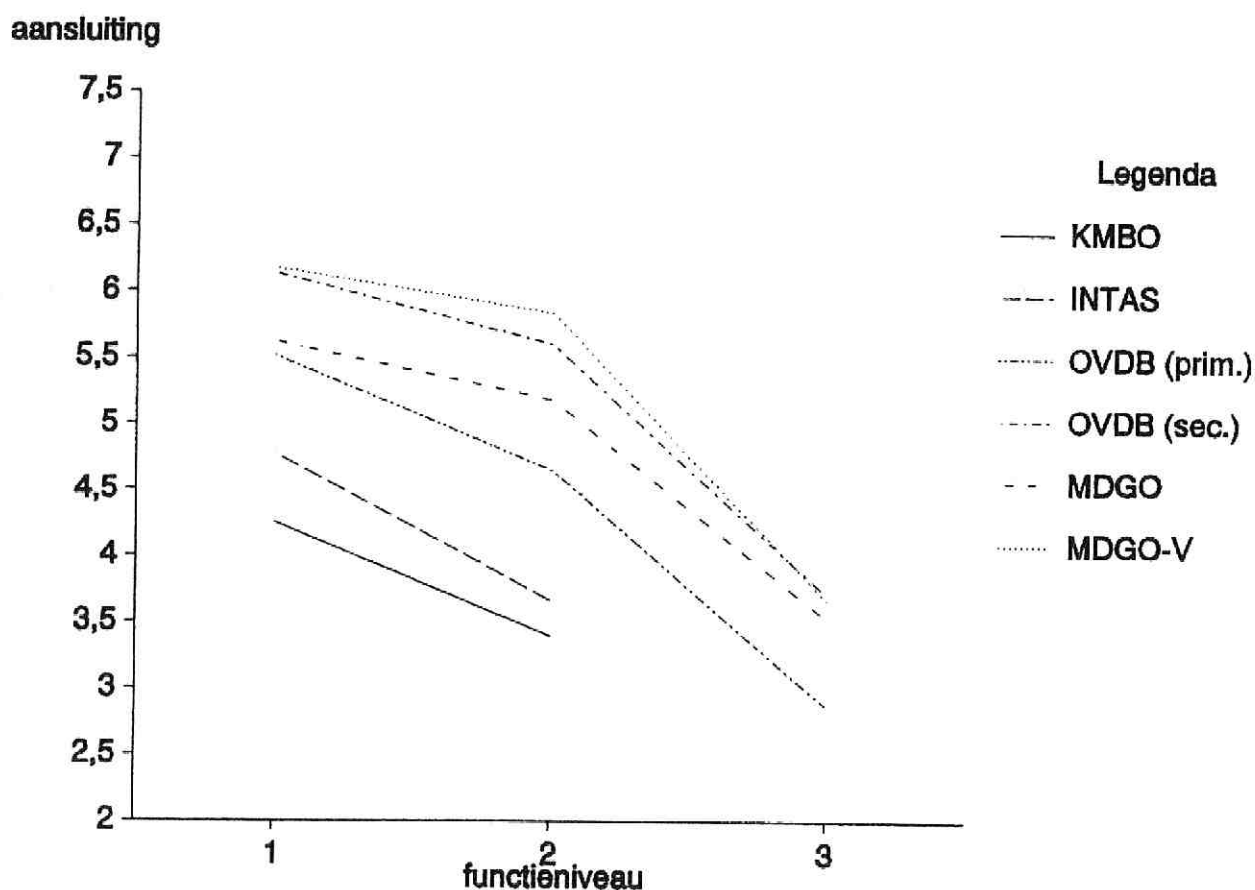


Fig. 3. Gemiddelde waardering van opleidingen voor aansluiting aan de functie-eisen naar functieniveau in de Verzorgingssector (1 = hulpkrachten; 2 = geschoolde functies; 3 = middenkader of gespecialiseerde functies).

naarmate zij meer aandacht besteden aan de praktijkcomponent. Het MBO wint het van de andere beroepsopleidingen of kan op zijn minst voor de lagere functieniveaus concurreren met het voortgezette leerlingwezen. Duale opleidingen winnen het daarentegen meestal van het KMBO. Dientengevolge ontstaat een waarderingsrangorde van KMBO, via primair en voortgezet LLW, naar MBO. Werkgevers hechten kennelijk meer waarde aan het ingeschatte theoretische niveau van de opleidingen dan aan hun duale karakter.

Er zijn echter verschillen tussen sectoren. Het geschetste algemene beeld geldt specifiek voor de sectoren van de bouw en de metaal. In de sector verzorging krijgen gedualiseerde opleidingen in tegenstelling tot de bouw en de metaal ook voor het hoogste functieniveau een relatief hoge waardering. In de administratieve sector worden de duale opleidingen ook voor de lagere functieniveaus lager gewaardeerd dan opleidingen in volle tijd. Zelfs algemeen vormend onderwijs scoort onverwacht hoog voor beroepen in de administratieve sector.

3.2 Verhouding tussen opleidingswaarderingen en functie-eisen, werving en selectie

Om na te gaan hoe de waardering voor de beroepsopleidingen zich verhoudt tot het proces van werving en selectie, zijn de waarderingen in verband gebracht met eisen die werkgevers hanteren voor een goede uitoefening van de functies, en met hun gedrag bij werving en selectie van nieuw personeel (selectiecriteria en zoekkanalen) onderzocht. Met name is van belang of de geconstateerde overeenkomsten en verschillen tussen sectoren in de relaties tussen functieniveaus en waarderingen voor opleidingen worden gereflecteerd door de eisen die aan functies worden gesteld, en door het feitelijke gedrag van werkgevers bij werving en selectie.

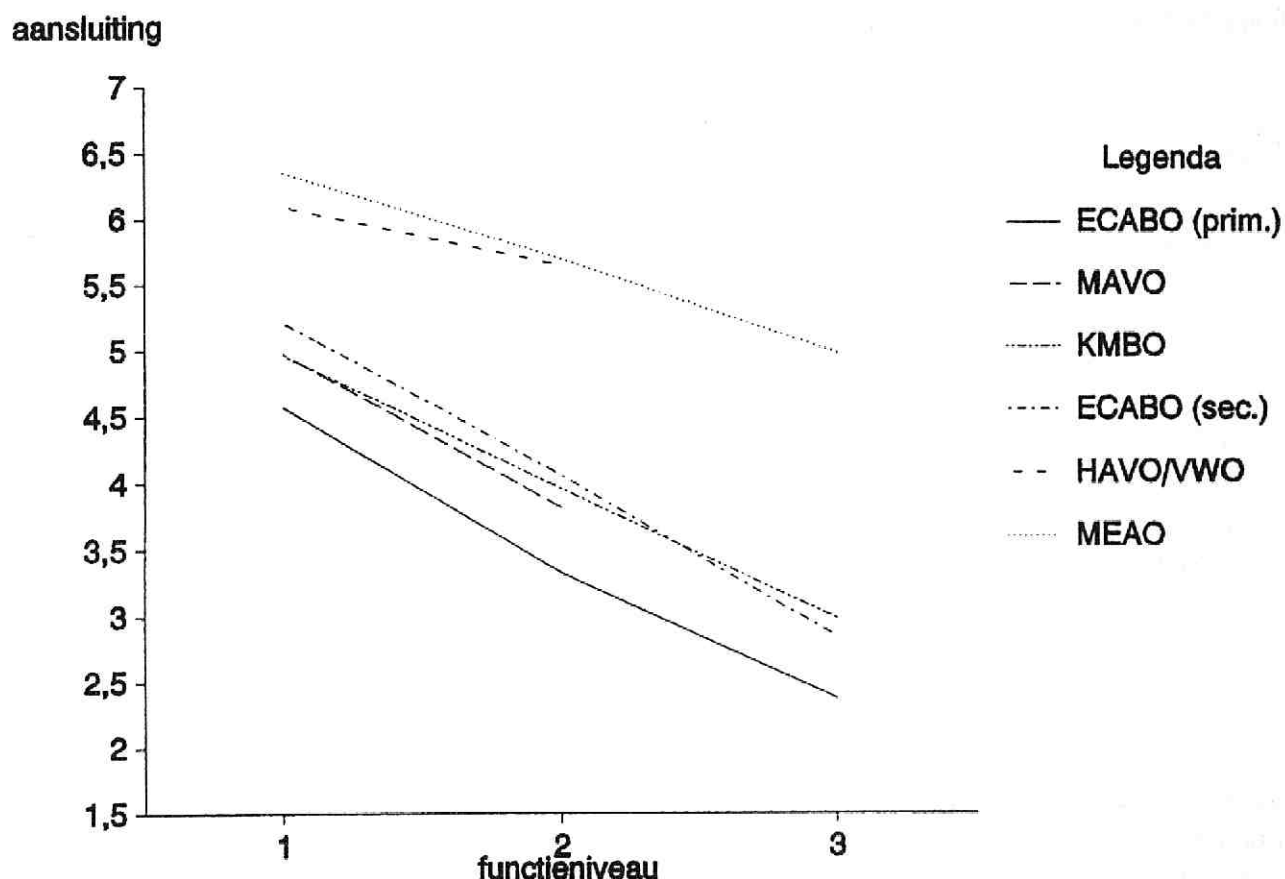


Fig. 4. Gemiddelde waardering van opleidingen voor aansluiting aan de functie-eisen naar functieniveau in de administratieve sector (1 = hulpkrachten; 2 = geschoolde functies; 3 = middenkader of gespecialiseerde functies).

Eisen voor de functie en waarderingen voor opleidingen

De voor de onderscheiden functieniveaus onderzochte kwalificatie-eisen hadden betrekking op: vakkennis en -inzicht, kunnen omgaan met machines en apparatuur, nauwkeurigheid, zelfstandigheid, samenwerking, omgangsvormen, het vermogen om nieuwe dingen te leren en het vermogen om leiding te geven. Uit factoranalyses bleek dat de kwalificatie-eisen voor elke sector en elk functieniveau een eigen factorstructuur hebben. De correlaties tussen het belang van kwalificatie-eisen voor een goede functie-uitoefening en de waardering voor opleidingen bleken echter voor elke sector en elk functieniveau te laag om systematische conclusies te kunnen trekken. Binnen elk van de vier beroepssectoren en voor elk van de drie functieniveaus waren het bovendien telkens andere kwalificatie-eisen waarvan het belang nu eens positief, dan weer negatief samenhang met de waardering van telkens andere opleidingen. In de correlaties tussen de waardering voor de opleidingen en het belang van kwalificatie-eisen viel dan ook geen patroon te ontdekken. Deze gegevens worden hier dan ook niet gerapporteerd (zie daarvoor Hoeben, 1993a).

Het is kennelijk niet het voldoen aan de afzonderlijke kwalificatie-eisen dat voor werkgevers aanleiding is om opleidingen hoger of lager te waarderen. Schulte-Lutter (1992) komt in een analyse op hetzelfde gegevensbestand van werkgeversvoorkeuren voor schoolverlaters tot de conclusie dat het gepercipiëerde algemene niveau van de opleiding de waardering beïnvloedt.

Feitelijke selectie-criteria

Er is ook gevraagd wat voor de werkgever de doorslag heeft gegeven bij de laatst aangenomen medewerker op elk van de drie functieniveaus: de leeftijd, de ervaring of de opleiding van de sollicitanten. Als de waardering voor een opleiding is ingegeven door het opdoen van praktijk-

Tabel 1. Factoren die doorslag gaven bij het aannemen van personeel naar functieniveau en sector.

	Bouw	Verzorging	Metaal	Administratie
Functieniveau 1				
leeftijd	9.8%	3.1%	6.0%	—
ervaring	16.4%	44.9%	29.9%	36.3%
opleiding	11.5%	8.2%	22.4%	24.2%
N	61	98	67	91
Functieniveau 2				
leeftijd	3.4%	0.8%	3.8%	1.0%
ervaring	57.6%	47.7%	38.0%	31.0%
opleiding	6.8%	20.0%	20.3%	39.0%
N	59	130	79	100
Functieniveau 3				
leeftijd	3.4%	—	—	3.8%
ervaring	50.0%	65.9%	45.7%	53.8%
opleiding	24.1%	16.5%	41.4%	29.5%
N	58	91	70	78

ervaring in de opleiding, dan zou, simpel geredeneerd, de ervaring van sollicitanten de doorslag moeten geven in het antwoord op deze vraag; als de waardering van een opleiding is ingegeven door het ingeschatte theoretische niveau van de opleiding, dan zou de opleiding de doorslag moeten geven. Uitgaande van de beschreven waarderingen van opleidingen zou ervaring het moeten winnen van opleiding voor alle drie functieniveaus in de verzorging en voor de twee laagste functieniveaus in de bouw en de metaal, terwijl opleiding het voor alle functieniveaus zou moeten winnen van ervaring in de administratieve sector en voor het hoogste functieniveau in de sectoren bouw en metaal.

Tabel 1 laat zien dat op één uitzondering na de ervaring het wint van de opleiding op alle functieniveaus in alle vier sectoren. De uitzondering betreft het tweede functieniveau in de administratieve sector. In drie gevallen in de sectoren bouw (niveau 1) en metaal (niveau 1 en 3) wint de ervaring het slechts met een kleine marge. In tegenstelling tot de voorspelling wint de ervaring in de administratieve sector op twee functieniveaus en in de bouw op het hoogste niveau met een verhoudingsgewijs grote marge van de opleiding, terwijl op het hoogste niveau in de metaal de ervaring het eveneens, maar met een kleine marge wint.

Alleen voor de sector verzorging wordt de eerder gerapporteerde waardering voor opleidingen met een praktijkcomponent gereflecteerd door de feitelijke selectiecriteria voor alle functieniveaus. Voor het hoogste functieniveau in de overige sectoren en voor het laagste administratieve functieniveau is de vereiste ervaring kennelijk iets anders dan ervaring die in een gedualiseerde opleiding kan worden opgedaan; we moeten waarschijnlijk denken aan eventueel langdurige werkervaring in een andere baan.

Zoekkanalen bij werving

Waardering voor de praktijkcomponent van opleidingen en de doorslag laten geven door ervaring bij de selectie van personeel vraagt om zoekkanalen bij de werving die zoveel mogelijk informatie geven over de praktische ervaring van kandidaten. Zoekkanalen waarbij een bedrijf of instelling over eigen interne informatie beschikt (zoals bij interne werving, werving via een eerste contract met uitzendbureaus, werving van mensen die een leerovereenkomst met of een stage bij het bedrijf hebben), hebben dan de voorkeur: de werkgever weet uit eigen ervaring hoe de kandidaat functioneert. De formele wegen via arbeidsbureau en personeelsadvertenties lijken de minst nauwkeurige informatie over de praktische ervaring op te leveren. Zoekkanalen

in het informele circuit via kennissen, via open, ongevraagde sollicitaties en via scholen liggen wat betreft de informatie over de kandidaten tussen deze twee uitersten in.

Aan de werkgevers is daarom ook gevraagd langs welke kanalen zij personeel voor elk van de drie functieniveaus zochten (1 = meestal, 2 = soms, 3 = zelden of nooit). In overeenstemming met de kwalificatieprofielen van opleidingen zouden voor functieniveaus 1 en 2 in de bouw en de metaal en voor alle functieniveaus in de verzorging de verschillende zoekkanalen waarbij het bedrijf al over informatie vooraf beschikt, het frequentste moeten worden gebruikt, en de formele zoekkanalen het minst frequent. In overeenstemming met de feitelijke selectiecriteria zou deze voorspelling ook kunnen worden gedaan voor het hoogste functieniveau in de bouw, de metaal en de administratie.

De voorspelde frequentieverschillen blijken niet te kloppen. Personeel wordt het meest geworven via informele zoekkanalen (waarbij vooral het kenniscircuit belangrijk is en informatie via scholen nauwelijks een rol speelt); dat springt het meeste naar voren bij de sectoren

Tabel 2. Gemiddelde frequentie van zoekkanalen bij de werving van personeel per functieniveau in de Bouw.

Functioniveau	1	2	3
Formele Zoekkanalen			
arbeidsbureau	2.9	2.1	2.2
advertenties	2.0	2.3	2.6
Informele Kanalen			
kennissen	1.7	2.1	2.0
open sollicitaties	2.5	2.3	2.5
via scholen	2.6	2.8	2.7
Interne Informatie			
interne werving	2.6	1.7	2.2
uitzendbureau	2.8	2.9	2.6
leerovereenkomsten	2.2	2.5	2.9
stagiairs	2.3	2.5	2.6

Tabel 3. Gemiddelde frequentie van zoekkanalen bij de werving van personeel per functieniveau in de Metaal.

Functioniveau	1	2	3
Formele Zoekkanalen			
arbeidsbureau	2.0	1.9	2.1
advertenties	1.8	2.2	2.6
Informele Kanalen			
kennissen	1.9	1.7	1.6
open sollicitaties	2.3	2.2	2.2
via scholen	2.6	2.6	2.6
Interne Informatie			
interne werving	2.2	2.2	2.4
uitzendbureau	1.8	2.1	2.2
leerovereenkomsten	2.6	2.5	2.8
stagiairs	2.5	2.5	2.5

Tabel 4. Gemiddelde frequentie van zoekkanalen bij de werving van personeel per functieniveau in de sector Verzorging.

Functieniveau	1	2	3
Formele Zoekkanalen			
arbeidsbureau	2.2	2.2	2.0
advertenties	1.7	2.3	2.6
Informele Kanalen			
kennissen	2.1	1.7	1.7
open sollicitaties	1.7	1.8	2.2
via scholen	2.6	2.5	2.9
Interne Informatie			
interne werving	2.3	2.2	2.5
uitzendbureau	2.4	2.5	2.8
leerovereenkomsten	2.4	2.4	2.9
stagiairs	2.0	2.1	2.8

Tabel 5. Gemiddelde frequentie van zoekkanalen bij de werving van personeel per functieniveau in de sector Administratie.

Functieniveau	1	2	3
Formele Zoekkanalen			
arbeidsbureau	2.0	1.8	1.8
advertenties	1.8	2.4	2.6
Informele Kanalen			
kennissen	2.5	1.7	1.7
open sollicitaties	2.4	2.5	2.5
via scholen	2.9	2.8	2.9
Interne Informatie			
interne werving	1.9	2.6	2.7
uitzendbureau	2.2	2.3	2.5
leerovereenkomsten	2.7	2.8	2.9

metaal en verzorging (tabellen 3 en 4); in de administratie worden de formele zoekkanalen het meest frequent genoemd (tabel 5). Kanalen waarbij men over interne informatie vooraf beschikt, zijn het minste in trek bij de werving van personeel; van de vier in aanmerking komende zoekkanalen voor de drie functieniveaus in de vier beroepssectoren (in totaal 48) hebben er slechts drie een hoge gemiddelde frequentie: interne werving heeft een keer in de bouw (niveau 2, tabel 2) en een keer in de administratie (niveau 2, tabel 5) een hoge gemiddelde frequentie en het uitzendbureau een keer in de metaal (niveau 1, tabel 3).

Concluderend kan worden gesteld dat personeel in het algemeen niet wordt geworven via die kanalen die de meest betrouwbare informatie geven over de praktijkervaring van de betrokkenen. Men gaat nog het meest af op informatie van kennissen. Met name de kanalen waarbij een bedrijf of instelling reeds over eigen informatie beschikt, worden weinig gebruikt. Een verklaring hiervoor zou kunnen liggen in de mogelijkheid dat de meeste van de betrokken bedrijven weinig uitzendkrachten, leerlingen of stagiairs hebben. Het onderzoek kan hier geen uitsluitsel over geven.

4. DISCUSSIE: CONCLUSIES EN BELEIDSIMPLICATIES

Schoolverlaters van het KMBO hebben de slechtste en die van het MBO de beste kansen op de arbeidsmarkt; de schoolverlaters van het primair en het voortgezetleerlingwezen zitten daar in die volgorde tussenin. Scholen voor beroepsonderwijs hebben nauwelijks toegevoegde waarde bij de verklaring van de arbeidsmarktposities van hun schoolverlaters, vergeleken met hun leeftijd, geslacht en vooropleiding, en vergeleken met kenmerken van de deelarbeitsmarkten, met name de waarderingen van opleidingen door werkgevers. Dat waren de voornaamste conclusies van het onderzoek naar arbeidsmarktposities van schoolverlaters uit het secundaire beroepsonderwijs die de aanleiding vormden om de waarderingen van opleidingen door werkgevers in meer detail te onderzoeken. Het onderzoek naar opleidingswaarderingen en het proces van werving en selectie geeft aanleiding om daar de volgende conclusies aan toe te voegen.

De waardering van opleidingen

De waardering van een opleiding door de werkgevers vanuit de aan functies gekoppelde kwalificatie-eisen kunnen worden geïnterpreteerd als een kwalificatieprofiel van de betrokken beroepsopleiding. De rangordes in de waardering van beroepsopleidingen volgen een vergelijkbaar patroon als de arbeidsmarktkansen van gediplomeerde schoolverlaters. Er is een rangorde van opleidingswaarderingen die loopt van KMBO, via primair en voortgezet LLW, naar MBO als de hoogst gewaardeerde opleiding. Uit de waarderingen van werkgevers in de verschillende beroepssectoren kan worden afgeleid dat werkgevers de opleidingen hoger waarderen naarmate het ingeschatte niveau van de opleidingen hoger is, en niet zozeer naarmate de opleidingen meer aandacht besteden aan de praktijkcomponent, c.q. gedualiseerd zijn.

Specifieke functie-eisen hangen niet met waardering voor de opleidingen samen. Functieniveaus beïnvloeden de waardering daarentegen wel. Voor de laagste twee functieniveaus in de sectoren bouw en metaal en voor alle drie functieniveaus in de verzorging geven werkgevers een waardering aan duale opleidingen, die kan concurreren met het MBO. Duale opleidingen worden zodoende meestal hoger gewaardeerd dan het KMBO. Voor alle administratieve functieniveaus en voor het hoogste functieniveau in de bouw en de metaal worden duale opleidingen lager gewaardeerd dan andere opleidingen. Deze verschillen in waardering van opleidingen voor bepaalde functies worden slechts gedeeltelijk weerspiegeld in de feitelijk gehanteerde selectiecriteria en zoekkanalen.

Selectiecriteria

Op functieniveaus 1 en 2 wordt personeel in de sectoren bouw, metaal en verzorging in overeenstemming met de bij die niveaus behorende waarderingen van de opleidingen vaker op basis van ervaring dan op basis van opleiding aangesteld. In tegenstelling tot de bijbehorende opleidingswaarderingen gebeurt dat ook op functieniveau 3 in de sectoren bouw en metaal; voor functies van dit niveau in deze sectoren is een MTS-opleiding kennelijk belangrijk, maar MTS-ers met ervaring krijgen de voorkeur.

Hoewel meer gedualiseerde opleidingen binnen de sector administratie systematisch lager worden gewaardeerd dan minder gedualiseerde of zelfs algemeen vormende opleidingen, wordt bij het aanstellen van personeel in deze sector verhoudingsgewijs vaak de doorslag gegeven door ervaring en wint de ervaring het van de opleiding bij aanstellingen op twee van de drie functieniveaus. Deze resultaten zouden er op kunnen wijzen dat praktijkervaring ook voor administratieve beroepen belangrijk is, maar dat de opleidingen met een praktijkcomponent in deze sector worden gediskwalificeerd.

Zoekkanalen

Opvallend is dat het informele kennissencircuit in alle sectoren als meest frequent kanaal naar voren komt om nieuw personeel te zoeken. Kanalen waarbij het bedrijf of de instelling vooraf over interne informatie beschikt, worden door de bank genomen het minst frequent gebruikt. De gegevens over het zoekgedrag zijn niet overeenkomstig de specifieke voorspellingen die zijn

gedaan op basis van de vooronderstelling dat zo betrouwbaar mogelijke informatie over praktische ervaring bij de werving voorop staat.

De gegevens over het zoekgedrag zijn globaal echter wel in overeenstemming met het accent op ervaring als doorslaggevende factor bij de selectie van personeel. Formele zoekkanalen zoals arbeidsbureaus en advertenties leggen het vrijwel altijd af van het informele circuit van kennissen. Evenals de gegevens over de selectie van personeel geven de gegevens over zoeken van personeel aanleiding tot een relativering van de kwalificatieprofielen van opleidingen, waarbij het niveau van de opleiding zwaarder lijkt te wegen dan de praktijkcomponent. Als puntje bij paaltje komt, hebben werkgevers de neiging om de praktische ervaring zwaarder te laten wegen dan (het niveau van) de opleiding.

Deze discrepantie zou verklaard kunnen worden door het gegeven dat op de arbeidsmarkt recente schoolverlaters moeten concurreren met mensen met vergelijkbare diploma's die reeds arbeidservaring hebben. Als werkgevers de keuze hebben tussen mensen met een vergelijkbare opleiding, waarbij de ene wel en de andere geen relevante arbeidservaring heeft, ligt het voor de hand dat praktische ervaring een groter accent krijgt in het feitelijk gedrag van werving en selectie.

Beleidsimplicaties en discussie

De onderzoeksresultaten roepen twijfels op over de effecten die door beleid en sociale partners worden verwacht van dualisering in het beroepsonderwijs, namelijk een betere aansluiting van het onderwijs op de arbeidsmarkt. De beste kansen op de arbeidsmarkt zijn immers verbonden met de niet gedualiseerde MBO-opleidingen; dat blijkt uit de feitelijke arbeidsmarktposities van schoolverlaters; het blijkt ook uit de waarderingen van de verschillende opleidingstypen die werkgevers hanteren bij hun werving en selectie van personeel. Ervaring geeft weliswaar vaker de doorslag dan opleiding bij de feitelijke selectie; deze ervaring heeft echter betrekking op de in een beroepsloopbaan opgedane werkervaring en niet op de praktijkcomponent van een opleiding.

De klacht dat schoolverlaters uit het beroepsonderwijs geen hamer meer kunnen vast houden, wordt duidelijk niet weerspiegeld door de waardering voor de meer theoretische vergeleken met de meer gedualiseerde opleidingen. Uitzonderingen daargelaten geven werkgevers de voorkeur aan zo hoog mogelijk opgeleid personeel ook voor de lagere functies; daarbij lijken ze een eventuele onderbenutting van de kwalificaties van hun werknemers voor lief te nemen. Het verdringingseffect op de arbeidsmarkt is zodoende een mechanisme dat niet alleen speelt bij sollicitanten die geen baan op hun eigen niveau kunnen vinden en naar lagere functies zoeken. Werkgevers lijken eveneens een actieve rol te spelen bij de verdringing door een voorkeur te geven aan personeel dat te hoog is geschoold voor met name de lagere functies.

Deze onderzoeksresultaten roepen twijfel op of de standpunten van vertegenwoordigers van de werkgevers in situaties van overleg over de aansluiting van onderwijs en arbeid wel een goede weerspiegeling vormen van het feitelijke gedrag en de feitelijke voorkeuren van hun achterban bij de werving en selectie van personeel. Als standpunten van werkgeversvertegenwoordigers niet overeenkomen met gedrag en voorkeuren van hun achterban, wordt tevens de vraag opgeroepen of enkele andere beleidsinstrumenten om de aansluiting van het beroepsonderwijs op de arbeidsmarkt te verbeteren, effectief zullen zijn. Het betreft vooral instrumenten waar werkgeversvertegenwoordigers namens het bedrijfsleven optreden, zoals de landelijke kwalificatiestructuur en het co-makenschap tussen opleidingen en bedrijfsleven. Deze beleidsinstrumenten verwachten namelijk gunstige effecten van veel en intensief overleg tussen de opleidingen en landelijke en regionale werkgeversvertegenwoordigers.

Als werkgevers de beroepsopleidingen blijven waarderen zoals in het onderzoek uit hun gedrag en voorkeuren is gebleken en als zij deze waarderingen bij hun werving en selectie een rol laten spelen, dan is het de vraag of de initiatieven die zijn genomen om de aansluiting van beroepsonderwijs op de arbeidsmarkt te verbeteren, veel effect zullen hebben op verbeterde kansen van schoolverlaters op de arbeidsmarkt.

NOTEN

1. Het onderzoek is mogelijk gemaakt door een subsidie van het Instituut voor Onderzoek van Onderwijs SVO (SVO-project 9042).
2. Het volledige rapport van het onderzoek of een uitgebreide samenvatting ervan, kan bij het GION (telefoon 050-636635) worden besteld onder vermelding van auteur en titel (samenvatting: f 10,00 rapport: f 32,50).

LITERATUUR

- Glebbeeck, A.C. (1988). De arbeidsmarktpositie van opleidingen. *Tijdschrift voor arbeidsvraagstukken*, 4, 75-89.
- Hoeben, W.Th.J.G. (1993a). *Beroepsonderwijs en arbeidsmarkt. De arbeidsmarktpositie van schoolverlaters uit verschillende vormen van secundair beroepsonderwijs*. (SVO-project 9042). Groningen: RION.
- Hoeben, W.Th.J.G. (1993b). Arbeidsmarktpositie van schoolverlaters secundair beroepsonderwijs: leeftijd, diploma en geslacht bepalen kansen op de arbeidsmarkt. *Didactief*, 23, 6, 5-6.
- Huijgen, F., Riesewijk, B.J.P. & Conen, G.J.M. (1983). *De kwalitatieve structuur van de werkgelegenheid in Nederland. Bevolking in loondienst en functieniveaustuctuur in de periode 1960-1977* (N.P.A.O.-publicatie nr 17). Den Haag: Staatsuitgeverij.
- Scheerens, J. (1989). *Wat maakt scholen effectief? Samenvatting en analyse van onderzoeksresultaten* (Balans van onderwijsonderzoek). 's-Gravenhage: SVO.
- Schulte-Lutter, S. (1992). *Hoe hoger hoe beter? Een studie naar voorkeuren van personeelsselecteurs voor sollicitanten met opleidingen van een bepaald niveau*. Doctoraalscriptie. Groningen: Sociologisch Instituut RUG.
- Thurow, L. (1975). *Generating inequality*. New York: MacMillan.

Manuscript ontvangen 5-12-1994

Definitieve versie ontvangen 11-4-1996

Impliciet versus expliciet onderwijs in bedrijfseconomische modellen

A.T.J. Vernooij

Instituut voor de Lerarenopleiding, Universiteit van Amsterdam

ABSTRACT

In Business Economics the teaching of accounting focusses on the transfer of knowledge about separate topics without much attention for their integration. As these topics are derived from different subdisciplines, which use conceptual models that are not consistent with one another, it remains unclear what students do, while they try to relate these topics. In a pilot study it was found that many students try to integrate the conceptual models, even when that is not possible. To investigate what exactly happens with students at secondary schools, two different instructional methods were compared within a pretest-posttest control group design in a group of 155 students. In one method the conceptual models were presented explicitly with the help of schemata whereas in the other the conceptual models were presented implicitly in cases presenting the required computations, as is the usual approach in Holland. The students of the group that received explicit instruction in conceptual models were supposed to perform better than the other group. And they did. But this was not due to the acquisition of new knowledge, but to the preservation of the knowledge acquired in an earlier stage.

INLEIDING

Docenten in het bedrijfseconomisch onderwijs op HAVO en VWO geven proefwerken over afgeronde eenheden nieuwe leerstof zonder dat zij de integratie van de nieuwe leerstof met de oude leerstof onderzoeken. Zij verwachten dat leerlingen de nieuwe stof als een zelfstandige eenheid opvatten en de kennis over de nieuwe leerstof toevoegen aan de kennis die zij reeds eerder hebben verworven. In een onderzoek naar de mogelijkheden om computers te gebruiken bij het bedrijfseconomisch onderwijs op het VWO (Vernooij, 1993b; Vernooij en Visch, 1994) kwam naar voren dat deze verwachting niet correct is. Veel leerlingen integreerden de nieuwe leerstof met reeds eerder verworven kennis. Nadere bestudering van de resultaten leidde tot de veronderstelling dat de leerlingen zich een onjuiste mentale voorstelling vormen van de bedrijfseconomie, wanneer de nieuwe leerstof niet consistent is met eerder behandelde leerstof.

De examenprogramma's voor bedrijfseconomie op HAVO en VWO zijn opgebouwd uit een samenstel van afzonderlijke subdisciplines. Deze zijn onderling niet consistent in het gebruik van de terminologie. Bij het onderdeel boekhouden schrijft het examenprogramma bij voorbeeld het gebruik van de term 'kostprijs verkopen' voor en verstaat daar 'de inkoopwaarde van de verkochte goederen' onder (zie appendix, schema 1). In deze 'kostprijs verkopen' zijn de inkoopkosten niet verwerkt, evenmin als de algemene bedrijfskosten die een ondernemer moet maken om zijn product te verkopen. Tegelijk is bij het onderdeel commerciële calculaties de term 'kostprijs per eenheid product' in gebruik. In tegenstelling tot de 'kostprijs verkopen' omvat de term kostprijs nu wel de directe inkoopkosten en een opslag voor algemene bedrijfskosten (zie appendix, schema 2). Meer verschillen zijn aanwezig (Vernooij, 1993a), maar deze blijven hier buiten beschouwing.

Op het Cals College in Nieuwegein is een pilot-onderzoek verricht in klas 5 VWO om hypothesen te ontwikkelen over de wijze waarop leerlingen met de inconsistenties in de leerstof omgaan. Bij dit onderzoek is gebruik gemaakt van het vigerende leerboek voor HAVO en VWO

(Hoogheid en Fuchs, 1987). Uit de resultaten bleek dat de onderzochte leerlingen op drie verschillende manieren omgingen met de inconsistenties tussen de boekhoudkundige benadering en de calculatorische benadering. Een klein aantal leerlingen hanteerde de twee benaderingen als naast elkaar staande begrippenapparaten. Een ander deel van de leerlingen herdefinieerde het *eerder* verworven boekhoudkundige begrippenapparaat op basis van het later verworven calculatorische begrippenapparaat. Een derde groep leerlingen herdefinieerden zowel het boekhoudkundige als het calculatorische begrippenapparaat (zie appendix, schema's 3 en 4). Het zoeken naar een samenhangend begrippenapparaat leidde bij hen niet alleen tot herdefiniëring van oude kennis, maar werkte ook belemmerend op het ontwikkelen van een correcte voorstelling van het nieuw aangeboden model.

Tijdens een hardop-denksessie merkte één van de leerlingen op: *"Als zij zeggen dat de kostprijs meer omvat dan de inkoopprijs en de inkoopkosten, dan doe ik dat wel, maar ik houd me wel vast aan: de opbrengst min de kostprijs is de brutowinst."* (Vernooij, 1993b).

Ten einde na te gaan of de gesignaleerde mentale processen ook plaatsvinden bij HAVO-leerlingen is een onderzoek opgezet naar de wijze waarop HAVO-leerlingen met deze problematiek omgaan. In de huidige leerboeken vindt op impliciete wijze onderricht in bedrijfseconomische modellen plaats. Auteurs presenteren vraagstukken waarin de modellen verwerkt zitten, zonder dat zij de modellen, los van de getallenvoorbeelden, beschrijven. Zij geven wel voorbeelden, maar maken niet duidelijk welke procedure ten grondslag ligt aan het voorbeeld.

In het kader van dit onderzoek is, op basis van de bevindingen uit de vakdidactiek natuurkunde, een alternatieve instructiewijze ontwikkeld voor het oplossen van bedrijfseconomische vraagstukken. Deze instructiewijze is gericht op het expliciteren van de verschillende bedrijfseconomische modellen. Daartoe is een wijze van schematiseren ontwikkeld die in de bedrijfseconomische appendix bij dit artikel is weergegeven.

In een vergelijkend onderzoek zijn de twee instructiewijzen met elkaar vergeleken. Een experimentele groep van circa 75 leerlingen kreeg expliciet onderwijs in bedrijfseconomische modellen en een controlegroep van circa 80 leerlingen kreeg impliciet onderwijs in diezelfde modellen. Onderzocht is in welke mate de leerlingen van beide groepen erin slaagden om zich de vereiste bedrijfseconomische modellen eigen te maken. Voor de operationalisering zijn verschillende testen ontwikkeld. In dit artikel komt slechts één van de gehanteerde testen aan de orde.

THEORETISCH KADER

Het onderzoek is opgezet vanuit de cognitieve psychologie. De Jong (1986) en Ferguson Hessler (1989) hebben in het kader van de vakdidactiek natuurkunde onderzocht welke band er bestaat tussen de vaardigheid om problemen op te lossen en de aanwezigheid van domeinspecifieke kennis. Zij formuleerden expliciet welke probleemschema's experts tot hun beschikking hebben.

Probleemschema's zijn een representatie (Jorna, 1989) van de voorstelling die experts zich maken van de problematiek binnen hun vakgebied (Norman, Gentner & Stevens, 1976; Chi, Feltovich & Glaser, 1981; Larkin, 1983; Anderson, 1985). Leerlingen dienen soortgelijke schema's te ontwikkelen door nieuwe kennis te relateren en te integreren met bestaande kennis (Ausubel c.s., 1978; Prawat, 1989). Norman (1983) maakte daarbij onderscheid tussen conceptuele modellen, mentale modellen en conceptualisering van de mentale modellen. Hij verstaat onder conceptuele modellen de kennisstructuren die docenten aan hun leerlingen willen overbrengen. De kennisstructuren die leerlingen daadwerkelijk ontwikkelen, noemt hij mentale modellen. Hij merkt daarbij op dat die mentale modellen slechts indirect zijn waar te nemen.

De conceptuele modellen die in het bedrijfseconomisch onderwijs centraal staan, zullen in dit onderzoek omschreven worden als 'bedrijfseconomische modellen'. Deze zijn opgebouwd uit handelingsvoorschriften die aangeven hoe de diverse bedrijfseconomische grootheden met

elkaar gecombineerd moeten worden, bij voorbeeld: *brutowinst = opbrengst verkopen minus kostprijs verkopen* (zie appendix, schema 1). Bij het uitwerken van vraagstukken zullen de leerlingen de juiste handelingsvoorschriften op het juiste moment moeten toepassen. Het achterhalen van de vereiste handelingsvoorschriften is de kern van het oplossingsproces.

Bij het probleemoplossen kunnen vier functionele stadia onderscheiden worden: oriëntatie, analyse, planning en uitwerking (Mettes en Pilot, 1980; De Jong, 1986; Ferguson Hessler, 1989; Schoenfeld, 1989; e.a.). De essentie zit in de analysefase. Om de samenhang tussen de data en de onbekende grootte in een specifiek vraagstuk op grond van economisch inzicht te kunnen vaststellen, moeten leerlingen dat vraagstuk in drie stappen analyseren: (1) vaststellen welk bedrijfseconomisch model van toepassing is op een bepaald vraagstuk, (2) vaststellen welk fragment uit dat model geïsoleerd moet worden op grond van de beschikbare gegevens en de onbekende grootte en (3) vaststellen hoe dat fragment getransformeerd moet worden tot een oplossingsschema voor de betreffende opgave.

In wezen gaat het bij bedrijfseconomische problemen om redactievraagstukken, die wiskundig zijn te definiëren als vergelijkingen van de nulde graad: er is één onbekende en een serie gegevens. De leerlingen moeten de beschikbare gegevens zo combineren dat zij de waarde van die ene onbekende grootte kunnen berekenen. De moeilijkheidsfactor zit niet in wiskundige bewerkingen, maar in het ontbreken van het oplossingsschema dat aangeeft op welke wijze de gegevens gecombineerd moeten worden tot een antwoord. Dat oplossingsschema moeten de leerlingen ontleen aan één van de bestaande bedrijfseconomische modellen en toevoegen aan de tekst van het vraagstuk om het oplosbaar te maken (Vernooij, 1996). In principe zouden de leerlingen eenduidige handelingsvoorschriften uit de bedrijfseconomische modellen moeten kunnen afleiden. Gezien de verschillen in het jargon van de afzonderlijke bedrijfseconomische subdisciplines is echter een grote mate van taalvaardigheid vereist om tot een juiste interpretatie van de gegeven vraagstukken te komen.

De gangbare bedrijfseconomische didactiek is niet gebaseerd op *explanation based learning* maar op *similarity based learning* (Ahn, Brewer en Mooney, 1992). In de impliciete instructiewijze moeten leerlingen een algemeen geldig bedrijfseconomisch model construeren door een gemeenschappelijk patroon te herkennen in een reeks gelijksoortige opgaven. Zij moeten bewust letten op overeenkomsten tussen analoge gevallen (Prawat, 1989). Vervolgens moeten zij zelfstandig mentale modellen opbouwen aan de hand van inferenties (Holland, Holyoak, Nisbett & Thagard, 1989). De hypothese van het uitgevoerde onderzoek houdt in dat 'explanation based learning' met behulp van expliciet gepresenteerde algemeen geldige modellen, tot mentale modellen leidt die een betere afspiegeling vormen van de gepresenteerde bedrijfseconomische modellen, dan 'similarity based learning' waarbij de modellen impliciet blijven.

Bij de ontwikkeling van mentale modellen kunnen misconcepties ontstaan (Chi, 1992). Alhoewel het ontstaan van misconcepties een normaal verschijnsel is in het onderwijs (Driver, Guesne en Tiberghien, 1985) is het onderkennen en classificeren ervan noodzakelijk om te voorkomen dat deze misconcepties blijven bestaan (Taconis, 1995). Taconis (1995) onderzocht recentelijk voor het vak natuurkunde in het voortgezet onderwijs hoe misconcepties kunnen ontstaan bij de verwerking van de leerstof.

Mogelijke verklaringen voor het ontstaan van misconcepties kunnen gegeven worden op basis van de theorie over 'conceptual change'. Chi, Slotta en De Leeuw (1994) definiëren conceptual change als gedachtenprocessen waarbij een concept van de ene categorie overgeplaatst moet worden naar een andere categorie, bij voorbeeld omdat een concept in een wetenschappelijke visie andere kenmerken heeft dan een soortgelijk concept in het dagelijkse spraakgebruik. Misconcepties zijn het gevolg van de plaatsing van nieuwe entiteiten in een categorie, waar deze op ontologische gronden niet thuis hoort. Bij voorbeeld, leerlingen plaatsen een proces dat niet aan een tijdsverloop gebonden is, in de categorie 'gebeurtenissen', die juist wel gekenmerkt worden door een begin- en een eindtijdstip.

Binnen de bedrijfseconomie is het onderscheid tussen voorraadgrootheden (met als dimensie 'guldens per stuk') en stroomgrootheden (met als dimensie 'guldens per periode') van belang. Leerlingen die dit ontologische verschil niet opmerken, raken in de problemen wanneer zij twee

begrippen laten samenvallen, terwijl een strikt onderscheid vereist is. Eerder in dit artikel is de opmerking van één leerling tijdens een hardop-denksessie aangehaald: *"Als zij zeggen dat de kostprijs meer omvat dan de inkoopprijs en de inkoopkosten, dan doe ik dat wel, maar ik houd me wel vast aan: de opbrengst min de kostprijs is de brutowinst."* In het eerste deel van de uitspraak gaat het om de kostprijs per product, uitgedrukt in 'guldens per stuk', en in het laatste deel gaat het om de kostprijs verkopen, uitgedrukt in 'guldens per periode'. De leerling probeert de term 'kostprijs' te hanteren zonder nadere verwijzing naar de vereiste dimensies (Vernooij, 1993c). De emotionele weerstand van de leerling die uit het citaat blijkt, vloeit onder andere voort uit het miskennen van het ontologische verschil tussen de begrippen 'kostprijs per product' en 'kostprijs verkopen'.

Toch is deze visie op conceptual change ontoereikend om de emotionele weerstand volledig te verklaren. Zij houdt in onvoldoende mate rekening met het geschetste onderscheid tussen de boekhoudkundige en de calculatorische benadering. Als de termen 'kostprijs per product' en 'kostprijs verkopen' uit hetzelfde begrippenapparaat komen, is te verwachten dat: $afzet \times kostprijs \text{ per product} = kostprijs \text{ verkopen}$. (Met als logische dimensioneel verband: $stuks \text{ per periode} \times guldens \text{ per stuk} = guldens \text{ per periode}$). Om te begrijpen dat dit niet het geval is, moeten de leerlingen niet alleen onderkennen dat de twee termen een verschillende dimensie hebben, maar ook dat zij afkomstig zijn uit twee afzonderlijke subdisciplines binnen de bedrijfseconomie. En dat aspect komt niet terug in de benadering van Chi *et al.*

Vosniadou (1994) biedt wel een verklaring waarin ruimte is voor de invloed van afzonderlijke subdisciplines. Zij onderkent twee vormen van 'conceptual change': verrijking en herziening van mentale modellen. Bij verrijking voegt een leerling nieuwe informatie toe aan een reeds bestaande conceptuele structuur. Bij herziening is de opslag van nieuwe kennis alleen mogelijk als het bestaande mentale model een verandering ondergaat door de vervanging van geldende vooronderstellingen door andere.

Het geschetste probleem van de geciteerde leerling is vanuit de visie van Vosniadou te verklaren als een poging om kennis van calculatorische berekeningen toe te voegen aan een conceptuele structuur die opgebouwd is uit boekhoudkundige termen. Pas als de leerling in staat is om te onderkennen dat er een andere conceptuele structuur bestaat die onafhankelijk van de boekhoudkundige structuur functioneert, is het mogelijk de nieuwe term in de daarbij vereiste conceptuele structuur te plaatsen. Zolang dat besef er niet is, zal de leerling trachten de boekhoudkundige en calculatorische termen tot een samenhangend geheel te brengen en aldus een mentaal model ontwikkelen dat geen goede afspiegeling is van de aangeboden leerstof.

Auteurs van leerboeken voor HAVO en VWO doen geen moeite om de verschillende hoofdstukken uit hun boek ten opzichte van elkaar te profileren als benaderingen die afkomstig zijn uit verschillende subdisciplines. Zij presenteren de leerstof als onderdelen die steeds tot een verrijking moeten leiden, in plaats van tot een herziening. Het uitgevoerde onderzoek had mede ten doel dit uitgangspunt ter discussie te stellen. Vanuit het oogpunt van conceptual change is het van belang te weten of onderwijs gebaseerd op 'explanation based learning' inferenties oproept die leiden tot meer adequate mentale modellen dan onderwijs gebaseerd op 'similarity based learning'. Leerlingen laten zich bij het oplossen van vraagstukken leiden door de mentale modellen die zij in hun hoofd hebben en de verwachting is dat het oplosproces succesrijker verloopt naar mate deze mentale modellen beter aansluiten bij de vereiste bedrijfseconomische modellen.

DOELSTELLING VAN HET ONDERZOEK

Met behulp van probleem-analyse-diagrammen (Vernooij, 1990) is voor een aantal hoofdstukken uit het genoemde leerboek van Hoogheid en Fuchs (1987) een conceptuele beschrijving gemaakt van de vraagstukken die de leerlingen moeten uitwerken. Vanuit deze diagrammen zijn de bedrijfseconomische modellen afgeleid die de leerlingen aangeboden krijgen. Explicitering van de modellen maakte duidelijk dat er tal van verschillen bestaan tussen de subdisciplines, waardoor de probleemschema's onderling niet verenigbaar zijn (Vernooij, 1993a).

Ten einde na te gaan hoe leerlingen uit 4 HAVO omgaan met dergelijke, onderling strijdige probleemschema's is een empirisch onderzoek opgezet, gesitueerd in een realistische klas-situatie. Twee instructiewijzen zijn vergeleken met behulp van een voortoets-natoets controle groep ontwerp aangevuld met een retentietoets. De experimentele groep kreeg expliciet onderwijs in bedrijfseconomische modellen met behulp van probleemschema's en de controlegroep kreeg volgens de gebruikelijke benadering impliciet onderwijs via voorbeeldopgaven. Meting van de verschillen tussen de twee groepen vond onder andere plaats aan de hand van vragen naar de omschrijving van handelingsvoorschriften die deel uitmaken van de gedoceerde bedrijfseconomische modellen. Toetsing van de verschillen in de gepresenteerde antwoorden vond plaats op basis van de navolgende onderzoekshypothese:

Leerlingen uit 4 HAVO die expliciet onderwijs hebben gehad in bedrijfseconomische modellen, formuleren na afloop van de instructie handelingsvoorschriften die dichter liggen bij de gedoceerde bedrijfseconomische modellen dan leerlingen die impliciet onderwijs hebben gehad.

Op grond van de resultaten uit het pilot-onderzoek is de verwachting dat de leerlingen uit de experimentele groep een hogere score behalen op de natoets dan de leerlingen uit de controlegroep, omdat zij beter in staat zijn de oude en de nieuwe leerstof te scheiden. De verwachting is voorts dat de leerlingen uit de controlegroep, meer dan de leerlingen uit de experimentele groep, op zoek gaan naar compromissen tussen de termen die in twee verschillende betekenissen worden gebruikt en dat zij daardoor een eigen variant op de bedrijfseconomische modellen ontwikkelen.

METHODE

Proefpersonen

Voor de selectie van proefpersonen zijn alle scholen in de regio Utrecht / Zuid Holland benaderd die op de HAVO gebruik maken van het 'Basisboek Bedrijfseconomie' van Hoogheid & Fuchs (1987). De keuze van deze methode maakte het mogelijk aansluiting te zoeken bij de resultaten van het pilot-onderzoek in 6 VWO. Het aantal scholen met de gezochte methode bleek echter beperkt te zijn omdat de meeste scholen de methode B-E in Balans van Hogenbirk (1988) op de HAVO gebruiken. Dit boek was echter minder geschikt voor het onderzoek, omdat de commerciële calculaties geheel ondergeschikt zijn gemaakt aan specifieke rekentechnieken, zodat veel onnodige inconsistenties in het taalgebruik optreden. In totaal bleken zes scholen de methode van Hoogheid & Fuchs op het HAVO te gebruiken en het hoofdstuk te behandelen conform de intenties van de auteurs. De docenten bedrijfseconomie van deze zes scholen waren allen bereid mee te werken aan het onderzoek.

Bij de scholen namen alle leerlingen die bedrijfseconomie als examenvak hadden gekozen, deel aan het onderzoek. Om de gelijke behandeling van leerlingen binnen een bepaalde school te waarborgen en vermenging van instructiewijzen door onderlinge informatie-uitwisseling tussen leerlingen te voorkomen, zijn de scholen in hun geheel toegewezen aan een bepaalde conditie. Twee scholen met een ongeveer gelijk aantal leerlingen waren afkomstig uit de binnenstad van resp. Den Haag en Rotterdam. Beide hadden veel allochtone leerlingen. Eén van deze scholen is toegewezen aan de experimentele conditie en de ander aan de controleconditie. De tweede school die toegewezen werd aan de experimentele conditie was de school die eerder had deelgenomen aan het onderzoek op het VWO. De docenten van deze school waren immers reeds op de hoogte van de experimentele instructiewijze.

Op de scholen uit de experimentele conditie hadden 78 leerlingen gekozen voor bedrijfseconomie als examenvak. Van deze leerlingen vielen er tijdens het onderzoek 3 af door absentie wegens ziekte. Bij de scholen uit de controleconditie hadden 86 leerlingen bedrijfseconomie als examenvak gekozen. Hiervan vielen er tijdens het onderzoek 6 af door absentie wegens ziekte of veelvuldig schoolverzuim.

De leerlingen vormden de eenheid van onderzoek. Aggregatie van resultaten heeft plaatsgevonden op het niveau van de groep, in casu de experimentele groep en de controlegroep, en niet op het niveau van de klas. De samenstelling van de klassen had in een eerder stadium plaatsgevonden op basis van eindexamenpakketten die de leerlingen hadden gekozen. Op de scholen waar twee of meer klassen deelnamen aan het onderzoek, zouden systematische verschillen tussen klassen niet op voorhand uitgesloten zijn. Wel vond aggregatie plaats op schoolniveau ten einde controleberekeningen uit te voeren.

Materialen

Het instructiemateriaal is gebaseerd op hoofdstuk 23 uit het genoemde leerboek. De leerlingen hebben in de voorafgaande hoofdstukken geleerd hoe ze via de boekhoudkundige benadering de nettowinst in een voorbije periode kunnen berekenen (zie schema 1 van de appendix). In de eerste paragraaf van hoofdstuk 23 staat uiteengezet hoe een handelaar de gewenste verkoopprijs van een product kan berekenen (zie schema 2 van de appendix). Daarop aansluitend volgt een afwijkende manier om de nettowinst in een periode te berekenen. In de tweede paragraaf van hoofdstuk 23 behandelen de auteurs de 'Direct Costing'-benadering, een berekeningswijze van de nettowinst die is overgenomen uit de Angelsaksische literatuur, met weer eigen karakteristieken.

Voor de controlegroep is op basis van het leerboek een aangescherpt cursusboek gemaakt door de theorie en de opgaven uit het gekozen hoofdstuk te zuiveren van slordige formuleringen. Zodoende bleven alleen de essentiële verschillen tussen de verschillende bedrijfseconomische benaderingen behouden. Vanuit dit cursusboek voor de controlegroep is een cursusboek ontwikkeld voor de experimentele groep. Alle rekenkundige voorbeelden zijn daarin vervangen door schema's waaruit de leerlingen zelfstandig de berekeningen konden afleiden. Voor de oefenopgaven zijn richtlijnen geformuleerd waarmee de leerlingen van de experimentele groep op systematische wijze de vraagstukken konden oplossen.

Toetsing van de kennis vond plaats met behulp van twee testen: de test Kennis van Handelingsvoorschriften en de test Kennis van Procedures. Tevens hebben 36 leerlingen deelgenomen aan een kaartsorteringsexperiment (Vernooij, 1993b). In dit artikel beperkt de beschrijving zich tot de test Kennis van Handelingsvoorschriften (KVH). Deze test is speciaal ontwikkeld voor dit onderzoek. Zij bestaat uit acht vragen die gericht zijn op de wijze waarop een handelaar bedrijfseconomische grootheden berekent. De eerste vier vragen gaan over de berekening van grootheden die van belang zijn voor de vaststelling van de nettowinst in een periode volgens de boekhoudkundige benadering (zie schema 1 in de appendix). De tweede reeks van vier vragen gaat over de berekeningsprocedure van de verkoopprijs volgens de calculatorische benadering (zie schema 2 in de appendix).

Procedure

Om de leersituatie zo realistisch mogelijk te maken, is het onderzoek geplaatst in een alledaagse klassesituatie. Dit houdt in dat de eigen docenten alle lessen gaven en dat de natoets meetelde als regulier proefwerk. Dit bracht noodzakelijkerwijze met zich mee dat niet voldaan kon worden aan de eisen die behoren bij een a-selecte toewijzing van onderzoekselementen aan de onderzoeksconditie. De invloeden van de docenten op hun klas en de sfeer die daar eventueel uit voortvloeide, zijn door deze keuze niet als mogelijke verklarende variabelen uit te sluiten. Een inperking van deze invloeden vond wel plaats doordat in totaal vier docenten meededen binnen de experimentele conditie en vijf docenten binnen de controleconditie, waardoor een zekere spreiding van subjectieve invloeden plaats vond. Uitschakeling van de invloed van de docent, door deze invloed op een of andere wijze mee te nemen in de covariantie-analyse, was niet mogelijk, omdat op één der scholen de leerlingen op grond van hun vakkenpakketkeuze ingedeeld waren in klassen die van verschillende docenten les kregen.

Het uitgevoerde onderzoek bevat een voortoets-natoets-ontwerp met een retentietoets. De voortoets vond plaats voor aanvang van het onderzoek, maar na afronding van de boekhoudkundige benadering. De natoets vond plaats direct na afronding van paragraaf 23.1. De retentie-

De onderzoeksopzet luidt:

1. Experimentele groep:	T_1	X_0	X_1	T_2	X_3	T_3
2. Controlegroep:	T_1		X_2	T_2	X_4	T_3

T_1 voortoets Kennis van Handelingsvoorschriften (KVH1);
 X_0 introductie in de schematechniek gedurende drie lesuren;
 X_1 zes lesuren expliciet onderwijs in de bedrijfseconomische modellen uit paragraaf 23.1;
 X_2 zes lesuren impliciet onderwijs in de bedrijfseconomische modellen uit paragraaf 23.1;
 T_2 natoets Kennis van Handelingsvoorschriften (KVH2);
 X_3 vier lesuren expliciet onderwijs in de bedrijfseconomische modellen uit paragraaf 23.2;
 X_4 vier lesuren impliciet onderwijs in de bedrijfseconomische modellen uit paragraaf 23.2;
 T_3 retentietoets Kennis van Handelingsvoorschriften (KVH3).

Fig. 1. Onderzoeksopzet aangevuld met het onderwijs- en toetsprogramma.

toets vond plaats na afronding van paragraaf 23.2 (zie figuur 1). Deze opzet leverde een beeld op van de ontwikkeling van de mentale voorstelling van de gedoeerde bedrijfseconomische modellen.

Instrumentatie

De onderzoekshypothese luidde dat de experimentele groep handelingsvoorschriften zou formuleren die dichter liggen bij de bedrijfseconomische modellen dan de controlegroep. Dit moest zich o.a. uiten in de natoets en de retentietoets van de test Kennis Van Handelingsvoorschriften doordat de experimentele groep tot een hogere score zou komen dan de controlegroep.

Om mogelijke alternatieve verklaringen voor de gevonden resultaten te minimaliseren zijn een serie testvariabelen mede in beschouwing genomen. Dit is temeer van belang omdat de leerlingen niet a-select zijn toegewezen aan één van de onderzoekscondities maar in totaliteit als school. Deze wijze van toedeling zou systematische verschillen tussen de groepen tot gevolg kunnen hebben die als verklarende variabelen een rol zouden kunnen spelen.

Voor alle leerlingen zijn beschrijvende gegevens verzameld over de samenstelling van hun vakkenpakket (i.c. de aanwezigheid van wiskunde A, wiskunde B en algemene economie in het pakket), over het geslacht van de leerlingen en over de vraag of zij doubleur zijn of afkomstig zijn uit 4 MAVO. Ook is hun analytisch vermogen gemeten met de DAT-analogieën test (Evers en Lucassen, 1984). Gezien de positionering van het onderzoek in een realistische klassensituatie was het niet mogelijk om meer onderdelen van de DAT af te nemen, omdat de tijd die de docenten beschikbaar stelden, dat niet mogelijk maakte. Tot slot is een nauwkeurige inventarisatie gemaakt van de tijd besteed aan huiswerk. Dit laatste is zodanig gebeurd dat de docenten geen inzage kregen in de gegevens van het huiswerk (Vernooij, 1993b).

Scoring en beoordeling van de test Kennis van Handelingsvoorschriften

Beoordeling van de test vond plaats door twee onafhankelijke beoordelaars. Zij kregen het werk van alle 155 leerlingen voorgelegd in alfabetische volgorde van de achternamen van de leerlingen zonder verwijzing naar de school van herkomst. De beoordelaars ontvingen een lijst van mogelijke antwoorden die correspondeerden met de verschillende benaderingen: de boekhoudkundige benadering (model B), de calculatorische benadering (model C), de alternatieve benadering die veelvuldig in het VWO was waargenomen (model A) of een niet-identificeerbare benadering (model X). Deze coderingslijst was ontwikkeld aan de hand van de resultaten uit het pilot-onderzoek in 5 VWO en bevatte de waargenomen nuancerings of slordigheden die kunnen optreden bij de formulering van de antwoorden.

De eerste vraag aan de beoordelaars was om vast te stellen vanuit welke model de leerlingen hun antwoorden geformuleerd hadden. De interbeoordelaarsbetrouwbaarheid, vastgesteld aan de hand van Cohen's kappa, was 0,62 op de voortoets, 0,65 op de natoets en 0,59 op de

retentietoets (Vernooij, 1993b). De tweede vraag aan de beoordelaars was om volledig correcte antwoorden te honoreren met twee punten en antwoorden die wel volgens het correcte model geformuleerd waren, maar onzuiverheden bevatten, te honoreren met één punt. Ook hier is de interbeoordelaarsbetrouwbaarheid gemeten. Deze bedroeg 0,68 op de voortoets, 0,70 op de natoets en 0,65 op de retentietoets. Ter afronding hebben de twee beoordelaars in onderling overleg een gezamenlijk eindoordeel vastgesteld, waarbij de richtlijn was om steeds de meest strenge interpretatie van de normen toe te passen.

RESULTATEN

De keuze voor afzonderlijke modellen

De eerste vraag aan de beoordelaars was om vast te stellen vanuit welke benadering de leerlingen hun antwoorden geformuleerd hebben. Dit artikel beperkt zich tot de gemaakte keuzen bij de berekening van de ‘kostprijs verkopen’ (item a) en de berekening van de ‘kostprijs per product’ (item e). Informatie over de andere items is elders beschikbaar (Vernooij, 1993a).

De instructie in de calculatorische benadering leidde bij de natoets en de retentietoets KVH tot een significant verschil ($\chi^2 = 16,8$ bij een $p \leq 0.001$) in correcte antwoorden tussen de experimentele groep en de controlegroep op een niveau van $\alpha = 5\%$ (tabel 1). Bij de retentietoets is eveneens een significant verschil ten gunste van de experimentele groep ($\chi^2 = 26,4$; $p \leq 0.001$) te constateren. Bij de experimentele groep stijgt in de retentietoets het aantal correcte antwoorden, terwijl dat bij de controlegroep niet het geval is.

Opvallend is de mate waarin leerlingen bij de natoets en de retentietoets de boekhoudkundige grootheid ‘kostprijs verkopen’ herdefinieerden in termen van de calculatorische benadering. Bij de natoets waren dat in totaal 96 van de 155 leerlingen. Van de experimentele groep handelden 36 van de 75 leerlingen zo en van de controlegroep 60 van de 80 leerlingen. Tussen de twee groepen bestaat een significant verschil ($\chi^2 = 16,8$; $p \leq 0.001$).

In tabel 2 is af te lezen dat bij de natoets en de retentietoets relatief meer leerlingen uit de controlegroep de berekening van de ‘kostprijs per product’ correct formuleerden dan leerlingen uit de experimentele groep. Bij de voortoets is een significant verschil op een niveau van $\alpha = 5\%$

Tabel 1. Geregistreerde codes op de vraag naar de berekening van de kostprijs verkopen.

	KVH1		KVH2		KVH3	
	EG	CG	EG	CG	EG	CG
model A	–	2	11	9	9	5
model B	32	26	18	1 (*)	27	2 (*)
model C	5	11	36	60	23	44
overige	38	41	10	9	16	29

Analyse van het verschil in correcte antwoorden:

KVH1: $\chi^2 = 1,3$ bij een $p \leq 0.254$ ten gunste van de experimentele groep;

KVH2: $\chi^2 = 16,8$ bij een $p \leq 0.001$ (*) ten gunste van de experimentele groep;

KVH3: $\chi^2 = 26,4$ bij een $p \leq 0.001$ (*) ten gunste van de experimentele groep;

(*) duidt op een verschil op een significantieniveau $\alpha = 5\%$.

Toelichting:

model A: alternatieve benadering;

model B: boekhoudkundige benadering: correct antwoord;

model C: calculatorische benadering.

EG: experimentele groep ($n = 75$), CG: controlegroep ($n = 80$).

Tabel 2. Geregistreerde codes op de vraag naar de berekening van de kostprijs per product.

	KVH1		KVH2		KVH3	
	EG	CG	EG	CG	EG	CG
model A	2	2	10	3	19	9
model B	19	3	–	–	1	1
model C	3	12 (*)	52	67	42	52
overige	51	60	13	10	13	21

Analyse van het verschil in correcte antwoorden:

KVH1: $\chi^2 = 4,2$ (*) bij een $p \leq 0.041$ ten gunste van de controlegroep;

KVH2: $\chi^2 = 3,7$ bij een $p \leq 0.054$ ten gunste van de controlegroep;

KVH3: $\chi^2 = 1,0$ bij een $p \leq 0.326$ ten gunste van de controlegroep.

(*) duidt op een verschil op een significantieniveau $\alpha = 5\%$.

Toelichting:

model A: alternatieve benadering;

model B: boekhoudkundige benadering;

model C: calculatorische benadering: correct antwoord.

EG: experimentele groep ($n = 75$), CG: controlegroep ($n = 80$).

te constateren. Bij de natoets is het verschil net niet significant op een niveau van $\alpha = 5\%$. Bij de retentietoets is het verschil in het geheel niet significant.

Een verklaring voor het gevonden verschil tussen de twee groepen in de voortoets werd gevonden bij de controleberekeningen van de resultaten per school. Uit deze analyse bleek dat de verschillen op de voortoets geheel te herleiden waren tot de resultaten van één school, nader aan te duiden als school Z. Een aantal van de 25 leerlingen van deze school maakten in de voortoets fouten, die leerlingen van de andere scholen pas in de tweede en derde toetsronde maakten. Telefonische navraag bij enkele leerlingen van school Z leverde de mededeling op dat de docent de stof uit het onderzoek reeds op voorhand besproken had. Bij de covariantie-analyses (weergegeven in tabel 3) zijn de leerlingen van deze school daarom buiten de berekening gehouden.

Uit de gepresenteerde resultaten blijkt dat meer leerlingen uit de experimentele groep overstappen naar het alternatieve model dan leerlingen uit de controlegroep. Nader onderzoek bracht aan het licht dat bij de natoets acht leerlingen zowel bij de berekening van de 'kostprijs verkopen' (item a) als de 'kostprijs per product' (item c) model A gebruikten. Bij de retentietoets waren dat tien leerlingen.

De behaalde scores

De verdeling van de behaalde scores

In tabel 3 staan de gemiddelde scores die de experimentele groep en de controlegroep behaalden op de test Kennis van Handelingsvoorschriften. Om na te gaan of een covariantie-analyse mogelijk was, is onderzocht of de behaalde scores redelijk normaal verdeeld zijn.

Bij de voortoets bleek dit niet het geval te zijn, althans wat betreft de totaalscore op de items a t/m h. Uitsplitsing van de score in twee delen leverde hiervoor een verklaring op. De score op de items e t/m h (berekening verkoopprijs per product volgens de calculatorische benadering) was niet normaal verdeeld, omdat veel leerlingen nul punten scoorden op deze items. Dit is conform de verwachting omdat de docenten de berekening van de verkoopprijs per product volgens de calculatorische benadering nog niet hadden behandeld. Een redelijk normale verde-

ling trad wel op bij de score op de items a t/m d (berekening nettowinst in een periode volgens de boekhoudkundige benadering). Deze items hadden immers betrekking op lesstof uit voorafgaande hoofdstukken.

Bij de natoets vertoonde de totaalscore over de items a t/m h een redelijk normale verdeling. Dit was echter niet het geval bij de score op de items a t/m d. Veel leerlingen uit de controlegroep scoorden nul punten op deze reeks van vier vragen, zodat een tweetoppige verdeling ontstond. Het verschil tussen het aantal leerlingen uit de experimentele groep en het aantal leerlingen uit de controlegroep dat nul punten scoorden, is significant ($\chi^2 = 15,4$; $p \leq 0.001$). Op de items e t/m h is wel een verdeling te constateren die benaderd kan worden met de normale verdeling. Voor de retentietoets traden soortgelijke verdelingen op als bij de natoets.

De betrouwbaarheid

De betrouwbaarheid van de test is onderzocht met behulp van Cronbach's alpha. Bij de voortoets zijn de items e t/m h buiten beschouwing gelaten omdat deze vragen betrekking hadden op leerstof die niet behandeld was en het doel van deze vragen was om eventuele voorkennis te achterhalen.

De resultaten van de betrouwbaarheidstest waren niet volledig bevredigend. Cronbach's alpha bedroeg 0.69 voor de voortoets, 0.53 voor de natoets en 0.70 voor de retentietoets. Uitsplitsing van de alpha voor de natoets liet zien dat de alpha voor de items a t/m d 0.63 bedroeg en de alpha voor de items e t/m h 0.60. Dit duidt erop dat van de leerlingen die bij de natoets hoog scoorden op de berekening van de nettowinst, een deel hoog scoorde op de berekening van de verkoopprijs maar een ander deel juist laag. Dit zou een aanduiding kunnen zijn dat de test twee verschillende vaardigheden meet. Gezien het feit dat Cronbach's alpha bij de retentietoets 0.70 is, zijn toch covarianties berekend voor de natoets en de retentietoets.

Covariantie-analyse

Na uitsluiting van school Z levert de voortoets als resultaat op dat het gemiddelde van de controlegroep niet significant afwijkt van het gemiddelde van de experimentele groep (zie tabel 3).

Uit tabel 3 blijkt voorts dat de score van de controlegroep op de vragen die betrekking hebben op de boekhoudkundige benadering sterk gedaald is bij de natoets. Bij de experimentele

Tabel 3. Resultaten op de test Kennis van Handelingsvoorschriften.

	score a t/m d		score e t/m h		score a t/m h	
	gem.	s.d.	gem.	s.d.	gem.	s.d.
<i>Voortoets:</i>						
experimentele groep	3.3	2.1	0.2	**	3.5	2.1
controlegroep excl. Z	3.4	2.4	0.1	**	3.5	2.4
<i>Natoets:</i>						
experimentele groep	3.0	2.4	3.8	2.2	6.8	3.2
controlegroep excl. Z	1.3	1.4	3.6	2.0	4.9	2.4
<i>Retentietoets:</i>						
experimentele groep	3.7	2.7	3.0	2.5	6.6	4.2
controlegroep excl. Z	1.1	1.2	2.7	2.0	3.8	2.3

ANCOVA natoets ($n = 130$): $F(1,98) = 10.6$ bij een $p \leq .001$.*

ANCOVA retentietoets ($n = 130$): $F(1,98) = 5.9$ bij een $p \leq .009$.*

* duidt op een verschil op een significantieniveau $\alpha = 5\%$ bij eenzijdige toetsing.

** de standaarddeviatie is niet meer relevant bij deze lage gemiddelde score.

groep is de score nagenoeg gelijk gebleven. Als gevolg daarvan treedt een significant verschil tussen de twee groepen op. Ter controle is een berekening uitgevoerd op schoolniveau om na te gaan of lokatieverschillen een storende factor vormen. Voorts is afzonderlijk voor de scholen uit de controleconditie en uit de experimentele conditie met behulp van het 95% betrouwbaarheidsinterval onderzocht of de gemiddelden en de standaarddeviaties van de afzonderlijke scholen afkomstig kunnen zijn van dezelfde populatie. Dit bleek het geval te zijn, zij het dat één school bij de retentietoets een duidelijk hogere gemiddelde score opleverde dan de andere scholen uit de controleconditie. Dit zou kunnen duiden op verschillen binnen de controlegroep die voortvloeien uit lokatieverschillen. Met behulp van de rangordenummering is vervolgens vastgesteld dat de gevonden conclusies toch plausibel zijn.

Op de totaalscore voor de items a t/m h is een covariantie-analyse uitgevoerd. Voor de analyse van de resultaten op de natoets zijn onder andere als covariaten ingevoerd: de score van de voortoets, de gerapporteerde tijd besteed aan het huiswerk tijdens de eerste onderzoeksperiode en de score op de DAT-Analogieën test. Naast de covariaten zijn als onafhankelijke variabelen ingevoerd: geslacht, doubleur, wiskunde A, wiskunde B of economie in het lespakket. Uit deze covariantie-analyse van de natoets komt een significant verschil ten gunste van de experimentele conditie naar voren ($F(1,98) = 10.6$ bij een $p \leq .001$). Voor de voortoets wordt een significant effect aangegeven ($F(1,98) = 9.2$ bij een $p \leq .005$), hetgeen in overeenstemming is met de verwachting. Voor de DAT-Analogieëntest is geen significante bijdrage op een niveau van $\alpha = 5\%$ gevonden, evenmin als voor de gerapporteerde tijd besteed aan het huiswerk en voor de onafhankelijke variabelen (Vernooij, 1993b).

De retentietoets leverde eveneens een significant verschil ($F(1,98) = 5.9$ bij een $p \leq .009$) op ten gunste van de experimentele conditie (tabel 3). Als covariaten zijn nu de score op de natoets, de score op de DAT-Analogieën test en de gerapporteerde huiswerktijd in de tweede periode opgenomen, evenals de eerder genoemde onafhankelijke variabelen. De score op de retentietoets gaf een significant effect ($F(1,98) = 75.5$ bij een $p \leq .001$) voor de score op de natoets. Het huiswerk leverde eveneens een significant effect op ($F(1,96) = 4.30$ bij een $p \leq 0.021$). Voor de onafhankelijke variabelen zijn bij de retentietoets geen significante effecten gevonden.

DISCUSSIE

Mogelijke alternatieve verklaringen

De gevonden resultaten bevestigen de geformuleerde hypothese. Leerlingen uit 4 HAVO die expliciet onderwijs hebben gehad in bedrijfseconomische modellen, formuleren na afloop van de instructie handelingsvoorschriften die dichter liggen bij de gedoeerde bedrijfseconomische modellen dan leerlingen die impliciet onderwijs hebben gehad. Zij behaalden een hogere score op de natoets dan de leerlingen uit de controlegroep.

Het gevonden verschil volgt niet uit betere resultaten op de nieuwe leerstof. Het verschil dat in de natoets tussen de experimentele groep en de controlegroep bestaat, is geheel te verklaren uit de terugval in de score van de controlegroep op de oude kennis. Dit is in overeenstemming met de geformuleerde verwachting. Significant meer leerlingen uit de controlegroep herdefiniëerden de termen uit het boekhoudkundige model in termen van het calculatorische model. Tegelijk is te constateren dat leerlingen uit beide groepen tot een alternatief model kwamen, dat noch aansloot bij de boekhoudkundige benadering, noch bij de calculatorische.

Gezien de opzet van dit onderzoek bestaat de mogelijkheid dat er alternatieve verklaringen zijn voor de gevonden resultaten. De eerste alternatieve verklaring ligt in de keuze om de leerlingen niet a-select toe te wijzen aan een conditie, maar via de school waar zij op zitten. Dit houdt in dat schooleffecten kunnen optreden die de gevonden resultaten geheel of gedeeltelijk kunnen verklaren. Deze mogelijkheid was op voorhand bekend, maar onvermijdelijk omdat bewust is gekozen voor onderzoek in een realistische klassesituatie. Wel is voor de diverse toetsen aan de hand van de gemiddelde en standaarddeviatie per school onderzocht of de daarmee correspon-

derende betrouwbaarheidsintervallen overlap vertonen. Met uitzondering van school Z, bleek dit het geval te zijn.

Bij de analyse van de resultaten van school Z is geconstateerd dat deze docent voorinformatie heeft gegeven aan zijn leerlingen. Onbedoeld leverde deze docent daarmee een extra bijdrage aan de resultaten van dit onderzoek. De afwijkende antwoorden van zijn groep leerlingen liepen vooruit op de ontwikkelingen die zich voordeden bij de andere leerlingen uit de controlegroep. Dat leidt tot een ondersteuning van de resultaten die gevonden zijn bij de controlegroep. De terugval in eerder opgedane kennis is reeds op voorhand te constateren bij school Z. Uit de resultaten op de voortoets blijkt dat de herdefiniëring van oude kennis in dit geval sneller ging dan de opbouw van nieuwe kennis.

Naast schooleffecten zijn ook docenteffecten als alternatieve verklaring voor de gevonden resultaten op te voeren. Deze effecten zijn zoveel mogelijk ingeperkt. Slechts één van de negen docenten heeft de keuze gehad tussen deelname aan de experimentele conditie of de controleconditie, zodat de toewijzing van de docenten niet afhankelijk was van persoonlijke interesse, maar van de school waar zij aan verbonden waren. Voorts is er enige spreiding geweest, omdat er vier docenten deelnamen binnen de experimentele conditie en vijf binnen de controleconditie. Alleen herhaling van onderzoek onder andere condities kan nader uitsluitsel geven.

Een volgende alternatieve verklaring is het mogelijke verschil in abstractievermogen tussen de leerlingen. In de tijd die beschikbaar was, is weliswaar de DAT-analogieëntest afgenomen, maar deze levert nog geen compleet beeld van het abstractievermogen van de leerlingen. Er was echter niet meer tijd beschikbaar om meer onderdelen van de DAT-test af te nemen.

De tijd besteed aan het huiswerk bleek in de tweede periode een significante invloed gehad te hebben op de resultaten van de retentietoets. Dit hing mede samen met een tweetal storende interventies die op één van de scholen uit de experimentele conditie plaatsvonden vlak voor het afnemen van de retentietoets. De eerste interventie bestond uit een feest dat op de avond voor de geplande retentietoets georganiseerd werd. Dit feest ging gepaard met huiswerkvrij voor de volgende dag. Bij de opzet van het onderzoekprogramma was deze feestavond niet voorzien. Zij noodzaakte tot uitstel van de retentietoets. De tweede storende interventie bestond uit de wijze waarop de drie docenten meenden de ontstane leemte in de tijd op te moeten vullen. Tijdens de vrijgevallen les hebben zij de problematiek van de nettowinstberekening besproken vanuit de micro-economische benadering. Daarmee introduceerden zij, zonder overleg met de onderzoeker, een nieuw model om de nettowinst te berekenen. Deze extra les, plus een les met de intentie om de effecten weer terug te draaien, is als extra huiswerktijd aangemerkt.

Een andere alternatieve verklaring betreft de inzet van de leerlingen. Voorstelbaar is dat leerlingen die deelnemen aan het experimentele programma zich meer inzetten dan leerlingen die volgens de gebruikelijke wijze les krijgen. Het is echter de vraag in hoeverre leerlingen in verband met het onderzoek bereid zijn voor één hoofdstuk over te stappen op een andere aanpak. Tijdens één van de tegelijkertijd uitgevoerde hardop-denksessies bracht een leerling uit de experimentele conditie haar reserves onder woorden:

Ik heb de opgaven zo min mogelijk met modellen gedaan, dus als het allemaal klopte, wist ik gewoon dat ik het snapte. (...) Als het echt niet anders kon, dan keek ik wel van 'oh ja, zo en zo' en dan ging ik weer verder werken zonder modellen. (...) Omdat het maar één hoofdstuk was, dacht ik, ik moet per sé niet met die modellen gaan werken (...).

Suggesties voor nader onderzoek

Het rendement van een meer conceptuele benadering zoals toegepast in de alternatieve instructiewijze, lijkt niet te liggen in het *verwerven* van nieuwe kennis, maar in het *integreren* van nieuwe kennis in de reeds aanwezige kennisstructuur. Nader onderzoek is nodig om in soortgelijke en andere situaties na te gaan of de gevonden resultaten generaliseerbaar zijn. Uitwerking volgens de lijnen die zijn uitgezet door Vosniadou (1994) lijkt daarbij een passende aanpak.

In dit onderzoek is de alternatieve instructiewijze behandeld als een eenheid, terwijl deze methode is opgebouwd uit diverse elementen. Het expliciteren van de bedrijfseconomische

modellen is er daar één van. Een ander element is de opbouw van de instructie volgens de experimentele conditie aan de hand van een cognitieve strategie (Van Hout-Wolters, 1992) waarin verschillende fasen in het probleemoplossingsproces zijn onderkend. Bij de opzet van het huidige onderzoek is geen rekening gehouden met de mogelijke invloed van de verschillende fasen. Nader onderzoek naar de afzonderlijke bijdragen van deze verschillende componenten is daarom wenselijk.

Uit de gevonden omschrijvingen (zie tabel 1 en 2) blijkt dat HAVO-leerlingen net als VWO-leerlingen tot herdefiniëring van kennis overgaan. Sommigen doen dat met terugwerkende kracht en herschrijven de oude kennis, terwijl anderen menen er goed aan te doen om zowel de oude als de nieuwe stof te herdefiniëren. Het onderzoek heeft geen indicaties opgeleverd, op grond van welke overwegingen de leerlingen hun keuze maakten. Nader onderzoek naar dit keuzeproces is wenselijk.

Meer kennis over de wijze waarop leerlingen met eerder verworven kennis omgaan, is niet alleen nodig vanuit leerpsychologisch oogpunt. Ook vanuit onderwijskundig oogpunt is het van belang meer zicht te krijgen op de ontwikkeling van mentale modellen tijdens de lessen bedrijfs-economie. In de aanvang van dit artikel is opgemerkt dat de docenten veronderstellen dat een proefwerk gericht kan zijn op het laatste hoofdstuk zonder dat zij de integratie van dat hoofdstuk met eerder verworven kennis testen. Als het opbouwen van nieuwe kennis echter gepaard gaat met het afbreken van oude kennis, is het noodzakelijk om te zoeken naar nieuwe toetsvormen die zowel de opbouw van nieuwe kennis meten als de integratie van nieuwe kennis met oude.

LITERATUUR

- Ahn, W.K., Brewer, W.F. & Mooney, R.J. (1992). Schema acquisition from a single example. *Journal of Experimental Psychology: Learning Memory and Cognition*, 2, 391-412.
- Anderson, J.R. (1985). *Cognitive psychology and its implications*. New York: W.H. Freeman and Company.
- Ausubel, D.P., Novak, J.D. & Hanesian, H. (1978). *Educational psychology, a Cognitive View*. New York: Holt, Rinehart and Winston.
- Chi, M.T.H., Feltovich, P.J. & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.
- Chi, M.T.H. (1992). Conceptual change within and across ontological categories: examples from learning and discovery in science. In R. Giere (ed.). *Cognitive models in science: Minnesota studies in philosophy of science*. University of Minnesota Press.
- Chi, M.T.H., Slotta, J.D. & Leeuw, N. de (1994). From things to processes: a theory of conceptual change for learning science concepts. *Learning and Instruction*, 4, 27-43.
- Driver, R., Guesne, E., & Tiberghien, A. (1985). *Children's ideas in science*. Milton Keynes, UK: Open University Press.
- Evers, A. & Lucassen, W. (1984). *DAT '83*. Lisse: Swets & Zeitlinger.
- Ferguson Hessler, M.G.M. (1989). *Over kennis en kunde in de fysica*. Eindhoven: proefschrift.
- Hogenbirk, J.C., Fijn, R.J., Hamers, H.H. & Reuver, W.J.M. de (1988/1991). *Moderne bedrijfseconomie voor het HAVO/VWO, B-E in balans*. Voorschoten: Uitgeverij Educatief.
- Holland, J.H., Holyoak, K.J., Nisbett, R.E. & Thagard P.R. (1986/89). *Induction. Processes of inference, learning and discovery*. Cambridge: MIT Press.
- Hoogheid, J.C. & Fuchs H. (1987). *Basisboek Bedrijfseconomie voor HAVO en VWO*. Groningen: Wolters-Noordhoff.
- Hout-Wolters, B.H.A.M. van (1992). *Cognitieve strategieën als onderwijsdoel*. Amsterdam: oratie.
- Jong, T. de (1986). *Kennis en het oplossen van vraagstukken*. Eindhoven: proefschrift.
- Jorna, R.J. (1989). *Kennisrepresentaties en symbolen in de geest*. Groningen: proefschrift.
- Larkin, J.H. (1983). The role of problem representation in Physics. In D. Gentner & A.L. Stevens (eds.). *Mental models*. Hillsdale, N.J.: Lawrence Erlbaum.
- Mettes, C.T.C.W. & Pilot, A. (1980). *Over het leren oplossen van natuurwetenschappelijke problemen. Een methode voor ontwikkeling en evaluatie van onderwijs toegepast op een cursus Thermodynamika*. Enschede: proefschrift.
- Norman, D.A., Gentner, D.R. & Stevens, A.L. (1976). Comments on learning schemata and memory representation. In D. Klahr (ed.). *Cognition and instruction* (pp. 177-197). Hillsdale N.J.: Lawrence Erlbaum.

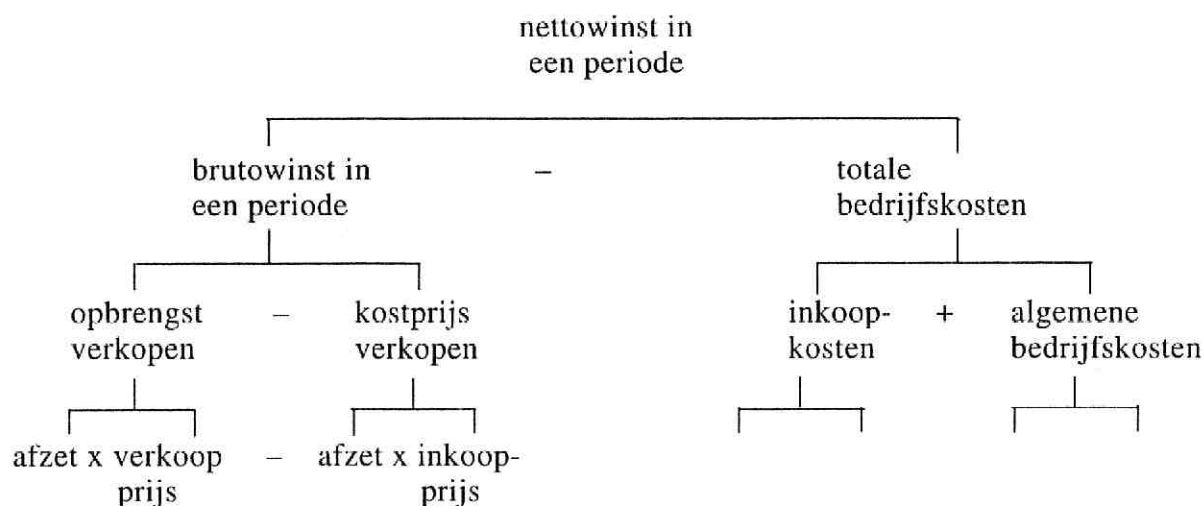
- Norman, D.A. (1983). Some observations on mental models. In D. Gentner & A.L. Stevens (eds.). *Mental models*. Hillsdale, N.J.: Lawrence Erlbaum.
- Prawat, R.S. (1989). Promoting access to knowledge, strategy, and disposition in students: a research analysis. *Review of Educational Research*, 59, 1-41.
- Schoenfeld, A.H. (1989). Teaching mathematical thinking and problem solving. In L.B. Resnick & L.E. Knopfer (Eds.). *Toward the thinking curriculum: Current cognitive research*. 1989 Yearbook of the Association for Supervision and Curriculum Development.
- Taconis, R. (1995). *Understanding based problem solving. Towards qualification-oriented teaching and learning in physics education*. Eindhoven: proefschrift.
- Vernooij, A.T.J. (1990). A systematic problem approach in Business Administration. In J.M. Pieters, K. Breuer, & P.R.J. Simons (eds.). *Learning environments, contributions from Dutch and German research*. Berlijn: Springer Verlag.
- Vernooij, A.T.J. (1993a). *Het leren oplossen van bedrijfseconomische problemen*. (proefschrift). Zutphen: Thieme.
- Vernooij, A.T.J. (1993b). Deelrapporten bij het onderzoek naar het leren oplossen van bedrijfseconomische problemen.
 - deelrapport 2: Het gebruik van spreadsheets in de lessen bedrijfseconomie. Verslag van een pilot-onderzoek in 5 VWO;
 - deelrapport 3: Instrumentatie van het onderzoek in 4 HAVO;
 - deelrapport 4: De test Kennis van Handelingsvoorschriften.
- Vernooij, A.T.J. (1993c). De toetsende tucht van de dimensie-analyse. *Maandblad voor Accountancy en Bedrijfseconomie*, 67 (10).
- Vernooij, A.T.J. & Visch, E. (1994). The use of spreadsheets in business education, *Economia, The Journal of the Association of European Economics Education*. 4, (1).
- Vernooij, A.T.J. (1996). *Elementaire bedrijfseconomische modellen*. Zutphen: Thieme.
- Vosniadou S. (1994). Capturing and modeling the process of conceptual change, *Learning and Instruction*, 4, (1), 45-69.

Manuscript ontvangen 1-9-1994

Definitieve versie ontvangen 11-4-1996

BEDRIJFSECONOMISCHE APPENDIX

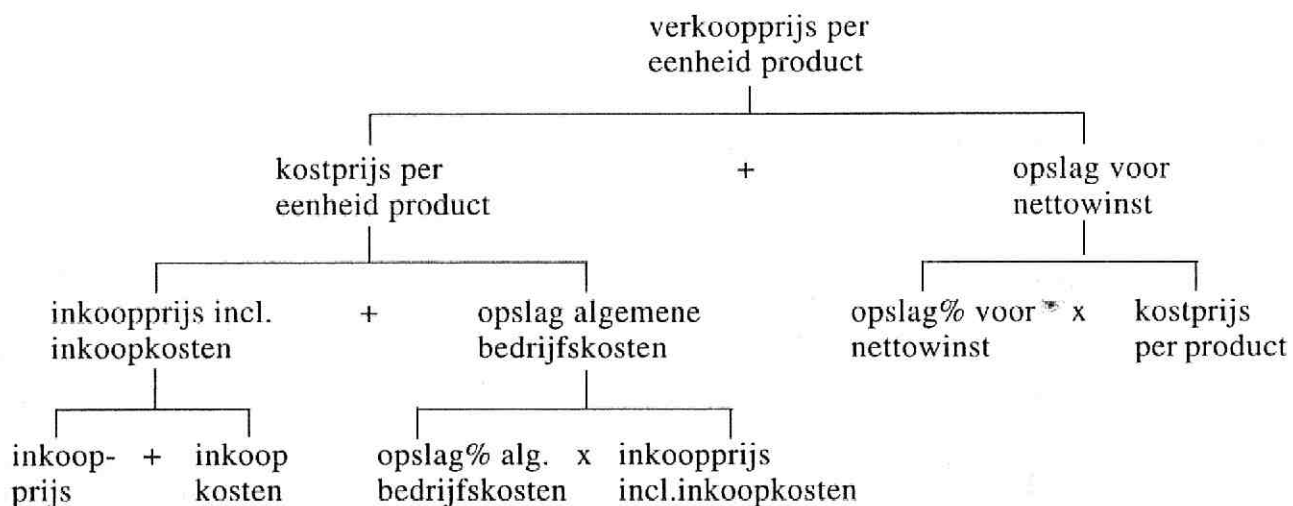
Schema 1: De boekhoudkundige benadering van de nettowinstberekening in een handelsonderneming (aangeduid als model B)



Vragen uit subtest 1 van de test Kennis van Handelingsvoorschriften. Het antwoord op elke vraag is van onder naar boven uit schema 1 af te lezen:

- Hoe berekent een ondernemer doorgaans de kostprijs verkopen?
- Hoe berekent een ondernemer doorgaans de opbrengst verkopen?
- Hoe berekent een ondernemer doorgaans de brutowinst in een periode?
- Hoe berekent een ondernemer doorgaans de nettowinst in een periode?

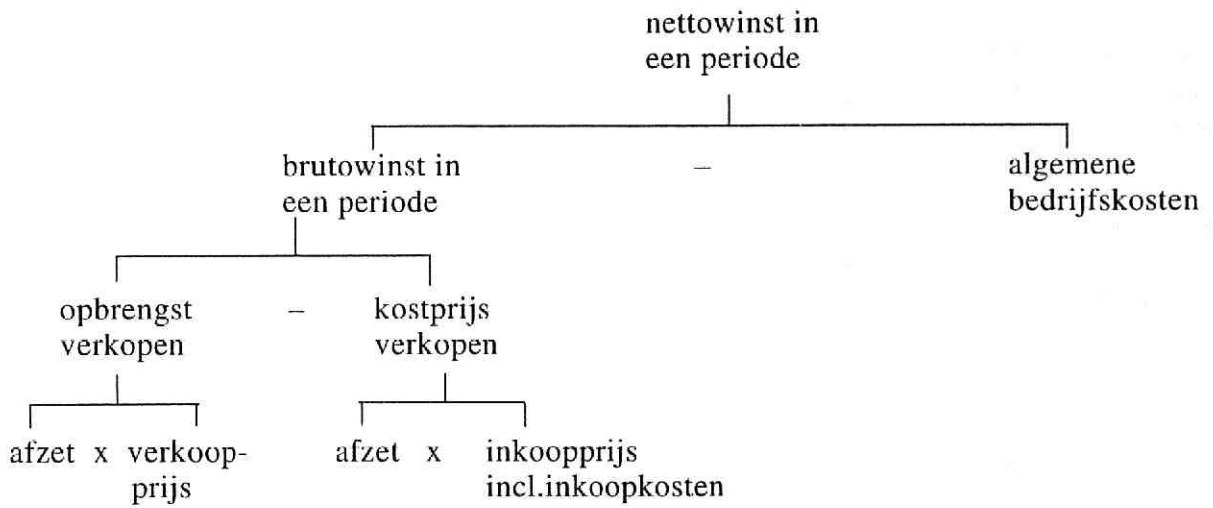
Schema 2: De berekening van de verkoopprijs volgens de calculatorische benadering (aangeduid als model C)



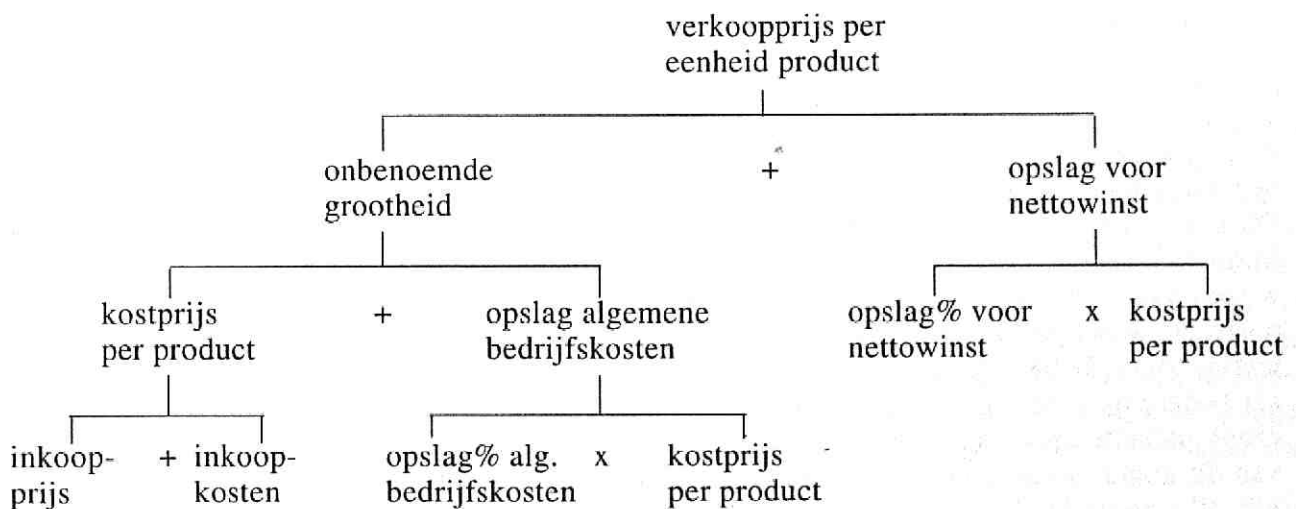
Vragen uit subtest 2 van de test Kennis van Handelingsvoorschriften. Het antwoord op elke vraag is van onder naar boven uit schema 2 af te lezen.

- Hoe berekent een ondernemer doorgaans de kostprijs van een product?
 - Hoe berekent een ondernemer doorgaans de verkoopprijs van een product?
 - Hoe berekent een ondernemer doorgaans de opslag nettowinst per product?
 - Hoe berekent een ondernemer doorgaans de opslag brutowinst per product? (*)
- (*) Antwoord: opslagpercentage brutowinst x inkoopprijs inclusief inkoopkosten.

Schema 3. Een alternatieve benadering van de nettowinstberekening in een handelsonderne-
ming (aangeduid als model A).



Schema 4. Een alternatieve benadering van de berekening van de verkoopprijs (eveneens aange-
duid als model A).



Schooladviezen van allochtone leerlingen

Hans Bosma¹ en Peter Cremers²

¹Rijksuniversiteit Limburg, Vakgroep Medische Psychologie, Maastricht

²Centraal Bureau voor de Statistiek, Afdeling Cohorten, Heerlen

ABSTRACT

This article aims to give a detailed insight into recommendations given to migrant pupils in the Netherlands about the type of secondary education most suited to them. Special attention is paid to the role of the school achievements. The analyses presented are based upon two separate data sets from Statistics Netherlands. The results show that pupils born in the Antilles, Surinam, the Molucca Islands, Turkey or Morocco on average had lower scores on school achievement tests and on average received lower school recommendations than their autochthon counterparts. When school achievements were controlled for, the results of the regression analysis show that the migrant pupils received higher recommendations than the autochthon pupils. This phenomenon has already been reported by other investigators. However, more detailed analyses reveal that the over-recommendation only appeared to occur at the lowest achievement levels. The statistically significant interaction between the ethnicity and the school achievements points at the highest achievement levels to an under-recommendation for migrant pupils compared to non-migrant pupils. Controlling for several covariates, among which social class, hardly affected the reported findings.

INLEIDING

Reeds enkele jaren wordt er door onderzoekers uitvoerig aandacht besteed aan de schoolloopbanen van allochtone leerlingen in vergelijking met die van autochtone leerlingen (van 't Hof & Dronkers, 1993; de Jong, 1987; Koeslag & Dronkers, 1994; van Langen & Jungbluth, 1990; Mulder & Tesser, 1992; Schouten, 1990). De bevindingen van deze onderzoekers lijken erop te wijzen dat er op één moment gedurende de schoolloopbaan sprake is van een zekere vorm van positieve selectie ten voordele van allochtone leerlingen. Dat moment vindt plaats wanneer de leerkracht op de basisschool tot een oordeel moet komen over het meest geschikte type voortgezet onderwijs voor zijn of haar leerlingen. De leerkracht baseert zijn of haar oordeel voor een groot gedeelte op de door de leerling geleverde schoolprestaties. Een thema dat in onderzoek van dit zogenaamde breekpunt in de schoolloopbaan regelmatig opduikt is de overadvisering van allochtone leerlingen ten opzichte van autochtone leerlingen. Bij gelijke schoolprestaties lijken allochtone leerlingen een hoger advies te krijgen dan autochtone leerlingen (van 't Hof & Dronkers, 1993; de Jong, 1987; Kohnstamm, 1993; Tesser & Mulder, 1990). Dit verschijnsel wordt meestal verklaard door de veronderstelling dat adviserende leerkrachten ervan uitgaan dat de eventuele (taal)achterstand van allochtone leerlingen later wel ingehaald wordt.

In een onlangs gepubliceerd artikel van Pijl worden er een aantal kanttekeningen geplaatst bij deze veronderstelde overadvisering van allochtonen (Pijl, 1994). Volgens Pijl berust de aangetroffen overadvisering grotendeels op een statistisch artefact, omdat in een aantal onderzoeken de gemiddelde schoolprestaties van allochtonen en autochtonen met elkaar worden vergeleken binnen categorieën van het schooladvies (Pijl, 1994). Statistisch gezien dient een omgekeerde weg gevolgd te worden; met andere woorden, gemiddelde schooladviezen van allochtone en autochtone leerlingen dienen met elkaar vergeleken te worden binnen categorieën van de schoolprestaties. Verder trekt Pijl de betrouwbaarheid en validiteit van de meting van

¹Rijksuniversiteit Limburg, Vakgroep Medische Psychologie, Postbus 616, 6200 MD Maastricht.

²Centraal Bureau voor de Statistiek, Afdeling Cohorten, Postbus 4481, 6401 CZ Heerlen.

schoolprestaties in twijfel. Tenslotte wijst Pijl op de mogelijkheid dat vooralsnog onbekende factoren de optredende overadvisering kunnen 'wegverklaren'.

In onderhavig artikel wordt, rekening houdend met Pijl's opmerkingen, een additionele kanttekening geplaatst bij de optredende overadvisering. We zullen met name aandacht schenken aan de interactie tussen de etnische herkomst en de schoolprestaties. Meer specifiek wordt de volgende vraagstelling bestudeerd: Is er op elk niveau van de schoolprestaties sprake van overadvisering van allochtone leerlingen? Voor de analyses die nodig zijn voor de beantwoording van deze vraagstelling wordt gebruik gemaakt van data materiaal afkomstig van het Centraal Bureau voor de Statistiek (CBS).

GEGEVENS EN ANALYSES

De gegevens voor dit onderzoek zijn afkomstig uit de door het CBS uitgevoerde cohortonderzoeken 1989 en 1993. In deze onderzoeken zijn in het schooljaar 1989/'90 en in het schooljaar 1993/'94 bij twee representatieve steekproeven van respectievelijk 19.524 en 20.331 leerlingen in de eerste klas van het voortgezet onderwijs een groot aantal (achtergrond) kenmerken vastgesteld, zoals het sociale milieu van herkomst van de leerlingen, het geboorteland van ouders en leerlingen en de schoolprestaties, het advies en de schoolkeuze van de leerlingen. Deze informatie is onder andere verkregen uit een uitvoerige vragenlijst die de ouders hebben ingevuld. De schoolprestaties zijn vastgesteld met behulp van speciaal daarvoor door het CITO ontwikkelde toetsen. Vervolgens is en wordt elk jaar vastgesteld op welke positie in het onderwijs de leerlingen zich bevinden. Zo kunnen relaties gelegd worden tussen de achtergrondkenmerken van de leerlingen enerzijds en de loopbaan door en de bereikte positie in het onderwijs anderzijds.

Alle analyses worden apart voor cohort 1989 en cohort 1993 gepresenteerd. Deze opzet maakt een cross-validatie van resultaten mogelijk. Etnische herkomst wordt in de analyses geoperationaliseerd door het geboorteland van de leerling, omdat het voor cohort 1993 nog niet mogelijk was ook het geboorteland van de ouders in de analyse te betrekken. De kleine aantallen in gedetailleerde categorieën van deze variabele noodzaakten ons tot een enigszins grove categorisering van het geboorteland: 1) Nederland, 2) Antillen, Suriname en Molukken (ASM), 3) Marokko, 4) Turkije en 5) ander land. Het CITO heeft een prestatietoets ontwikkeld met als subtoetsen de onderdelen taal, rekenen en informatieverwerking, elk bestaand uit 20 items. De totaalscore 'schoolprestatie' is berekend door de correcte items van alledrie de toetsen op te tellen. Cronbach's α is gebruikt om de betrouwbaarheid van de metingen van schoolprestaties vast te stellen. Het schooladvies is geoperationaliseerd door, bij meer dan één advies, het hoogst verkregen advies te nemen. Hierdoor ontstaan de volgende categorieën: (1) vbo, (2) mavo, (3) havo en (4) vwo.

De analyses starten met het tonen van de bivariate relaties tussen het geboorteland van de leerling enerzijds en de schoolprestaties en het advies anderzijds. Vervolgens is de relatie tussen het geboorteland en het advies bestudeerd binnen quintielen (20% groepen) van de schoolprestaties. Tenslotte zijn een aantal regressie-vergelijkingen geschat met behulp van de kleinste kwadraten methode. In deze analyses is het geboorteland van de leerling gedummificeerd, waarbij de leerlingen die in Nederland geboren zijn als referentiegroep dienen. In het bijzonder zijn een vijftal modellen geschat: 1) een model waarbij het advies geschat wordt door het geboorteland van de leerling; 2) een model waarbij het advies geschat wordt door de schoolprestaties; 3) een model waarbij het advies geschat wordt door het geboorteland van de leerling en de schoolprestaties; 4) een model waarbij het advies geschat wordt door het geboorteland van de leerling, de schoolprestaties en de interactie tussen beide kenmerken; 5) een model waarbij de invloed van andere variabelen op de relaties in voorgaande modellen wordt bestudeerd. De laatste vier modellen zijn met elkaar vergeleken met een F-toets voor verschillen.

Vergelijking van de resultaten van model 1 en 3 geeft een indicatie van de mate waarin er in cohort 1989 en cohort 1993 sprake is van overadvisering. Wanneer in model 3 de coëfficiënten van het geboorteland (van allochtone leerlingen ten opzichte van in Nederland geboren leerlin-

gen) niet meer, zoals in model 1, negatief zijn, maar (statistisch significant) positief zijn geworden kunnen we spreken over overadvisering van allochtone leerlingen. Of er sprake is van een interactie tussen het geboorteland en de schoolprestaties wordt duidelijk in model 4. In dit model wordt bestudeerd of het effect van de schoolprestaties op het advies verschilt tussen leerlingen met een verschillend geboorteland. Omgekeerd betekent dit een analyse van de samenhang tussen het geboorteland en het advies voor verschillende schoolprestatieniveau's. In deze analyse staat centraal of er op alle niveau's van de schoolprestaties sprake is van overadvisering.

Om na te gaan in hoeverre de aangetroffen relaties op andere kenmerken berusten worden in model 5 tenslotte verschillende andere kenmerken aan model 4 toegevoegd. Met andere woorden hier wordt bestudeerd in hoeverre de gevonden verbanden berusten op schijnrelaties. Hierbij is een groot aantal variabelen onderzocht. Daarbij is niet altijd gebruik gemaakt van gegevens uit beide cohorten. Soms is informatie alleen in één cohort beschikbaar. De volgende variabelen zijn steeds apart toegevoegd aan het model: sociaal milieu van herkomst (geoperationaliseerd door het beroepsniveau van de ouders); het geslacht van de leerling; de gezinssamenstelling (één- en twee-oudergezinnen); opgelopen vertraging (doublures); cognitieve vaardigheden (psb3 en psb8) en het cultureel gezinsklimaat (ouderlijk bezoek aan musca, lezen van boeken en onderwijsondersteuning door ouders). Bovendien wordt in model 5 ook aandacht geschonken aan meer motivationele aspecten van de leerlingen. Het is mogelijk dat de leerkracht, naast de schoolprestaties, ook deze houdingen betreft in het komen tot een schooladvies. Voor de analyse van de invloed van dit kenmerk van leerlingen is een schoolbelevingsschaal (15 items) gebruikt en een schaal die de 'leergierigheid' (21 items) meet. Meer informatie over de meting van genoemde kenmerken kunnen bij de auteurs opgevraagd worden.

In beide cohorten is ook onderzocht in hoeverre allochtonen, in vergelijking met autochtonen, hogere adviezen krijgen bij gelijke cognitieve vaardigheden (psb3 'denkvaardigheid' en psb8 'denkbeeldig abstraheren'). Ook de interactie tussen etnische herkomst en de cognitieve vaardigheden is daarbij betrokken. Voor de achtergrondkenmerken: sociaal milieu, geslacht en de gezinssamenstelling tenslotte is ook bestudeerd of er sprake is van overadvisering bij gelijke schoolprestaties. Ook de interactie tussen deze achtergrondkenmerken en de schoolprestaties is getoetst op significantie.

RESULTATEN

In tabel 1 wordt de samenhang getoond tussen enerzijds het geboorteland van de leerlingen en anderzijds de schoolprestaties en het advies. Cronbach's α van de afzonderlijke 60 schaalitems van de prestatietoets in cohort 1993 is 0,91. Duidelijk is te zien dat de in Nederland geboren leerlingen de hoogste schoolprestaties hebben. Leerlingen die in Marokko zijn geboren scoren in cohort 1989 het slechtst; in cohort 1993 zijn dit de leerlingen die in de ASM landen zijn geboren. De in Turkije geboren leerlingen verschillen niet substantieel van deze leerlingen. De leerlingen die in een ander land zijn geboren nemen een tussenpositie in. Een analyse voor de aparte toetsen voor taal, rekenen en informatieverwerking levert vergelijkbare resultaten op. Een vergelijkbare rangorde wordt aangetroffen ten aanzien van het verkregen advies. De laagste adviezen worden verkregen door de leerlingen die in de ASM landen, Marokko of Turkije zijn geboren. Gemiddeld krijgen zij een advies dat tussen het vbo en mavo in ligt. In Nederland of in een 'ander land' geboren leerlingen krijgen de hoogste adviezen, namelijk gemiddeld iets hoger dan het mavo.

De relatie tussen het geboorteland en de schooladviezen binnen de quintielen van schoolprestaties laat een complex beeld zien (tabel 2). Duidelijk is te zien dat de veronderstelde overadvisering, dat wil zeggen bij gelijke prestatieniveau's hogere adviezen voor leerlingen die niet in Nederland geboren zijn, zich niet consistent in elk quintiel van de prestatiescores manifesteert. Overadvisering komt eigenlijk met name voor bij de leerlingen die in een 'ander land' zijn geboren; zowel in cohort 1989 als in cohort 1993 krijgen deze leerlingen op elk prestatieni-

Tabel 1. Gemiddelde schoolprestaties en schooladviezen (standaarddeviaties) naar geboorteland van de leerling.

		schoolprestaties ¹	schooladvies ²
	steekproef-aantal	gemiddelde (sd)	gemiddelde (sd)
Cohort 1989			
geboorteland			
Nederland	15884	34,73 (11,17)	2,22 (1,03)
ASM ³	143	26,97 (10,87)	1,87 (0,79)
Marokko	227	21,83 (9,56)	1,44 (0,71)
Turkije	159	22,66 (9,75)	1,55 (0,76)
ander land	374	33,24 (11,81)	2,33 (1,07)
onbekend	2737	31,96 (11,52)	2,14 (1,06)
Cohort 1993			
geboorteland			
Nederland	19076	35,38 (11,29)	2,25 (1,05)
ASM ³	114	22,93 (10,09)	1,60 (0,88)
Marokko	208	23,23 (9,50)	1,49 (0,68)
Turkije	147	23,78 (10,50)	1,58 (0,78)
ander land	585	31,84 (32,61)	2,13 (1,07)
onbekend	201	40,65 (9,15)	3,22 (0,94)

¹Maximum aantal punten = 60.

²Het schooladvies heeft de volgende waarden: 1: vbo (het vroegere lbo), 2: mavo, 3: havo en 4: vwo

³Antillen, Suriname of de Molukken.

veau een hoger advies dan de in Nederland geboren leerlingen. Voor de andere allochtone groepen geldt dat er niet consistent hogere adviezen gegeven worden dan aan in Nederland geboren leerlingen die vergelijkbare schoolprestaties hebben. Voor cohort 1989 blijkt dat met name leerlingen die geboren zijn in de ASM landen op het laagste prestatieniveau duidelijk geconfronteerd worden met overadvisering, terwijl dezelfde groep leerlingen, alsmede de leerlingen die in Marokko of Turkije zijn geboren, op het hoogste prestatieniveau te maken hebben met onderadvisering. In cohort 1993 is er bij de laagst presterende leerlingen geen substantieel verschil, terwijl er bij de hoogst presterende leerlingen die in de ASM landen of Marokko zijn geboren sprake is van onderadvisering. In een gecombineerd bestand met grotere aantallen allochtone leerlingen in de hoogste prestatieklassen laten de cijfers duidelijk zien dat er op de hoogste prestatieniveau's geen sprake is van overadvisering, maar van onderadvisering. Bij de laagst presterende leerlingen krijgen met name de leerlingen die in de ASM landen zijn geboren hogere adviezen dan de in Nederland geboren leerlingen. De analyses die apart gedaan zijn voor de taal-, reken- en informatieverwerkingstoets resulteerden in vergelijkbare bevindingen.

De volgende resultaten zijn afkomstig van de verschillende onderscheiden regressie-modellen (tabel 3)². De resultaten in cohort 1989 en cohort 1993 komen in grote mate met elkaar overeen. Model 1 laat zien dat de niet in Nederland geboren leerlingen lagere adviezen krijgen dan de in Nederland geboren leerlingen. Wanneer de schoolprestaties aan dit model worden toegevoegd (model 3), dan verandert de richting van de coëfficiënten van het geboorteland van negatief naar positief. Dit houdt in dat niet in Nederland geboren leerlingen, bij gelijke schoolprestaties, (iets) hogere adviezen krijgen dan leerlingen die in Nederland zijn geboren. Leerlingen die in de ASM landen zijn geboren blijken, bij gelijke prestaties, een ongeveer 0,2 schooltype hoger advies te krijgen dan de in Nederland geboren leerlingen. Model 4 laat echter zien dat deze overadvisering niet op alle niveau's van de schoolprestaties plaatsvindt. Voor de lagere

Tabel 2. Gemiddelde schooladviezen¹ naar geboorteland van de leerling en de schoolprestaties in quintielen.

Prestatieniveau's in 20%-groepen										
1 laagste		2		3		4		5 hoogste		
aantal	M ²	aantal	M	aantal	M	aantal	M	aantal	M	
Cohort 1989										
geboorteland										
Nederland	2610 1,25	3496 1,70	3173 2,15	2840 2,74	2502 3,50					
ASM ³	53 1,51	32 1,88	19 2,16	15 2,47	6 3,33					
Marokko	113 1,29	41 1,46	11 2,00	7 2,86	6 2,67					
Turkije	66 1,30	32 1,59	12 1,83	4 2,50	5 3,20					
ander land	70 1,33	79 1,80	70 2,46	53 3,00	54 3,56					
Cohort 1993										
geboorteland										
Nederland	3298 1,24	3526 1,71	3556 2,17	3501 2,76	3173 3,50					
ASM ³	60 1,28	20 1,50	7 2,71	3 3,33	5 3,20					
Marokko	103 1,26	33 1,73	22 1,91	10 2,40	2 2,50					
Turkije	74 1,20	25 1,80	10 1,80	8 2,63	5 3,80					
ander land	150 1,29	99 1,88	63 2,44	83 2,86	69 3,61					
Combinatie										
geboorteland										
Nederland	6246 1,26	6684 1,71	5743 2,12	6489 2,65	6513 3,45					
ASM ³	117 1,41	48 1,71	21 2,29	22 2,55	12 3,25					
Marokko	219 1,28	71 1,59	31 1,90	17 2,59	10 2,60					
Turkije	146 1,25	51 1,73	21 1,81	11 2,46	12 3,42					
ander land	225 1,31	173 1,85	107 2,37	146 2,84	139 3,55					

¹Het schooladvies heeft de volgende waarden: 1: vbo (het vroegere lbo), 2: mavo, 3: havo en 4: vwo.

²M = gemiddelde schooladvies.

³Antillen, Suriname of de Molukken.

presteerders blijkt er sprake te zijn van een duidelijke overadvisering; bij hogere prestatieniveau's nemen de verschillen in advies tussen de onderscheiden etnische groepen echter af, hetgeen erin resulteert dat er op de hoogste prestatieniveau's zelfs sprake is van onderadvise-ring. Alhoewel de R^2 alleen drie cijfers achter de komma verschilt ten opzichte van model 3, laat een F-toets op verschillen tussen model 3 en 4 een statistisch significant verschil zien, hetgeen er op duidt dat toevoeging van de interactie termen een verbetering van het model betekent. Alleen de specifieke interactie term ten aanzien van de in een 'ander land' geboren leerlingen blijkt in beide cohorten niet statistisch significant te zijn. Tenslotte zijn verschillende variabelen aan het voorgaande model toegevoegd. Geen enkele van deze variabelen kan de interactie tussen het geboorteland van de leerlingen en de schoolprestaties 'wegverklaren'. Sommige kenmerken reduceren de over- en onderadvise-ring, maar doen ze niet teniet. Dit wordt voor het beroepsniveau van de ouders getoond in model 5. Het kleine aantal allochtone leerlingen verhinderde een simultane analyse van meerdere variabelen en hun onderlinge interacties.

Wanneer de analyses ook gedaan worden voor de aparte toetsen van taal, rekenen en informatieverwerking, dan blijkt dat ook hier er zeker geen sprake is van een consistente overadvise-ring. Op de hoogste prestatieniveau's vindt er een onderadvise-ring plaats. Het dient op deze plaats ook te worden opgemerkt dat het geboorteland van de ouders, waarover informatie be-

schikbaar is in cohort 1989, op dezelfde manier gerelateerd is aan het advies. Kinderen met ouders die in de ASM landen, Turkije of Marokko zijn geboren ondervinden bij een laag prestatieniveau een bepaalde mate van overadvisering, terwijl ze bij een hoog prestatieniveau geconfronteerd worden met onderadvisering ten opzichte van kinderen van in Nederland geboren ouders. Wanneer de analyses gerepliceerd werden met de cognitieve vaardigheden van de

Tabel 3. Ongestandaardiseerde regressie-modellen.

	Model 1	Model 2	Model 3	Model 4	Model 5 ²
Cohort 1989					
intercept	<u>2,22</u>	<u>-0,09</u>	<u>-0,12</u>	<u>-0,14</u>	<u>0,25</u>
schoolprestaties		<u>0,07</u>	<u>0,07</u>	<u>0,07</u>	<u>0,06</u>
geboorteland					
ASM ³	<u>-0,36</u>		<u>0,19</u>	<u>0,92</u>	<u>0,88</u>
Marokko	<u>-0,78</u>		<u>0,12</u>	<u>0,77</u>	<u>0,69</u>
Turkije	<u>-0,67</u>		<u>0,09</u>	<u>0,69</u>	<u>0,66</u>
ander land	<u>0,10</u>		<u>0,18</u>	<u>0,20</u>	<u>0,13</u>
interacties					
ASM*prestaties				<u>-0,03</u>	<u>-0,02</u>
Marokko*prestaties				<u>-0,03</u>	<u>-0,02</u>
Turkije*prestaties				<u>-0,03</u>	<u>-0,02</u>
ander land*prestaties				<u>-0,00</u>	<u>0,00</u>
gecorrigeerde R ²	0,01	0,54	0,54	0,54	0,56
F-toets voor verschillen (t.o.v. vorige model)			F(4,15363)= 9	F(4,15359)= 16	F(5,15264)= 115
Cohort 1993					
intercept	<u>2,32</u>	<u>-0,15</u>	<u>-0,15</u>	<u>-0,17</u>	
schoolprestaties		<u>0,07</u>	<u>0,07</u>	<u>0,07</u>	
geboorteland					
ASM ³	<u>-0,69</u>		<u>0,19</u>	<u>0,46</u>	
Marokko	<u>-0,73</u>		<u>0,08</u>	<u>0,80</u>	
Turkije	<u>-0,69</u>		<u>0,10</u>	<u>0,41</u>	
ander land	<u>-0,12</u>		<u>0,17</u>	<u>0,25</u>	
interacties					
ASM*prestaties				<u>-0,01</u>	
Marokko*prestaties				<u>-0,03</u>	
Turkije*prestaties				<u>-0,01</u>	
ander land*prestaties				<u>-0,00</u>	
gecorrigeerde R ²	0,01	0,55	0,55	0,55	
F-toets voor verschillen (t.o.v. vorige model)			F(4,17899)= 9	F(4,17895)= 10	

Model 1: advies naar geboorteland; model 2: advies naar schoolprestaties; model 3: advies naar geboorteland en schoolprestaties; model 4: advies naar geboorteland, schoolprestaties en de interactie tussen beide; model 5: advies naar geboorteland, schoolprestaties en de interactie tussen beide, alsmede het beroepsniveau van de ouders (gedummificeerd)¹

¹ Onderstreepte coëfficiënten hebben een *p*-waarde kleiner dan 0.05

² Model 5 kon niet geschat worden voor cohort 1993, aangezien de informatie over het beroepsniveau van de ouders vooralsnog ontbreekt

³ Antillen, Suriname en Molukken

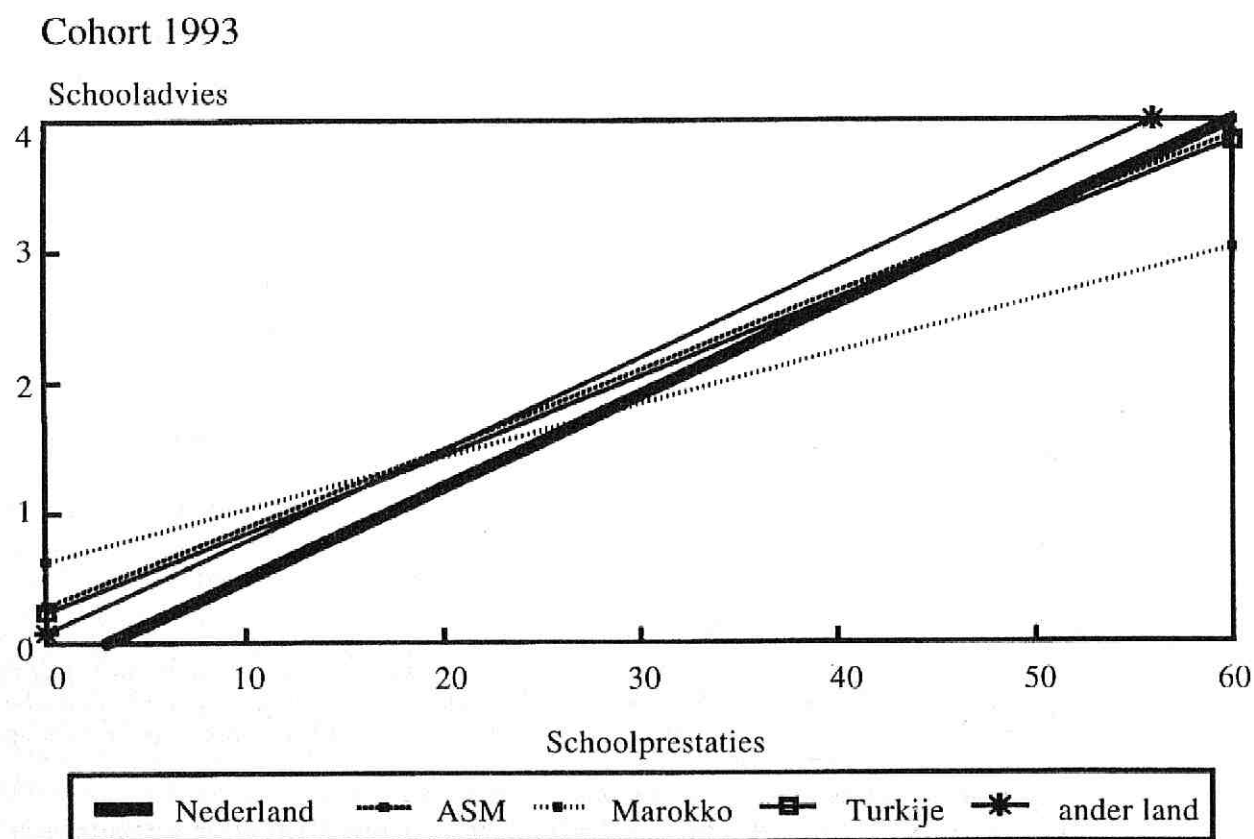
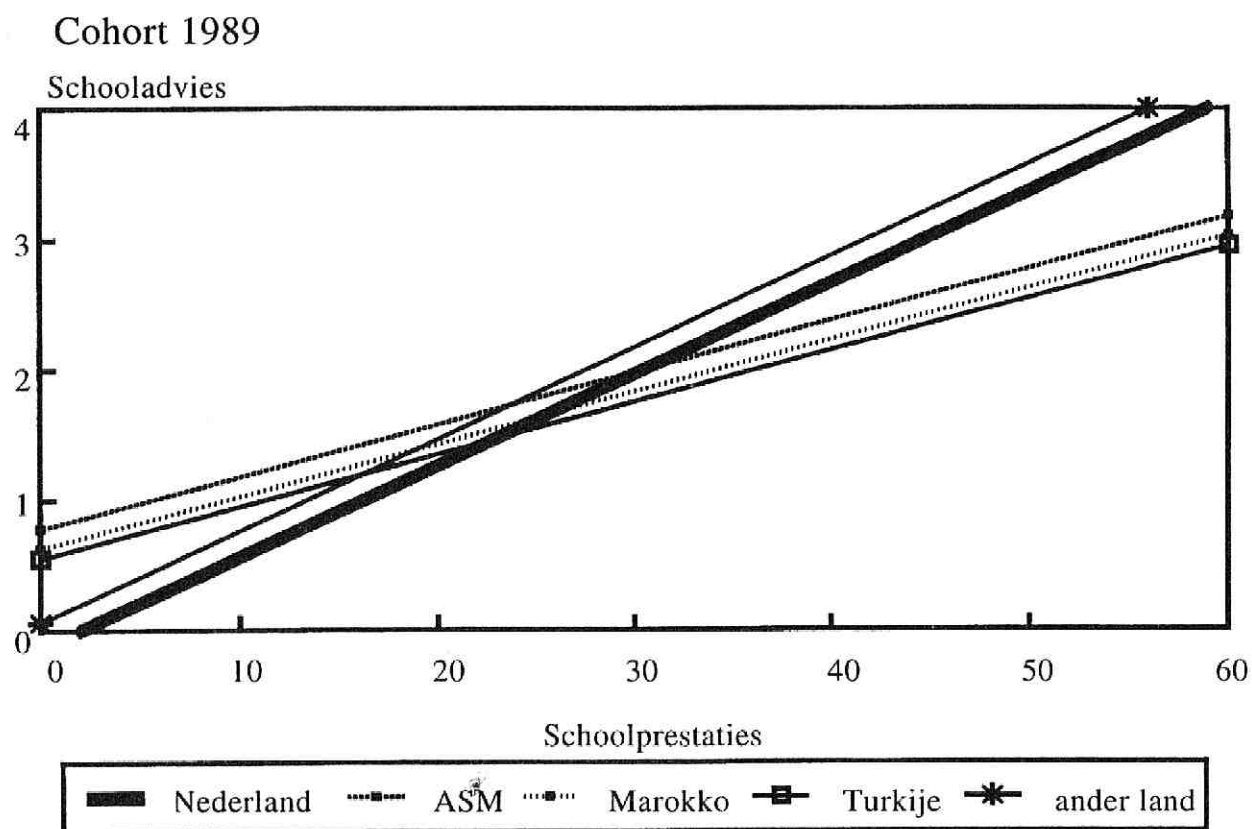


Fig. 1. De door het regressie-model geschatte interactie tussen het geboorteland en de schoolprestaties.

leerlingen (psb3 en psb8) in plaats van de schoolprestaties, dan blijkt er voor dit kenmerk van leerlingen geen over- of onderadvisering op te treden. Bovendien hadden deze vaardigheden geen verstorende werking op de voorgaande bevindingen. Ten aanzien van leerlingen uit de lagere milieus, jongens of kinderen uit één-oudergezinnen trad er geen overadvisering op. Bovendien bestond er geen interactie tussen enerzijds het sociaal milieu van herkomst, het geslacht en de gezinssamenstelling en anderzijds de schoolprestaties.

De resultaten van het regressie-model 4 zijn aanschouwelijk gemaakt in Figuur 1. In beide cohorten blijkt er sprake te zijn van een klassieke interactie tussen het geboorteland en de schoolprestaties (de regressie-lijnen met betrekking tot het effect van de schoolprestaties op het advies van de allochtone en autochtone leerlingen kruisen zich). Alleen voor de leerlingen die in een 'ander land' zijn geboren geldt dat ze op alle niveau's van de schoolprestaties een iets hoger advies krijgen dan de in Nederland geboren leerlingen.

CONCLUSIE EN DISCUSSIE

De gegevens van twee grote CBS cohorten laten zien dat er ten aanzien van de schooladvisering van allochtone leerlingen geen sprake is van een eenduidig beeld. Duidelijk is wel dat leerlingen die geboren zijn in de Antillen, Suriname, Molukken, Turkije of Marokko gemiddeld genomen zowel lagere schoolprestaties als lagere schooladviezen hebben dan in Nederland geboren leerlingen. Bij de tot nu toe gebruikelijke modellering door middel van regressie-analyse blijkt verder ook dat deze allochtone leerlingen bij gelijke schoolprestaties (iets) hogere adviezen krijgen dan leerlingen die in Nederland geboren zijn. Bij nadere beschouwing blijkt deze overadvisering echter niet over de hele range van de schoolprestaties plaats te vinden. Een regressie-analyse waarbij ook de interactie tussen het geboorteland en de schoolprestaties van de leerling wordt gemodelleerd, toont aan dat de aangetroffen overadvisering zich alleen manifesteert op de laagste prestatieniveau's. Op de hoogste prestatieniveau's blijken allochtone leerlingen juist lagere adviezen te krijgen dan autochtone leerlingen. De enige uitzondering wordt gevormd door de groep leerlingen die in andere dan bovengenoemde landen zijn geboren. Dit zijn met name leerlingen die in andere landen rond de Middellandse Zee of Azië zijn geboren. Deze groep krijgt over de hele linie iets hogere adviezen dan in Nederland geboren leerlingen. Ook de analyse waarbij de schoolprestaties ingedeeld zijn in quintielen ondersteunt de bevinding dat er geen sprake is van een consistente overadvisering van allochtone leerlingen. De kans op een toevalstreffer wordt aanzienlijk verkleind doordat de bevinding niet toegeschreven kan worden aan de eigenaardigheden van één steekproef. De analyses op zowel cohort 1989 als cohort 1993 komen tot hoogst vergelijkbare resultaten. Wel dient hier te worden opgemerkt dat de gevonden effecten relatief klein zijn.

Rekeninghoudend met Pijl's opmerkingen ten aanzien van het fenomeen overadvisering, moeten we dus tot de conclusie komen dat er, afhankelijk van het prestatieniveau, sprake is van zowel over- als onderadvisering (Pijl, 1994). De samenhang tussen het geboorteland van de leerlingen en het advies is gecontroleerd voor de prestaties en niet andersom. De prestaties zijn betrouwbaar vastgesteld, gezien de hoge Cronbach's α van de prestatieschaal. Tenslotte blijkt er bij controle voor een groot aantal andere kenmerken (sociaal milieu, cognitieve vaardigheden, motivationele aspecten etc.) geen sprake te zijn van een schijnsamenhang. Dit betekent onder andere dat de van in Nederland geboren leerlingen afwijkende advisering van allochtone leerlingen niet verklaart kan worden, doordat de leraar op de basisschool aan hun systematisch een andere (prestatie-)motivatie of leergierigheid toeschrijft. De meest invloedrijke variabele is het sociaal milieu van herkomst. Rekeninghoudend met dit kenmerk, blijft de interactie tussen het geboorteland van de leerling en de schoolprestaties toch een statistisch significante invloed behouden op de schooladvisering. Of deze bevinding overeind blijft, wanneer meerdere variabelen tegelijkertijd in het model opgenomen worden óf wanneer interacties tussen het geboorteland en de controle variabelen in het model opgenomen worden, moet toekomstig onderzoek met grotere aantallen allochtone leerlingen uitwijzen.

Uitgangspunt bij de analyses was het gebruikelijke Wisconsin model, waarbij de schoolprestaties verondersteld worden vooraf te gaan aan de schooladvisering en de schoolkeuze. De schoolprestaties zijn in beide cohorten echter na de schoolkeuze vastgesteld. Redenerend vanuit de plausibiliteit van het Wisconsin model lijkt het echter onaannemelijk dat de resultaten hierdoor vertekend zijn. Controle voor het schooltype waarop de leerling zich bevindt heeft bovendien geen invloed op de aangetroffen interactie. Terzijde dient te worden opgemerkt dat er geen enkele indicatie kon worden aangetroffen voor het fenomeen 'over-positionering'. Bij een gelijke advisering (en prestatieniveau) gaan allochtone leerlingen in zowel cohort 1989 als cohort 1993 naar een lager schooltype dan hun in Nederland geboren medeleerlingen. Tenslotte dient in dit kader te worden opgemerkt dat de invloed van de adviseringsstrategie van de leraar op de basisschool niet kon worden vastgesteld, omdat het advies pas in de eerste klas van het voortgezet onderwijs is gemeten. Multi-level analyses met betrekking tot de invloed van de op klas- of schoolniveau geaggregeerde gegevens over het schooladvies of de etnische samenstelling van de klas konden dan ook niet uitgevoerd worden.

Over de verklaring van onze bevindingen met betrekking tot de (niet in een 'ander land' geboren) allochtone leerlingen kunnen we alleen speculeren. Het lijkt niet onwaarschijnlijk dat de leraar op de basisschool allochtone leerlingen met een laag prestatieniveau wat vaker het voordeel van de twijfel geeft. Leraren verwachten misschien dat deze allochtone leerlingen hun achterstand op latere leeftijd inhalen, waardoor het in de ogen van de leraar zonde zou zijn om hun voorbarig een te laag schooladvies te geven. Aan de andere kant worden de mogelijkheden en capaciteiten van allochtone leerlingen met een hoog prestatieniveau wat vaker lager ingeschat dan die van in Nederland geboren medeleerlingen. Deze onderschatting van de capaciteiten van hoog presterende allochtone leerlingen kan gebaseerd zijn op hun relatief geringe aantal binnen de groep hoog presterende leerlingen. Het beeld dat leraren hebben van allochtone leerlingen als gevolg van hun in het algemeen minder goede schoolprestaties kan wellicht een nadelige invloed hebben bij de beoordeling van individuele hoog presterende allochtone leerlingen. Het lijkt daarom aanbevelenswaardig de motiveringen die de leraar hanteert bij het komen tot een advies aan allochtone leerlingen nader onder de loep te nemen.

SAMENVATTING

Het doel van dit artikel is om meer inzicht te krijgen in de schooladvisering van allochtone leerlingen. Met name wordt er aandacht geschonken aan de rol van de schoolprestaties. De in dit artikel beschreven analyses zijn uitgevoerd op twee data bestanden van het Centraal Bureau voor de Statistiek. De resultaten tonen aan dat leerlingen die geboren zijn in de Antillen, Suriname, Molukken, Turkije of Marokko gemiddeld lagere schoolprestaties hebben en gemiddeld lagere schooladviezen krijgen dan hun autochtone medeleerlingen. Regressie-analyse toont verder aan dat deze allochtone leerlingen bij gelijke prestaties hogere adviezen krijgen dan autochtone leerlingen. De aangetroffen overadvisering blijkt echter alleen op de laagste prestatieniveau's plaats te vinden. De statistisch significante interactie tussen de etnische herkomst van de leerling en de schoolprestaties laat zien dat er bij bovengenoemde allochtone leerlingen met de beste schoolprestaties juist een substantiële onderadvisering plaatsvindt ten opzichte van de best presterende autochtone leerlingen. Controle voor andere variabelen, waaronder het sociaal milieu van herkomst, blijkt nauwelijks invloed te hebben op deze bevindingen.

NOTEN

1. Hans Bosma was ten tijde van dit onderzoek werkzaam bij het Centraal Bureau voor de Statistiek, afdeling Cohorten. Peter Cremers is hoofd van deze afdeling. De in dit artikel weergegeven opvattingen zijn van de auteurs en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek.

- 2 In de geanalyseerde modellen waren de residuen steeds normaal verdeeld. Bovendien bleek de samenhang tussen de schoolprestaties en het advies bij benadering lineair te zijn. Het verwijderen van uitschieters leidde in geen enkel geval tot zwakkere samenhangen dan gepresenteerd. Een replicatie van de analyses met het logistische regressie model, waarbij de afhankelijke variabele 'schooladvies' gedichotomiseerd werd, leverde tenslotte vergelijkbare resultaten op.

LITERATUUR

- Hof, L. van 't & Dronkers, J. (1993). Onderwijsachterstanden van allochtonen: klasse, gezin of etnische cultuur? *Migrantenstudies*, 9, 2-25.
- Jong, M.J. de (1987). *Herkomst, kennis en kansen. Allochtone en autochtone leerlingen tijdens de overgang van basis- naar voortgezet onderwijs*. Lisse: Swets & Zeitlinger.
- Koeslag, M. & Dronkers, J. (1994). Overadvisering en de schoolloopbanen van migrantenleerlingen en autochtone leerlingen in het voortgezet onderwijs. *Tijdschrift voor Onderwijsresearch*, 19, 240-258.
- Kohnstamm, R. (1993). Hogerop. *NRC-Handelsblad*, 25 februari.
- Langen, A. van & Jungbluth, P. (1990). *Onderwijskansen van migranten. De rol van sociaal-economische en culturele factoren*. Lisse: Swets & Zeitlinger.
- Mulder, L. & Tesser, P. (1992). *De schoolkeuzen van allochtone leerlingen*. Nijmegen: ITS.
- Pijl Y.J. (1994). De overschatten overadvisering van allochtonen. *Tijdschrift voor Onderwijswetenschappen*, 24, 240-246.
- Schouten, S.P. (1990). Schoolloopbanen van Turken en Marokkanen in het voortgezet onderwijs. In M. du Bois-Reymond & L. Eldering (red.). *Nieuwe oriëntaties op school en beroep: de rol van sexe en etniciteit* (pp. 163-176). Lisse: Swets & Zeitlinger.
- Tesser, P. & Mulder, L. (1990). Concentratie en prestatie van allochtone leerlingen in het basisonderwijs. *Migrantenstudies*, 6, 31-45.

Manuscript ontvangen 17-11-1995

Definitief aanvaard 22-3-1996

Notities en Commentaren

SCORINGSREGELS EN TOETSINSTRUCTIE

A. Dirkzwager

Het is een goed gebruik om voorafgaand aan de toetsafname een toetsinstructie te geven over hoe proefpersonen een – voor hen – maximaal haalbare score kunnen krijgen. Vanuit de scoringsregel kan mathematisch afgeleid worden welke aanwijzingen dat moeten zijn. Het is zeer wel mogelijk dat daardoor een antwoord-strategie bevorderd wordt die in strijd is met het doel dat men met de toets voor ogen heeft: door gissen wordt de betrouwbaarheid en validiteit van de metingen verlaagd. Derhalve worden regelmatig zonder voldoende mathematische fundering andere scoringsregels met andere toetsinstructies, die intuïtief meer aanspreken, voorgesteld en in praktijk gebracht. Dit kan ten nadele van de proefpersonen gaan die, als zij de instructies opvolgen, hun te verwachten score niet meer maximaal maken.

Mathematisch ligt de zaak eenvoudig. Een bepaalde proefpersoon heeft, gegeven zijn kennis van zaken, zijn algemene intelligentie en zijn taalgevoel, doorgaans een kans p , groter dan "toeval", om een vraag goed te beantwoorden. Hij geeft het antwoord r dat hem het meest waarschijnlijk juist lijkt. Voor die vraag krijgt hij een itemscore $S(r)$: meestal een score $S(t)$ als zijn antwoord "goed" is en een score $S(f)$ als zijn antwoord "fout" is, waarbij $S(t) > S(f)$. De *verwachte* itemscore is nu:

$$E(S) = p \times S(t) + (1-p) \times S(f) \quad (1)$$

De (ruwe) *totaalscore* op de test is de som van de itemscores; voor $S(t)=1$ en $S(f)=0$ is dit bijvoorbeeld het aantal goed. De *verwachte totaalscore* is in dit geval de som van de verwachte itemscores. Om de verwachte totaalscore te maximaliseren is het voldoende als men voor elk item de verwachte itemscore maximaliseert.

We zullen voor enkele scoringsregels nagaan of de bijbehorende toetsinstructie hiermee in overeenstemming is.

MEERKEUZEVRAGEN, RUWE SCORE

De toetsinstructie is: "Streep altijd het antwoord aan dat je het meest waarschijnlijk juist lijkt, laat geen van de vragen onbeantwoord, als je het niet weet moet je gokken".

Deze instructie is correct. Als men hem opvolgt maximaliseert men zijn itemscores en dus de totaalscore. Immers, als men niets invult is men zeker van een score van nul punten, vult men iets in dan leert toepassing van (1) dat de verwachte score gelijk is aan de kans p die groter of gelijk aan nul is. Men verliest dus niets en wint meestal als men een alternatief aanstreept.

MEERKEUZEVRAGEN MET CORRECTIE VOOR GISSEN

De meest gebruikelijke correctie voor gissen is het aantal goed te verminderen met een proportie van het aantal fout

$$S = G - 1/(k-1) \times F \quad (2)$$

Dit betreft de totaalscore. De vraag is echter wat de (te maximaliseren) *itemscore* in dit geval is. De volgende scoringsregel geeft dezelfde “gecorrigeerde” totaalscore als (2):

$$\begin{aligned} S(t) &= 1 && \text{(goed ingevuld)} \\ S(o) &= 0 && \text{(niet ingevuld)} \\ S(f) &= -1/(k-1) && \text{(fout ingevuld)} \end{aligned} \quad (5)$$

Substitueren van (5) in (1) geeft:

$$E(S) = p - (1-p)/(k-1) = px(1 + 1/(k-1)) - 1/(k-1)$$

of:

$$E(S) = (pxk)/(k-1) - 1/(k-1) \quad (6)$$

Voor $p \geq 1/k$ is dit altijd positief; men kan dus altijd winst verwachten als men iets invult. Als men niets invult is men zeker van een score gelijk aan nul. De beste strategie is hier dus ook om altijd iets in te vullen, ook al moet men gissen.

Om dit gissen te ondervangen zijn voorstellen voor “zekerheidsscore” voorgesteld en in praktijk gebracht. Deze zullen we nu bespreken.

SYSTEEM “MAASTRICHT”

Aan de medische faculteit van de universiteit van Maastricht heeft men een bijzonder interessante onderwijskundige benadering: probleemgestuurd onderwijs. Hiermee gepaard gaat een beoordelingssysteem waarbij vanaf het eerste tot het laatste studiejaar problemen over het gehele vakgebied gepresenteerd worden in de vorm van een uitspraak waarvan de student aan moet geven of hij juist dan wel onjuist is, met een derde antwoordoptie “vraagteken” als indicatie dat men het niet weet. Een goede beoordeling geeft een punt, bij foutieve beoordeling krijgt men -1 punt, een vraagteken levert nul punten op. De instructie is, dat men het vraagteken moet aanstrepen als men het antwoord niet (zeker) weet. We zullen nagaan in hoeverre deze instructie juist is (tot maximalisering van de verwachte score leidt).

“Vraagteken” geeft met zekerheid nul punten. Als “juist” of “onjuist” wordt aangestreept is de kans p dat men het goede antwoord geeft 50% of meer (tenzij het een “strikvraag” is). Toepassing van formule (1) leert dat als men een keuze maakt de verwachte score altijd hoger of gelijk aan nul is, de score die men zeker krijgt als men het vraagteken kiest. De verwachte score wordt dus gemaximaliseerd als men *nooit* het vraagteken kiest -hoe onzeker men over zijn antwoord ook moge zijn. Toevoeging van de vraagteken-optie heeft dus weinig zin en de gebruikte instructie is misleidend.

Het systeem “Maastricht” is een (niet geslaagde) poging om door middel van zekerheidsindicaties de effecten van (gedwongen) gokken – onbetrouwbaarheid en inflatie van ruwe score – te beteugelen. Andere dergelijke benaderingen zullen we hieronder bespreken.

MEERKEUZEVRAGEN MET ZEKERHEIDSAANDUIDING

Als representatief voor deze benadering nemen we het scoringssysteem dat door Sandbergen bij zijn proefschriftonderzoek gebruikt is. Hij liet proefpersonen het (volgens hen meest waarschijnlijk) juiste antwoord kiezen en vervolgens opgeven of zij “zeker” of “onzeker” waren.

De scoringsregel was:

- goed en zeker: 2 punten.
- goed en onzeker: 1 punt
- ?: niet geweten: 0 punten

fout en onzeker:	-1 punt
fout en zeker:	-2 punten

Dit lijkt een zeer redelijk scoringssysteem, maar laten we eens exact bekijken wat voor consequenties het heeft voor de te verwachten score.

Het meest waarschijnlijke alternatief wordt gekozen. De kans op juist is p . Dan zijn er drie antwoordopties: "zeker", "onzeker" en "vraagteken". Volgens (1) zijn de verwachte scores:

$$\text{als "zeker": } E(s) = 2p - 2(1-p) = 4p - 2 \quad (7)$$

$$\text{als "onzeker": } E(s) = p - (1-p) = 2p - 1 \quad (8)$$

$$\text{als "vraagteken": } E(s) = 0 \quad (9)$$

Hieruit blijkt dat voor "zeker" $E(s) \geq 0$ als $p \geq .50$. Dit geldt ook voor de keuze "onzeker". Dit betekent dat men altijd "zeker" (of "onzeker") moet invullen als de kans-op-goed groter is dan 50%, en dat men als die kans kleiner is dan 50% moet kiezen voor het vraagteken. Dit betekent nauwelijks een terugdringen van gokken: zodra men een hogere kans dan 50% op "goed" heeft moet men gokken om zijn verwachte score te maximaliseren; "zeker" en "onzeker" maken geen verschil in verwachte score. "Zekerheidsscore" geeft dus ook geen oplossing van het probleem dat door gokken de score onbetrouwbaarder wordt.

MEERKEUZEVRAGEN MET KANSSCHATTINGEN

"Zekerheidsaanduiding" heeft een aantal nadelen:

1. De schaal waarop zekerheid aangegeven wordt is subjectief. Waar de een zegt dat hij "zeker" is zal de ander aangeven dat hij "onzeker" is; tussen beide categorieën is geen duidelijke objectieve grens.
2. Er is geen objectief criterium om te checken of iemand zijn "zekerheid" correct heeft gerapporteerd: het is zeer wel mogelijk dat men opgeeft "zeker" te zijn terwijl men in feite zeer "onzeker" is.
3. Men wordt geforceerd een alternatief te kiezen, ook als men bijv. tussen twee van de vijf alternatieven aarzelt. Als een van die twee het goede is zou men een hogere score moeten krijgen dan wanneer beide fout zijn. Dit wordt bij zekerheidsscore niet gehonoreerd: als men van de twee toevallig het verkeerde gokt krijgt men zelfs een negatieve itemscore.

Deze bezwaren kunnen ondervangen worden als men de proefpersoon voor elk alternatief laat aangeven hoe waarschijnlijk het is dat het het juiste is. Dat moet hij voor zichzelf bij meerkeuzevragen ook doen voordat hij een keuze maakt: onwaarschijnlijke alternatieven worden uitgesloten en het meest waarschijnlijke alternatief wordt aangestreept.

Het is duidelijk welke toetsinstructie men wenst: "Probeer je kansen zo goed mogelijk in te schatten en rapporteer ze eerlijk". Niet zo duidelijk is onder welke scoringsregel het opvolgen van die instructie tot een maximaal te verwachten score leidt.

Een betere methode dan een trial-and-error benadering is om de scoringsregel bij een gewenste test-instructie wiskundig af te leiden. We stellen dan als eis dat de verwachte score maximaal wordt als men de kansen $r(i)$ eerlijk en realistisch rapporteert. De itemscore wordt met de scoringsregel F berekend uit de gerapporteerde kansen $r(i)$: $F(r(i))$ is dus de itemscore als i het juiste antwoord is. De verwachte score is dan:

$$E(s) = \sum p(i) \times F(r(i)) \quad (11)$$

Dit moet maximaal zijn voor $r(i)=p(i)$ voor alle alternatieven i .

We differentiëren (11) naar $r(i)$ onder de bijvoorwaarde $\sum r(i)=1$ en vinden, als we de eerste afgeleide op nul stellen voor alle $p(i)=r(i)$, dat deze afgeleide gelijk is aan:

$$F'(r(i)) = 1/r(i) \quad (12)$$

Dit betekent dat F een lineaire functie moet zijn van de logaritme van $r(i)$: alleen voor $G(n) = A \ln(n) + B$ is $G' = 1/n$.

De itemscore moet dus berekend worden als:

$$S = A \ln(r(t)) + B \quad (13)$$

waarbij t het juiste antwoord is en A en B vrij te kiezen constanten.

Interessant is dat de verwachte score $E(s)$ gelijk is aan de "uncertainty" zoals gedefinieerd in Shannon en Weaver's informatietheorie voor $A = 1/\ln(2)$ en $B=0$.

Voor toetsingsdoeleinden geven we aan andere waarden voor A en B de voorkeur. We stellen twee eisen:

1. Bij perfecte kennis ($r(t)=1.0$) krijgt men de maximale itemscore van 100 punten.
2. Bij "perfecte" onwetendheid ($r(i)=1/k$ voor alle k alternatieven i , dus ook: $r(t)=1/k$) krijgt men als itemscore nul punten.

Hiermee kunnen twee vergelijkingen worden opgesteld waaruit A en B kunnen worden opgelost. We krijgen dan:

$$S(i) = 100 \times (\ln(r(i))/\ln(k) + 1) \quad (13)$$

de itemscore $S(i)$ als i het juiste alternatief is.

OPTIMALE TOETSINSTRUCTIE EN BIJBEHORENDE SCORINGSREGEL

Het schatten van persoonlijke kansen is geen alledaagse bezigheid en de scoringsregel is weinig inzichtelijk. Daarom moet de toetsinstructie vanuit een andere invalshoek gegeven worden. We maken gebruik van het feit dat tussen $S(i)$ en $r(i)$ een een-op-een relatie bestaat.

In plaats van de kansen $r(i)$ te laten opgeven kunnen we de proefpersoon ook vragen zelf de $S(i)$ te bepalen: het aantal punten winst (of verlies) dat ze willen krijgen als een bepaald alternatief i het juiste zou blijken te zijn.

Dit is alleen mogelijk als de toets met een interactief computerprogramma wordt afgenomen. De computer zorgt dan automatisch dat aan (13) voldaan wordt onder de voorwaarde dat de som van de kansen $r(i)$ gelijk is aan 1. Terwijl de proefpersonen hun inzet bepalen (de vraag beantwoorden) zien zij direct hoeveel punten ze winnen cq. kunnen verliezen voor elk alternatief *als dat alternatief het juiste zou blijken te zijn*. Ze zien nu direct welke scores ze kunnen verwachten: een groot verlies op een alternatief dat zeker onjuist is kan altijd geriskeerd worden; op alternatieven die mogelijk juist zouden kunnen zijn moet men geen verlies riskeren; voor alternatieven die men waarschijnlijk juist acht kan men winst claimen.

Een toetsinstructie in deze geest is veel begrijpelijker dan een instructie voor het schatten van kansen. Bovendien kan men bij computerpresentatie van de toets direct na de beantwoording laten zien welk antwoord het juiste is en hoeveel punten men op die vraag gewonnen cq. verloren heeft. Dit is een "spel" dat proefpersonen al doende snel leren spelen, ook als de eerste instructie nog wat abstract was. Het resultaat is per item een meting op een ratio-schaal (de gebruikelijke scoringsregel bij meerkeuzevragen geeft slechts een dichotome "meting": goed of fout.)

BETROUWBAARHEID VAN DE METING ONDER VERSCHILLENDE TOETSINSTRUCTIES

We willen vaststellen wat de meetfout is *op item-niveau* bij de *individuele meting* van (de kennis

betreffende dit item van) *een bepaalde proefpersoon*. Dat gaat als volgt:

De verwachte score voor een item is gelijk aan:

$$E = \sum p(i) \times S(i) \quad (14)$$

De variantie is:

$$V = \sum p(i) \times (S(i) - E)^2 \quad (15)$$

Aangenomen dat $p(i)$ de ware kansen zijn en dat de scoringsregel correct is (d.w.z. dat we geloven dat $S(i)$ een goede maat is voor wat we met het item meten), kunnen we E de "ware" score noemen en V de error-variantie (meetfout). We zullen nu eerst voor een paar extreme gevallen voor verschillende toetsinstructies met bijbehorende scoringsregels nagaan hoe groot deze error variantie is.

In het eerste geval, dat de proefpersoon het antwoord op de vraag perfect weet is $p(i)=1.0$ en $p(i)=0$ voor alle i ongelijk i . In dat geval wordt $E=S(i)$ en $V=0$. We krijgen in dit geval dus een perfecte meting ongeacht de scoringsregel, de meetfout is altijd nul.

In het tweede geval nemen we aan dat de proefpersoon geen notie heeft wat het juiste antwoord zou kunnen zijn: $p(i) = 1/k$ voor alle i . In dat geval is:

$$E = (1/k) \times S(i) \quad (16)$$

en:

$$V = (1/k) \times \sum (S(i)-E)^2 \quad (17)$$

Bij tradionele scoring van meerkeuzevragen wordt dit: $E = 1/k$ en

$$V = (1/k) \times ((k-1)/k)^2 + ((k-1)/k) \times (1/k)^2 \quad (18)$$

of:

$$V = 2 \times (k-1)/k^3 \quad (19)$$

Voor een lage variantie (nauwkeuriger meting) moet men k zo groot mogelijk kiezen. Dit geeft echter praktische problemen omdat het zeer moeilijk is om bij een Multiple Choice vraag een groot aantal voor een onwetende even aantrekkelijke alternatieven te bedenken. Voor een gebruikelijke $k=5$ wordt $E=.20$ en $V=.064$, voor $k=2$ is $E=.50$ en $V=.25$, een zeer grote meetfout.

Bij de toetsinstructie/scoringsregel die hierboven is geschetst, waarbij de proefpersoon zelf zijn score (punten winst cq. verlies) bepaalt, afhankelijk van zijn persoonlijke waarschijnlijkheden $r(i)$, is, wanneer hij het antwoord geheel niet weet en wanneer hij eerlijk en realistisch antwoordt, voor alle i $r(i)=p(i)$ en $S(i)=0$ voor $p(i)=1/k$. Bij deze instructie en scoringsregel wordt dus $E=0$ en $V=0$. We krijgen ook in dit geval dus -in tegenstelling tot traditionele scoring waar de meetfout gelijk is aan $V = 2 \times (k-1)/k^3$ (17)- een perfect accurate meting zonder meetfout.

REALISME TOETSBAAR PER PROEFPERSOON

Onder de nulhypothese dat een proefpersoon volkomen realistisch is ($r(i)=p(i)$ voor alle i) is de itemscore $S(i)$ een stochastische variabele met gemiddelde en variantie zoals in (14) en (15) gegeven. We transformeren de itemscore tot een z-score: $z = (S(i)-E)/\sqrt{V}$. Als we deze z-scores sommeren over een voldoende groot aantal (N , $N>6$) items krijgen we (onder de nulhypothese) een bij benadering normaal verdeelde variabele $z(\text{totaal})$ met gemiddelde nul en variantie N . Hiermee kan de nulhypothese tweezijdig getoetst worden.

Als de nulhypothese verworpen wordt zijn er twee mogelijkheden: of $z(\text{totaal}) < 0$ of $z(\text{totaal}) > 0$.

Als $z(\text{totaal})$ negatief is, is de verkregen score lager dan de (persoonlijk) verwachte score (de score die de proefpersoon probeerde te halen): de proefpersoon heeft zichzelf *overschat*, eventueel door met (onterechte) hoop op winst zwaar te gokken.

Als $z(\text{totaal})$ positief is, is de verkregen score hoger dan de (persoonlijk) verwachte score (de hoogste score die de proefpersoon dacht te kunnen halen): de proefpersoon heeft zichzelf *onderschat*, hij is te bescheiden of te voorzichtig geweest, met zijn kennis had hij meer winst kunnen claimen en meer verlies kunnen riskeren om zo een hogere score te krijgen.

Ook als men niet op de dichotome beslissing uit is om de nulhypothese of te accepteren of te verwerpen, geeft $z(\text{totaal})$ een goede maat voor zelf-overschatting vs. zelf-onderschatting. Een waarde in de buurt van nul betekent dat de proefpersoon realistisch was.

Het is een empirisch te beantwoorden vraag of- en hoe snel- proefpersonen kunnen leren realistisch te zijn.

SLOTOPMERKING

Essentieel voor de scoringsregel is dat -onder bepaalde wel-gedefinieerde restricties- de proefpersoon zelf de gewichten voor de alternatieven bepaalt: het aantal punten dat hij wint (of verliest) als dat alternatief juist is. Proefpersonen met veel kennis kunnen veel winst claimen en veel verlies "riskeren"; proefpersonen, die geen idee hebben wat het juiste antwoord is, claimen geen winst en riskeren dan ook geen verlies.

Essentieel voor deze toetsmethode is dat proefpersonen realistisch zijn in de risico's die zij nemen: zij moeten niet "gokken" en nooit meer winst claimen of meer verlies riskeren dan vanuit wat zij weten gerechtvaardigd is. Of een proefpersoon realistisch is kan statistisch getoetst worden en het is met geëigende feedback mogelijk proefpersonen te leren meer realistisch te worden.

De methode (inclusief trainingsfeedback) is geïmplementeerd in het interactieve toetsprogramma "TestBet" dat door de uitgeverij Zwijsen op de markt gebracht wordt. Het geeft een zeer volledige resultatenrapportage tot op het (diagnostisch belangrijke) itemniveau. Aansluitend bevat het ook procedures voor item- en test-analyse. Een volledige demo is op aanvraag beschikbaar bij de auteur: E-mail: aried@xs4all.nl

Ontvangen 11-10-1995

Notities en Commentaren

HET TOEKOMSTPERSPECTIEF VAN DE ONDERWIJSKUNDE¹

Len F.W. de Klerk²

Katholieke Universiteit Brabant

INLEIDING

Het is inmiddels veertien jaar geleden dat de ORD voor de eerste keer in Tilburg plaatsvond. Mij was toen verzocht de openingsrede te houden, niet als rector namens de Katholieke Universiteit Brabant, maar als onderwijspsycholoog namens de toenmalige vakgroep onderwijspsychologie. De titel van mijn rede luidde: *Uitdaging en teleurstelling*.

Gaarne heb ik voldaan aan het verzoek van het organiserend comité om in het kader van de opening van de onderwijsresearchdagen 1996 een toespraak te houden. De titel van deze rede luidt: *Het toekomstperspectief van de onderwijskunde*. Ik maak van de gelegenheid gebruik om met u stil te staan bij enkele saillante ontwikkelingen die zich in de loop van de laatste veertien jaar hebben voltrokken.

De punten waarbij ik wil stilstaan zijn achtereenvolgens: interdisciplinariteit, internationalisering en onderzoekscholen. Ik sluit af met een perspectief, dat niet zozeer is gebaseerd op teleurstelling, maar waaruit wel bezorgdheid spreekt over de toekomst van de onderwijskunde. Ik hoop dat mijn verhaal voldoende uitdagend is om tot die acties over te gaan die de toekomst van de onderwijskunde veilig stellen.

INTERDISCIPLINARITEIT

De onderwijsresearch vindt in belangrijke mate plaats aan die universiteiten waar opleidingen gevestigd zijn die het onderwijs tot object van studie hebben en die studenten voorbereiden op een functie in het veld van onderwijs. In veel gevallen gaat het daarbij om een *interdisciplinaire opleiding onderwijskunde*.

Dit type opleidingen heeft het niet gemakkelijk gehad de afgelopen jaren. Er is een tendens waarneembaar die kan worden omschreven als 'back to the basics', terug naar de basisdisciplines.

Mede onder invloed van allerlei beoordelingsprocedures die de laatste jaren in zwang zijn gekomen en waarbij de nadruk veelal ligt op individuele disciplinaire excellentie, is er binnen de universitaire opleidingen de neiging om het disciplinaire onderzoek te versterken, hetgeen ten nadele is van de multi- en interdisciplinaire opleidingen. Deze tendens wordt versterkt door de reacties uit het beroepsveld waarop de opleidingen zich richten. Toen universiteiten voornamelijk werden afgerekend op het aantal studenten dat instroomde, groeide het aantal opleidingen enorm. Er was sprake van een volstrekt uit de hand gelopen differentiatie van opleidingen. Het ging hierbij zonder uitzondering om opleidingen die claimden maatschappelijke vraagstukken te kunnen oplossen. De kennis die hiervoor nodig is, is afkomstig van interdisciplinair onderzoek dat een sterk toegepast karakter heeft. Juist deze bijna vanzelfsprekende koppeling tussen interdisciplinair en toegepast is naar mijn mening de grootste zwakte gebleken van deze

¹ Gecorrigeerde openingstoespraak in het kader van de ORD'96 te Tilburg (4-6 juni 1996).

² Rector Magnificus van de Katholieke Universiteit Brabant.

opleidingen. Het bedrijfsleven heeft duidelijk gemaakt dat universiteiten met de geconstateerde proliferatie van maatschappelijk relevante opleidingen op de verkeerde weg waren omdat zij onvoldoende konden aantonen wat:

1. het vakmanschap is van de nieuwe afgestudeerde;
2. de innovatieve kracht is van de nieuwe tak van wetenschap;
3. het oplossend vermogen is voor de problematiek van het maatschappelijk veld waarop men zich richt.

Ook de minister van onderwijs heeft gereageerd op de roep vanuit het veld om het aantal opleidingen drastisch terug te dringen. In het plan voor hoger onderzoek en onderwijs voor de komende twee jaar (HOOP'96), heeft hij aangegeven hieraan serieus aandacht te willen besteden. Inmiddels heeft de VSNU een werkgroep in het leven geroepen (waarvan ik lid ben) die concrete voorstellen ontwikkelt om het geheel van ruim 300 universitaire opleidingen terug te brengen tot beneden de honderd. De reacties op het eerste voorstel van deze commissie, zowel van de minister, als van de werkgevers en van de universitaire bestuurders, zijn bepaald niet negatief.

Onderwijskunde blijft volgens het voorstel van de commissie als een afzonderlijke opleiding gehandhaafd, ofschoon er ook stemmen waren die pleitten voor samenvoeging met pedagogiek. Deze mededeling dient als een waarschuwing.

Onderwijskunde wordt geacht het brede veld van het onderwijs te bestuderen vanuit verschillende gezichtspunten. Om dit doel te bereiken zijn de opleidingen interdisciplinair opgezet. Kenmerkend voor interdisciplinaire opleidingen is dat *integratie* tot stand wordt gebracht tussen kennis, inzichten en benaderingswijzen uit verschillende disciplines. Deze kennis, inzichten en benaderingswijzen komen voort uit onderzoek waarbij sprake is van intensieve samenwerking tussen verschillende disciplines met koppeling en terugkoppeling tussen de verschillende bijdragen.

Indien de genoemde kennis, inzichten en benaderingswijzen wel een rol spelen, maar integratie geen expliciet doel is, dan spreken we gewoonlijk van multidisciplinariteit. Het verschil tussen beide heb ik onlangs iemand als volgt horen formuleren: "In het geval van multidisciplinariteit is één en één twee en bij interdisciplinariteit is één en één drie."

Hoe dan ook, in beide gevallen doet zich de vraag voor wat de balans is tussen breedte en diepte. Welke kennis en vaardigheden moeten de studenten zich eigen maken? Welk overzicht moeten zij hebben van de verschillende relevante kennisdomeinen en in hoeverre moeten zij vertrouwd zijn met de benaderingswijzen die in die domeinen worden toegepast?

Naar mijn oordeel zullen de kracht en daarmee ook de overlevingskansen van de interdisciplinaire universitaire onderwijskunde sterk afhangen van de mate waarin individuele wetenschappers excelleren in één van de betreffende basisdisciplines. Zij moeten vertrouwd zijn met de voornaamste noties, theorieën, methoden en technieken van die discipline en in staat zijn hoogwaardig wetenschappelijk onderzoek te verrichten dat bijdraagt aan de kennis en inzichten op het betreffende gebied. Maar dit is niet alles. Sterker nog, dat kan een gevaar inhouden. Vaak hebben de excellente onderzoekers iets 'arrogants' dat te maken heeft met de cultuur van het deelgebied waaruit zij afkomstig zijn. We kennen dat in verschillende disciplines: de functie-leerpsycholoog die wat neerkijkt op de klinisch psycholoog; de econometrist die zich verheven voelt boven de bedrijfseconoom, de taalkundige die vindt dat hij serieuzer genomen moet worden dan de letterkundige en de logicus die denkt dat hij een wezenlijker bijdrage levert aan de wijsbegeerte dan een sociaal filosoof. Hier gaat het nog om verschillen binnen disciplines. Om interdisciplinaire wetenschapsbeoefening tot een succes te maken moet er naast de wetenschappelijke diepgang ook de bereidheid en het vermogen zijn om samen te werken. Men moet zich werkelijk aangesproken voelen door de problematiek die moet worden aangepakt, bereid zijn om over de grenzen van de eigen discipline heen te kijken, kunnen communiceren en dus

ook kunnen luisteren en aanvaarden wat anderen hebben te melden, ook als ze een andere en soms wat minder heldere taal gebruiken.

De kwaliteit van universitaire opleidingen wordt van tijd tot tijd tegen het licht gehouden door de visitatiecommissies. Het is in mijn ogen opmerkelijk dat de *onderwijs*visitatiecommissie die op het terrein van de onderwijskunde aan het werk is geweest, de multi- of interdisciplinariteit niet expliciet als toetsingscriterium heeft gehanteerd, zoals dat bijvoorbeeld bij milieukunde-opleidingen wel is gebeurd. Zij heeft het vakgebied simpelweg ingedeeld in een aantal niveaus en voorts een onderscheid gemaakt tussen wel en niet regulier onderwijs. Binnen de aldus ontstane rubrieken heeft zij gekeken hoe de vlag erbij hing. Het oordeel was uitermate mild. De conclusie, zoals geformuleerd in het visitatierapport, luidt: "Alles overziende is de visitatiecommissie van oordeel dat de opleidingen Onderwijskunde aan de Nederlandse universiteiten in het algemeen van redelijk tot goed niveau zijn."

Blijkbaar is de opleiding onderwijskunde in vier jaar goed te doen. Dit in tegenstelling tot de interdisciplinaire opleiding milieukunde, waarvoor een langere studieduur noodzakelijk werd geacht. De afgestudeerden bleken hun titel waard te zijn, omdat zij meer vakken hebben gevolgd dan het programma voorschreef.

INTERNATIONALISERING

In het voorafgaande ging het om de onderwijsvisitatie. We zullen nu stilstaan bij de *onderzoeks*-visitatie.

Laat ik beginnen met een opmerking vooraf. De ervaringen die met het visitatiestelsel tot nu toe zijn opgedaan hebben duidelijk gemaakt dat kwaliteitsbeoordeling in hoge mate afhankelijk is van de disciplinaire context waarbinnen beoordeeld wordt. Voor juridisch onderzoek zullen andere normen en regels moeten worden gehanteerd dan voor psychologisch onderzoek. Hier doet zich al direct een probleem voor bij interdisciplinair onderzoek. Bij welke discipline moet bijvoorbeeld onderwijskundig onderzoek worden ondergebracht? Vanuit de ene discipline kan de waardering voor een bepaald type onderzoek een geheel andere zijn dan vanuit de andere.

Toen professor Frijda voor de eerste keer met pensioen ging werd hem gevraagd welk wetenschapsgebied hij onder de maat vond. Dat vond hij een lastige vraag omdat daarop meerdere antwoorden mogelijk zijn. Hij wilde er slechts één noemen: onderwijskunde.

De onderwijskunde heeft er verstandig aan gedaan om buiten de invloedssfeer van de onderzoeksvisitatie psychologie te blijven die hij na zijn pensionering heeft gepresideerd. Zoals u allen ongetwijfeld weet is onderwijskunde gevisiteerd samen met de pedagogiek.

Sinds de invoering van het visitatiestelsel is er meer nadruk komen te liggen op de internationalisering van het onderzoek. Nederlandse publikaties tellen nauwelijks meer mee. Wellicht is dit de voornaamste reden dat de jaarlijkse bundels die naar aanleiding van de voordrachten tijdens de onderwijsresearchdagen werden uitgebracht niet langer meer verschijnen. Ze lonen de moeite niet meer. Ook het aantal engelstalige dissertaties begint flink toe te nemen. Misschien is het aardig te vermelden wat mij enkele jaren geleden is overkomen als lid van de jury voor de VOR-prijs. De door mij geselecteerde dissertatie werd door de overige juryleden niet gehonoreerd. Eén van de juryleden noemde toen als reden dat het boek niet in het Nederlands was gepubliceerd.

Voor zover het om research gaat is er altijd een zekere spanning tussen fundamenteel en toegepast of beleidsgericht onderzoek. In het eerste geval gaat het gewoonlijk om 'science driven' onderzoek dat gericht is op vermeerdering van kennis. De wetenschappelijke kwaliteit staat centraal. In het tweede geval staan maatschappelijke problemen voorop. Door middel van onderzoek wordt getracht deze op te lossen. Niet het wetenschappelijk, maar het maatschappelijk belang staat centraal.

Gesteld moet worden dat de grens tussen beide soorten onderzoek moeilijk te trekken is. Bovendien kan onderzoek zowel theoretisch als maatschappelijk relevant zijn en voor onder-

wijskundig onderzoek is dat ook vaak het geval. Ofschoon onderwijskunde in belangrijke mate toepassingsgericht is geweest, wordt, mede onder invloed van het visitatiestelsel, het fundamentele onderzoek de laatste jaren benadrukt. Dit komt omdat het stelsel meer is opgezet vanuit de basisdisciplines dan vanuit interdisciplinaire opleidingen. Zo was er voor onderwijskunde geen afzonderlijke visitatiecommissie en is de kwaliteit beoordeeld door de commissie die ook de pedagogiek heeft moeten beoordelen.

Binnen de basisdisciplines is er in het algemeen een hoge waardering voor internationale, hoogwaardige publikaties, waarbij de fundamentele bijdragen zwaarder wegen dan de praktische. Door de met de visitatie samenhangende beloningen (in termen van punten) is er een zekere druk om meer fundamentele problemen aan te pakken.

De vraag is of we daarin zijn geslaagd. Uit het visitatierapport valt af te leiden dat het antwoord op deze vraag is: "neen". Geen van de door de onderzoekscommissie bekeken programma's zijn echt uitstekend als je kijkt naar de impact die ze hebben binnen de internationale 'scene'. Waar we goed in zijn is het gebruiken en toepassen van door buitenlanders bedachte theorieën of benaderingen. Voorzover wij de resultaten van onze onderzoeksinspanningen publiceren deden we dat veel te weinig in het Engels en zelden in vooraanstaande internationale tijdschriften.

ONDERZOEKSCHOLEN

Gedurende de afgelopen veertien jaar zijn er tal van *onderzoekscholen* ontstaan. Enerzijds dienen deze om de aio-opleiding te verzorgen en anderzijds om de kwaliteit van de onderzoeksprogramma's te verbeteren.

In feite zijn er twee soorten scholen, facultaire en federatieve scholen. De keuze is niet geheel vrij, maar hangt in belangrijke mate af van de grootte van de faculteit. Als er op een bepaald terrein te weinig onderzoekscapaciteit is binnen een faculteit dan moet men samenwerken met collegae uit andere universiteiten. In een aantal gevallen is door de KNAW nadrukkelijk bepleit te streven naar landelijke dekkendheid in de vorm van federatieve werkverbanden.

Een voorbeeld van een *federatieve onderzoekschool* is het ICO (Interuniversitair Centrum voor Onderwijskundig Onderzoek). Hierin is onderzoek ondergebracht van onderwijswetenschappelijke opleidingen van verschillende universiteiten.

Zoals gezegd is de onderzoekschool het opleidingsinstituut voor aio's. Destijds is afgesproken dat een school minstens 40 aio's zou moeten tellen. De kwaliteit van de opleiding blijkt onder andere uit de aio-rendementen. Die zijn een goede graadmeter voor het functioneren van onderzoekscholen.

De school is tevens een onderzoeksinstituut. Het gaat hier niet alleen om het proefschriftonderzoek van de aio's, maar ook om het kwalitatief betere onderzoek van de staf dat in de school is ondergebracht. Ook hier speelt de output een belangrijke rol. Het ICO-bestuur is om deze reden begonnen met het verzamelen van kengetallen, waaronder de aio-rendementen en de wetenschappelijke output per fte.

Bezien we de resultaten van de onderzoeksvisitatie dan komen we in Nederland voor de onderwijswetenschappen gemiddeld uit op een produktie van 1.2 publikaties per fte per jaar, met een variatiebreedte van .25 tot 3.5. Dit betekent dat er heel wat wetenschappers rond lopen die zeer weinig presteren. Het door het ICO-bestuur geïnitieerde rendementsonderzoek is erop gericht aan te tonen dat de produktiviteit van degenen die werkzaam zijn binnen de school aanmerkelijk hoger ligt dan het landelijk gemiddelde, zowel in kwantitatief als in kwalitatief opzicht. Van de uitslag van dit onderzoek zal het mede moeten afhangen of met de school moet worden doorgegaan indien de termijn van vijf jaar zal zijn verstreken.

Een onderzoekschool behoort meer te zijn dan een verzameling individuele projecten. De projecten behoren deel uit te maken van een programma. In een onderzoekschool worden een aantal

programma's ondergebracht die onderling met elkaar samenhangen. Op deze manier wordt voorkomen dat het onderzoek versnipperd over het land wordt uitgevoerd.

Als ICO-bestuurder pleit ik voor een onderscheid tussen consistentie *binnen* de programma's en samenhang en coördinatie *tussen* de programma's van de onderzoeksschool.

Over de interne consistentie heeft de genoemde onderzoeksvisitatiecommissie al de nodige pittige opmerkingen gemaakt. Er werd opgemerkt dat verscheidene onderzoeksprogramma's niet alleen internationale excellentie missen maar bovendien geen interne samenhang vertonen. Zulke programma's fungeren in de ogen van de commissie als handige paraplu's voor uiteenlopende en afzonderlijke onderzoekslijnen. Het ICO is bedoeld om hierin verandering aan te brengen.

Binnen het verband van de onderzoeksschool kan de consistentie worden bevorderd door regelmatig seminars en workshops te organiseren. In die zin zijn de onderwijsresearchdagen bijzonder nuttig. Eigenlijk zouden hier alleen onderzoeksprogramma's gepresenteerd moeten worden, zodat ook de rode draad en de onderlinge samenhang ter discussie kunnen staan. Ook verdient het aanbeveling tegen het eind van de planperiode een overzichtsartikel te produceren, waarin onder meer de produktiviteit, de consistentie en de relaties tussen de verschillende programma's tot uitdrukking worden gebracht. Dit zou kunnen dienen als een basis voor een eventuele vervolgvraag aan het einde van de looptijd van het huidige onderzoekprogramma. Over de coördinatie tussen programma's is, althans voor wat betreft ICO, nog weinig bekend. Ook dit heeft de aandacht van het ICO-bestuur.

De coördinatie tussen programma's kan worden bevorderd via een regelmatig programmaleidersoverleg. Dit zou onder andere moeten dienen om een gezamenlijke tekst voor het jaarverslag te produceren, waarin speciaal aandacht wordt besteed aan samenhang en afstemming. In de huidige jaarverslagen ontbreekt een paragraaf over dit onderwerp.

PERSPECTIEF

Dames en heren; in het voorafgaande heb ik me hier en daar wat kritisch en uitdagend uitgelaten over de interdisciplinaire onderwijskunde. Dit is vooral een gevolg van mijn bezorgdheid. Degenen die verantwoordelijk zijn voor de inrichting van de universitaire opleidingen onderwijskunde dienen zich voortdurend af te vragen hoe de opleiding eruit moet zien en in het bijzonder welke keuze moet worden gemaakt ten aanzien van wat ik als het hoofd-dilemma zie, namelijk *breedte versus diepte*. In feite gaat het om de vraag wat de relatie is ten aanzien van de basisdisciplines. Naarmate meer wordt gekozen voor de breedte-optie zal de onderwijskunde ook meer kwetsbaar zijn. Als een afgestudeerde aan een universitaire interdisciplinaire opleiding onderwijskunde geen psycholoog is, en geen socioloog en geen pedagoog, om maar enkele van de basisdisciplines te noemen, dan is hij eigenlijk helemaal niets. Deze stelling is des te meer waar naarmate de mix van disciplines, waarmee de studenten tijdens hun opleiding onderwijskunde worden geconfronteerd, groter is.

De status van het vakgebied wordt voor een belangrijk deel bepaald door de mate waarin de beroepsbeoefenaren erin slagen om de majeure problemen die zich in de praktijk voordoen tot een oplossing te brengen. In een groot aantal gevallen kan dit worden gerealiseerd vanuit een mono-discipline. Vraagstukken die te maken hebben met individuele verschillen in leervermogen en daarmee verband houdende capaciteiten kunnen heel goed door psychologen worden aangepakt. Dat geldt ook voor het vraagstuk van selectie en van het vinden van remedies als zich problemen voordoen, bijvoorbeeld van emotionele of motivationele aard.

Pas als een probleem niet door een monodisciplinaire aanpak kan worden opgelost kan de moeilijke weg van de samenwerking tussen disciplines worden ingeslagen. Deze leidt alleen dan tot succes als aan een aantal basisvoorwaarden wordt voldaan:

1. er moet een gemeenschappelijke doelstelling zijn voor de verschillende teamleden;
2. de teamleden moeten zich aangesproken voelen door dezelfde problematiek, zij het dat dit om verschillende motieven het geval kan zijn. De een zal meer vanuit een theoretische belangstelling participeren en de ander meer vanuit een maatschappelijke betrokkenheid.

3. elke individueel teamlid moet over goede disciplinaire kennis en vaardigheden beschikken zodat de verschillende teamleden qualitate qua gelijkwaardig zijn.

De vraag is of de onderwijskunde op den duur als een initiële opleiding moet blijven bestaan. Verdient het niet de voorkeur om onderwijspsychologen, -sociologen en -pedagogen af te leveren, die een ferme disciplinaire basis hebben en die tijdens het afstudeerjaar van hun opleiding kunnen kiezen voor een afstudeerrichting gericht op die onderwijsvraagstukken die in de context van de eigen discipline relevant zijn. In deze fase zou ook aandacht besteed moeten worden aan vaardigheden die te maken hebben met samenwerken. Het onderwijs tijdens het afstudeerjaar zou in belangrijke mate aangestuurd kunnen worden door landelijke onderzoekscholen. Daarin participeren geselecteerde onderzoekers die - in het ideale geval - zowel bezig zijn met disciplinaire verdieping als verbreding.

Dames en heren, hoe dan ook, naar mijn volle overtuiging zal het voortbestaan van de onderwijswetenschappen aan de Nederlandse universiteiten afhankelijk zijn van de mate waarin zij erin zullen slagen:

1. bij te dragen aan de *vergroting van het kennisdomein* dat door de onderscheiden onderwijswetenschappen wereldwijd wordt gedekt;
2. de verworven kennis en inzichten te *toetsen op bruikbaarheid* ten behoeve van de beroepspraktijk, mede met het doel:
3. bij te dragen aan de verdere *ontwikkeling van de samenleving* als geheel.

We beleven met de onderwijskunde cruciale tijden. We hebben enkele weken geleden kunnen constateren dat vele miljarden zijn verspild aan het wegwerken van achterstanden die vele leerlingen aan het begin van de schoolloopbaan bleken te hebben opgelopen. In de zeventiger jaren werd gestart met een onderwijsstimuleringsbeleid voor kinderen van arbeiders. In de tachtiger jaren is er een cultureel minderhedenbeleid gevoerd voor de kinderen van gastarbeiders en enkele jaren geleden is er opnieuw veel geld besteed aan het achterstandsbeleid. Mevrouw Mulder, verbonden aan het ITS, heeft aangetoond dat dit beleid volledig is mislukt. De schuld werd voortsnog niet gezocht bij de onderwijskundigen maar bij overheid en scholen.

Momenteel doet zich echter een nieuwe ontwikkeling voor waarbij de wetenschappers niet langer buiten schot zullen blijven. Hun onderzoek naar zelfstandig leren heeft de overheid gestimuleerd tot het vernieuwen van de tweede fase van het voortgezet onderwijs. Het studiehuis is geïntroduceerd. De komende jaren zullen leren hoe serieus de onderwijswetenschappen moeten worden genomen. Het gaat hier dan precies om de drie zaken die ik zojuist formuleerde:

1. de *kennis*; is de kennisbasis voldoende stevig om het te bouwen studiehuis te kunnen dragen?
2. de *toetsing*; is de kennis getoetst op bruikbaarheid voor de professionals die het studiehuis moeten bouwen, inrichten en onderhouden?
3. de *ontwikkeling van de samenleving*; zal het gemiddeld niveau van degenen die het voortgezet onderwijs verlaten de vergelijking met andere landen kunnen doorstaan?

Ik hoop dat er tijdens deze onderwijsresearchdagen voldoende gelegenheid is voor een dergelijke discussie. Eigenlijk zouden alleen voorstellen besproken moeten worden in het licht van deze drie criteria. Uiteraard gaat het dan niet om individuele projecten, maar om programma's. Die bepalen de impact en daarmee de toekomst van de onderwijswetenschappen.

Ik wens u een vruchtbaar congres toe!

Notities en Commentaren

CONSTRUCTIEVE EN DESTRUCTIEVE DISCUSSIES OVER DE INTERDISCIPLINAIRE ONDERWIJSKUNDE. EEN REACTIE OP DE KLERK

B. Creemers¹, N. Lagerweij² en N. Verloop³

Het zal niet vaak voorkomen dat de deelnemers aan een congres welkom worden geheten door de rector van een universiteit die daarbij en passant het bestaansrecht van het wetenschapsgebied van de deelnemers ter discussie stelt. Dit gebeurde de deelnemers aan de ORD '96 bij de opening van hun conferentie in Tilburg, waarbij rector (en onderwijspsycholoog) De Klerk een somber beeld schetste van de stand van zaken en toekomst van de onderwijskunde. De dagbladen maakten gretig gebruik van het persbericht dat tegelijkertijd, met de toevoeging "rector magnificus van de Katholieke Universiteit Brabant", verspreid was. "Onderwijskunde kan opgeheven" kopte de NRC de volgende dag en de andere dagbladen lieten zich in soortgelijke bewoordingen uit.

Wat was de aanleiding tot deze commotie? Was er een onoplosbare crisis in de onderwijskunde ontstaan? Was er een vernietigend visitatie-rapport verschenen of bevatte het zojuist verschenen rapport van de Verkenningcommissie Onderwijsonderzoek verontrustende bevindingen? Niets van dat alles blijkt het geval te zijn. Het blijkt te gaan om een mening van De Klerk, met name wat betreft het belang van monodisciplinair versus interdisciplinair onderzoek. Wij gunnen ieder zijn mening over een dergelijk vraagstuk. Hoewel dit debat enkele decennia geleden in de onderwijswetenschappen al langdurig en grondig is gevoerd, is het denkbaar dat hierover opnieuw een discussie wordt geopend. Een symposium tijdens de ORD zou daarvoor bijvoorbeeld een uitgelezen gelegenheid zijn geweest. In plaats van een discussie in het wetenschappelijk forum is hier gekozen voor het (via een persbericht) naar buiten brengen van een particuliere mening.

We zullen hieronder ingaan op enkele punten uit het betoog van De Klerk. Daarbij baseren we ons op de tekst die tijdens de ORD is uitgesproken en onder de pers is verspreid en niet op de tekst die later voor dit tijdschrift ter beschikking is gesteld. We concentreren ons hier op het thema 'interdisciplinariteit', omdat daaraan de meest zware twijfels aan het bestaansrecht van de onderwijskunde worden ontleend en omdat dit punt ook in de pers de meeste aandacht heeft gekregen.

De Klerk pleit, kort gezegd, voor een sterkere monodisciplinaire orientatie, ten koste van de momenteel dominante inter- of multidisciplinaire orientatie van de onderwijskunde. Hij stelt onder meer "Naar mijn oordeel zal de kracht en daarmee ook de overlevingskansen van de interdisciplinaire onderwijskunde sterk afhangen van de mate waarin individuele wetenschappers excelleren in een van de betreffende basisdisciplines". En verder: "Als een afgestudeerde aan een universitaire interdisciplinaire opleiding onderwijskunde geen psycholoog is, en geen socioloog, en geen pedagoog, om maar enkele basisdisciplines te noemen, dan is hij eigenlijk helemaal niets". Zo, daar kan de onderwijskundige het mee doen.

We zullen hieronder een drietal argumenten naar voren brengen die onzes inziens aantonen dat een dergelijke opstelling onlogisch en niet ter zake is. Zij hebben achtereenvolgens betrekking op de aard van het probleemgebied, de huidige werkwijze van de onderwijswetenschappers en de kwaliteit van de output.

¹(RUG), directeur onderwijs onderzoekschool ICO

²(UU), voorzitter Kamer Onderwijskunde VSNU

³(RUL), voorzitter VOR

Ten eerste de *aard van het probleemgebied* waarop de onderwijskunde zich richt, het onderwijs als maatschappelijk relevant probleemveld. Omdat het destijds gaandeweg duidelijk werd dat de afzonderlijke bestaande wetenschapsdisciplines onvoldoende in staat waren de complexe vragen vanuit dit probleemveld te 'bestrijken', is, met vallen en opstaan, de interdisciplinaire onderwijskunde ontwikkeld. Dat men in de onderwijskunde zo goed mogelijk gebruik moet maken van de inzichten en methoden uit de basisdisciplines is niet ter discussie. Deze basisdisciplines vormen dan ook de hoofdmoot van de eerste jaren van de onderwijskunde-opleidingen. Het is echter curieus om te stellen dat de kwaliteit van de onderwijskunde als interdiscipline een functie is van deskundigheid in een van de basisdisciplines. Kenmerkend voor het maatschappelijk probleemveld is nu juist dat de meest essentiële vragen zich niet laten vertalen in vraagstellingen die met een van de basisdisciplines opgelost kunnen worden. Het ligt dus in het geheel niet voor de hand dat grotere deskundigheid in de basisdisciplines op zich bepalend zou zijn voor de kwaliteit van de onderwijskundige. Veel meer voor de hand liggend is, dat het vermogen tot integreren van uiteenlopende deskundigheden en inzichten ten behoeve van de oplossing van een voor het onderwijs relevant probleem, de specifieke deskundigheid van de onderwijswetenschapper zal uitmaken. Daarbij moet zeker ook vermeld worden dat de onderwijskunde de afgelopen decennia een aantal eigen aanpakken en theorieën heeft ontwikkeld waarin de 'basisdisciplines' als zodanig nauwelijks meer een rol spelen.

Kwaliteit van de onderwijskundige definiëren als een functie van zijn deskundigheid als psycholoog of pedagoog is even onzinnig als het definieëren van de kwaliteit van de werktuigbouwkundige in termen van zijn kennis op het terrein van de natuurkunde of de mechanica. Natuurlijk kan de werktuigbouwkundige niet zonder kennis van deze basisdisciplines, maar zijn specifieke deskundigheid ligt nu juist in het combineren van deze inzichten ter oplossing van praktische problemen.

Het tweede argument heeft betrekking op de *huidige werkwijze van de onderwijswetenschappers*. Geconstateerd kan worden dat op dit moment tal van onderwijswetenschappers met tal van achtergronden (inclusief de door De Klerk bepleite monodisciplinaire achtergrond) met elkaar samenwerken aan de oplossing van onderwijskundige vraagstellingen. Geconstateerd kan bovendien worden dat men de afgelopen jaren een werkwijze heeft ontwikkeld waarbij de participanten steeds meer gebruik maken van elkaars inzichten en aanpakken. Kenmerkend voor de huidige stand van zaken is dat men van collega's vaak niet eens meer weet wat zijn of haar disciplinaire achtergrond is. Men kent vaak de aard (en de kwaliteit!) van iemands wetenschappelijke prestaties, zonder te weten of iemand nu 'eigenlijk' van oorsprong pedagoog, psycholoog of socioloog is. Men kan dit betreuren. Wij zijn in Nederland experts in het voeren van 'bloedgroepen'-discussies. Enkele decennia geleden werd deze discussie ook op het terrein van de onderwijswetenschappen gevoerd. Dat we in de onderwijswetenschappen dit stadium te boven zijn beschouwen we als winst. In elk geval is er op grond van datgene wat 'op de werkvloer' gebeurt geen enkele behoefte aan het oprakelen van deze discussies. Dit geldt overigens ook op grond van de internationale stand van zaken. Op het internationale gebied van de educational research wordt het volstrekt oninteressant gevonden of iemand nu iets als pedagoog, als psycholoog of als wat dan ook te berde brengt. Het gaat dan uitsluitend om de kwaliteit van het onderzoek of de theorievorming die men op tafel legt.

Overigens blijkt dat de Nederlandse onderwijskundigen hierin een zeer substantiële inbreng hebben. Iedereen die internationale congressen bezoekt en internationale tijdschriften leest kan constateren dat de Nederlandse onderwijskunde (samen met die uit Israël en uit enkele Scandinavische landen) verhoudingsgewijs niet onderdoet voor die uit het Angelsaksische taalgebied. Dit brengt ons op het volgende argument.

Het derde argument betreft de *kwaliteit van de output*. Over de kwaliteit van de output van de onderwijskunde worden door De Kerk nogal wat persoonlijke meningen naar voren gebracht. Over dergelijke punten kan gemakkelijk een welles-nietes discussie ontstaan. We verkeren in de gelukkige omstandigheid dat er recentelijk een drietal grondige analyses van de kwaliteit van de onderwijskunde heeft plaatsgevonden, namelijk een onderwijsvisitatie, een onderzoeksvisitatie en een zogenaamde 'Verkenning' van het Nederlandse onderwijsonderzoek (in opdracht van de

door de minister ingestelde Overlegcommissie Verkenningen). Zonder nu te willen stellen dat al deze analyses vlekkeloos zijn verlopen, mag toch, gezien de grote hoeveelheid energie die hierin door tal van gerenommeerde wetenschappers is gestoken, verwacht worden dat deze bevindingen op hoofdlijnen juist zijn. Uiteraard kunnen we hier slechts de algemene conclusies uit deze rapporten memoreren. De eindconclusie van de onderwijsvisitatiecommissie is dat "de opleidingen Onderwijskunde aan de Nederlandse universiteiten in het algemeen van redelijk tot goed niveau zijn". De onderzoeksvisitatiecommissie concludeert met betrekking tot de 'educational sciences' (d.w.z. de Onderwijskunde): "With the exception of one small programme, the quality of all programmes is rated as 'satisfactory', 'good' or 'very good'". Ook het in mei van dit jaar verschenen rapport van de 'Verkenningscommissie Onderwijsonderzoek', ten slotte, geeft geen enkele aanleiding tot het doen van verontruste uitspraken over het Nederlandse onderwijsonderzoek. In dit rapport zijn, in het kader van onze discussie, met name de uitspraken over de oriëntatie op de basisdisciplines interessant. Zo wordt op pag. 27 bijvoorbeeld gesteld: "De indruk van de commissie is evenwel dat kennis en inzichten uit de basisdisciplines in voldoende mate doordringen tot het onderwijsonderzoek" (zie verder ook pag. 30).

We kunnen hier uiteraard niet op allerei details van de drie rapporten ingaan. De algemene tendens is echter volstrekt duidelijk: De kwaliteit van de Nederlandse onderwijskunde is redelijk tot goed, zowel op het terrein van het onderwijs als van het onderzoek.

Uit het bovenstaande blijkt dat er op zich geen enkele concrete aanleiding is om de kwaliteit van de Nederlandse onderwijskunde serieus ter discussie te stellen, hetgeen overigens natuurlijk niet betekent dat er op onderdelen geen verbeteringen noodzakelijk zouden zijn.

Zoals gezegd, wij gunnen ieder zijn mening omtrent het vraagstuk van de interdisciplinariteit. Wij zijn echter van mening dat De Klerk met het op deze wijze naar buiten brengen van zijn opvatting de onderwijskunde onterecht en onnodig schade heeft gedaan. Dit wordt nog versterkt doordat hij in zijn betoog de onderwijskunde op een hoop veegt met de tientallen modieuze studies die de laatste jaren zijn ontstaan en die als voornaamste doel hadden meer studenten binnen te halen. Het is juist dat de VSNU een commissie (waarvan De Klerk lid is) opdracht heeft gegeven tot het doen van voorstellen voor reductie. De onderwijskunde bestaat al zo'n 25 jaar en heeft met deze proliferatie van modieuze studierichtingen niets van doen.

Wanneer De Klerk behoefte heeft aan een debat over de interdisciplinariteit van de onderwijskunde staan daartoe de geëigende wegen open (en hij is bij dezen uitgenodigd voor een discussie hierover tijdens de volgende ORD).

Wie echter een standpunt hieromtrent koppelt aan het in twijfel trekken van de kwaliteit of zelfs het bestaansrecht van de onderwijskunde, dient hiervoor gedegen en controleerbare argumenten te hebben. Deze ontbreken.

Door de wijze waarop de zaak nu in de publiciteit is gebracht is stemmingmakerij ontstaan waartegen de onderwijskunde zich per definitie niet met rationele argumenten kan verweren. Dat laatste is ongetwijfeld niet de bedoeling van collega De Klerk geweest. Hij had, zeker als rector magnificus, beter moeten weten.

Mededeling van de uitgever

Dit is het laatste nummer van het Tijdschrift voor Onderwijsresearch dat wordt uitgegeven door Swets & Zeitlinger. Auteurs en abonnees kunnen voortaan hun correspondentie richten aan:

Dr Martin Mulder
Vereniging voor Onderwijsresearch
Faculteit Toegepaste Onderwijskunde
Universiteit Twente
7500 AE Enschede
Tel.: 053 489 3652
Fax: 053 489 3759
E-mail: mulder@edte.utwente.nl

Inleiding op het thema leerstijlen

Jan Vermunt¹ en Johan van der Sanden²

¹ICLON, Rijksuniversiteit Leiden, ²Katholieke Universiteit Brabant en Technische Universiteit Eindhoven

Het leerproces van leerlingen en studenten heeft de laatste jaren een steeds centraler plaats gekregen in het denken over de kwaliteit van onderwijs. Waar in het verleden onderwijsontwerp-theorieën, didactische modellen en handelingsvoorschriften voor docenten vooral gebaseerd waren op het kennisoverdracht-model, vinden veel hedendaagse ontwerptheorieën en dergelijke hun inspiratie in het kennisconstructie-model. Eén reden voor deze verschuiving is van epistemologische aard. Onderzoeksresultaten wijzen erop dat de kwaliteit van kennis verworven door actieve kennisconstructie door lerenden, beter is (toegankelijker, samenhangender, bruikbaar) dan die verworven door passieve kennisopname. Een andere reden is van maatschappelijke aard. De snelle veranderingen in beroep, technologie en maatschappij maken het noodzakelijker dan voorheen dat mensen in staat zijn ook na hun schoolloopbaan zelf nieuwe kennis te blijven verwerven. Het ligt voor de hand dat ze de vaardigheden en inzichten die hiervoor nodig zijn op school aanleren.

Deze toegenomen aandacht voor het leerproces van leerlingen en studenten heeft zich ook gemanifesteerd op het terrein van de onderwijsresearch. Zo is bijvoorbeeld het verloop van concrete leerprocessen bij het bestuderen van studieteksten of het oplossen van opgaven nauwkeurig onderzocht met behulp van hardop-denken protocollen en computerregistraties in experimentele onderzoeksopzetten (bijv. De Jong, 1992). Andere onderzoeken richtten zich meer op het leergedrag van leerlingen en studenten in de context van hun eigen school, studie of opleiding (bijv. Lonka en Lindblom-Ylänne, 1996). In beide gevallen is een belangrijk uitgangspunt voor het initiëren van dergelijk onderzoek dat meer kennis over de wijze waarop leerlingen en studenten leren, een noodzakelijke voorwaarde is om onderwijsprogramma's te kunnen maken waarin actieve kennisconstructie door lerenden centraal staat.

De 'wijze waarop leerlingen en studenten leren' kan op verschillende manieren worden geoperationaliseerd. Het meest concrete niveau is dat van de *uitwendig waarneembare leeractiviteiten*. Hieronder valt bijvoorbeeld leergedrag als onderstrepingen of markeringen in studieteksten aanbrengen, college-aantekeningen maken, of discussiëren met medestudenten. Wat minder concreet, maar vanuit psychologisch gezichtspunt zeker zo interessant, is het niveau van de *inwendige, mentale leeractiviteiten*. Dit betreft bijvoorbeeld activiteiten als nadenken over wat het studieboek en de uitleg door de docent met elkaar te maken hebben, zichzelf vragen stellen om te toetsen of men de stof beheerst, en het toeschrijven van goede cijfers aan eigen capaciteiten. Een uitwendig waarneembare activiteit als onderstrepen van tekstgedeelten kan in functie staan van verschillende mentale activiteiten, bijvoorbeeld het reduceren van de hoeveelheid stof die uit het hoofd gekend moet worden, of het markeren van die uitspraken in het boek waarmee men het niet eens is. Een derde niveau van operationalisatie van de 'wijze waarop leerlingen en studenten leren' is dat van de *leerstrategieën*. Dit betreft combinaties van leeractiviteiten die door lerenden vaak in samenhang worden gebruikt om een bepaald doel te bereiken. Zo wordt de leerstrategie 'diepte-verwerking' gekenmerkt doordat lerenden actief zoeken naar samenhangen in de leerstof, proberen daar structuur in aan te brengen en er overzicht over te krijgen, en kritisch over de aangeboden stof nadenken, met het doel tot een zo goed mogelijk inzicht in die stof te komen. Wanneer een leerling een leeractiviteit of -strategie goed kan uitvoeren, heeft deze leerling *vaardigheid* in het gebruiken van die activiteit of strategie. Een vierde niveau van operationalisatie is dat van *leerstijl*. Dit begrip wordt gehanteerd om consistentie in leergedrag mee aan te duiden. Wanneer leerlingen en studenten gewoonlijk gebruik maken van dezelfde leerstrategieën vertonen ze een bepaalde leerstijl. Evenals verwante be-

grippen als levensstijl en bouwstijl is leerstijl een prototypisch begrip waarvan een leerling of student veel tot weinig kenmerken kan vertonen. Een levensstijl betekent niet dat iemand zich elke dag precies hetzelfde gedraagt. Evenzo betekent leerstijl ook niet dat iemand in elke situatie op precies dezelfde wijze te werk gaat bij het leren. Wel betekent het dat, in de onderwijssituaties waarin iemand doorgaans vertoeft, er consistentie te ontdekken is in dat leergedrag: wat een leerling vandaag doet in de klas lijkt op wat hij of zij gisteren deed in die klas.

Een leerstijl representeert dus de leergewoonten en gedragsneigingen die de persoon meebrengt in een leersituatie. In de smalle betekenis van de term leerstijl valt alleen het leergedrag onder dit begrip: de leeractiviteiten die iemand ontplooit, de leerstrategieën die hij of zij gebruikt. In de brede betekenis vallen ook de kennis en opvattingen die mensen hebben over leren en onderwijzen (mentale leermodellen) en de persoonlijke doelen en leermotieven (leeroriëntaties) onder dit begrip. Een leerstijl representeert dan een samenhangend geheel van leerstrategieën die mensen gewoonlijk gebruiken, hun mentale leermodellen en hun leeroriëntaties, een geheel dat kenmerkend is voor hen in een bepaalde periode. Vermunt (1992) vond bij studenten in het hoger onderwijs vier van dergelijke leerstijlen: een ongerichte, een reproductiegerichte, een betekenisgerichte en een toepassingsgerichte. Hoewel leerstijlen dus een bepaalde consistentie impliceren in leergedrag over de tijd en over situaties heen, betekent dit niet dat ze onveranderbaar zijn. Ze worden hier niet opgevat als persoonlijkheidstrekken, maar als het resultaat van de interactie tussen persoonsgebonden en omgevingsgebonden invloeden. Veranderingen in deze persoons- en omgevingsgebonden invloeden kunnen dus op den duur ook veranderingen in leerstijlen bewerkstelligen.

Vanuit epistemologisch perspectief is het van belang dat onderwijs zich richt op het bevorderen van leerprocessen gekenmerkt door actieve kennisconstructie door de lerende. Vanuit maatschappelijk oogpunt is het van belang dat de school er zorg voor draagt dat een dergelijke manier van leren ook tot stijl van de lerende wordt. Op deze wijze worden leerlingen en studenten erop voorbereid ook na hun schoolloopbaan zelf actief nieuwe kennis te blijven verwerven. Vanuit theoretisch gezichtspunt is het van belang dat meer kennis beschikbaar komt over een aantal tot nu toe niet opgehelderde kwesties, zoals de ontwikkeling van leerstijlen van de kleuterleeftijd tot in de volwassenheid, de compatibiliteit van verschillende soorten leerstrategieën, de relatie tussen leerstijlen en persoonsgebonden variabelen (zoals sekse, identiteit, persoonlijkheid), omgevingsgebonden variabelen (zoals gehanteerd onderwijsconcept, manier van toetsing, vakgebied) en intermediaire variabelen (zoals perceptie van de leeromgeving), de interne structuur van leerstijlen in verschillende ontwikkelingsfasen van leerlingen, en de relatie tussen leerstijlen en actueel leergedrag in bepaalde leersituaties.

Dit themanummer is bedoeld om enkele van deze theoretische kwesties nader uit te diepen en de praktische consequenties ervan te doordenken. Het bevat vier bijdragen van onderzoekers die, elk vanuit een andere invalshoek, het verschijnsel 'leerstijl' in hun onderzoek hebben betrokken. Gemeenschappelijk is dat alle bijdragen uitgaan van dezelfde leerstijl-conceptualisatie zoals hierboven geschetst. Ook gemeenschappelijk is dat ze als een van hun meetinstrumenten gebruik maken van hetzelfde diagnostisch leerstijlinstrument, de *Inventaris Leerstijlen*. Hierdoor zijn de resultaten van de verschillende studies vergelijkbaar op het thema leerstijlen. Doel van het themanummer is dan ook het concept leerstijl in te bedden temidden van verscheidene andere variabelen en de aard van leerstijlen bij verschillende populaties te verhelderen. De diverse studies verschillen in vrijwel alle andere aspecten. Ze verschillen wat betreft populatie: studenten in het hoger onderwijs en volwassen en adolescente leerlingen in het voortgezet onderwijs. Ze verschillen wat betreft de andere variabelen die in het onderzoek zijn betrokken: sekse en sekse-identiteit, perceptie van de leeromgeving, leerprestaties, selectie- en zoekgedrag bij een studietekst, zelfstandig leren, en de ontwikkeling van leerstijlen. De studies verschillen ook wat betreft primair doel: ontwikkeling van conceptuele modellen, theorievorming, inzicht in leerprocessen, en bijdragen aan de inrichting van het studiehuis.

Severiens en Ten Dam hebben verschillen in leerstijlen in relatie tot sekse en sekse-identiteit onderzocht bij volwassen leerlingen in het voortgezet onderwijs. Ook vergeleken zij de aard van de leerstijlen bij hun populatie met die van studenten in het hoger onderwijs. Zij wijzen

onder meer op het belang van de conceptualisatie van sekse-identiteit in vergelijking met biologische sekse.

Wierstra en Beerends stellen de relatie tussen leerstijl en leeromgeving centraal. Zij presenteren een conceptueel model waarin perceptie van de leeromgeving een centrale plaats inneemt. In hun empirisch onderzoek bij studenten sociale wetenschappen onderzochten zij relaties binnen dit model. Op basis hiervan komen ze tot een typologie van leeromgevingen, leerstijlen en interacties tussen deze variabelen.

Stoutjesdijk en Beishuizen bestudeerden relaties tussen leerstijlen en selectie- en zoekgedrag bij een studietekst aangeboden in een hypertext leeromgeving bij eerstejaars Psychologie-studenten. Met name onderzochten zij of de mate van concordantie tussen verwerkings- en regulatiestrategieën effect had op de prestaties in een dergelijke taak.

Roosendaal en Vermunt onderzochten de structuur van leerstijlen bij leerlingen in de bovenbouw van het voortgezet onderwijs. Vanuit een vergelijking van deze structuur met die van brugklassers en van studenten in het hoger onderwijs schetsen ze een beeld van de ontwikkeling van leerstijlen gedurende de studieloopbaan.

Als redacteurs van dit themanummer zijn we ervan overtuigd dat de artikelen, ieder afzonderlijk maar vooral ook in hun onderlinge samenhang, bijdragen aan een rijkere kennis van het verschijnsel 'leerstijl' vanuit verscheidene perspectieven.

LITERATUUR

- De Jong, F.P.C.M. (1992). *Zelfstandig leren. Regulatie van het leerproces en leren reguleren: een procesbenadering*. Academisch proefschrift, Katholieke Universiteit Brabant, Tilburg.
- Lonka, K., & Lindblom-Ylänne, S. (1996). Epistemologies, conceptions of learning, and study practices in medicine and psychology. *Higher Education*, 31, 5-24.
- 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.

Verschillen in leerstijlen: Sekse en sekse-identiteit

Sabine Severiens en Geert ten Dam

Instituut voor de Lerarenopleiding, Universiteit van Amsterdam

ABSTRACT

Two recent reviews of gender differences in learning styles showed heterogeneous results in a variety of studies. In the reviewed studies, only biological gender was considered, even though gender is usually assumed to be a cultural phenomenon. In the present study on learning styles we do not only use gender as a dichotomous variable but gender-identity, as it is socially constructed, as well. 103 adult students in general secondary education completed the Inventory of Learning Styles (ILS) and the Dutch Sex Role Inventory. Only the first three factors which are usually found in a factor-analysis of the ILS were also found in our sample. The fourth factor appears to be typical for adult education: it can be interpreted as a learning style aimed at obtaining a certificate for that particular subject. Gender differences were found on the meaning directed learning style: men more often use such a learning style compared to women. Besides, correlations appear between this learning style and male gender identity, and also between the female gender identity and the undirected learning style. Regression analyses show that gender identity is a better predictor of the meaning directed learning style compared to gender. We conclude this article by elaborating on gender identity and the viewpoint it offers in the research area of gender and learning.

1. INLEIDING

Leerstijlen en leerstrategieën¹ zijn in de afgelopen jaren door diverse onderzoekers in verband gebracht met sekseverschillen. De resultaten geven geen eenduidig antwoord op de vraag in hoeverre leerprocessen seksspecifiek zijn (Richardson & King, 1991; Severiens & ten Dam, 1994). Met behulp van een meta-analyse op de onderzoeken waarin gebruik is gemaakt van Kolb's *Learning Style Inventory* (1976) en Entwistle's *Approaches to Studying Inventory* (1981) hebben we getracht een meer coherent beeld te geven van de eventuele sekseverschillen in leren (ten Dam & Severiens, 1995). In de meta-analyse werd duidelijk dat 'sekse' globaal gezien, dat wil zeggen over de verschillende onderzoeken heen, slechts in geringe mate differentieert naar leerstijlen/leerstrategieën.

Tussen de diverse onderzoeken bestaan echter, naast deze kleine gemiddelde verschillen, aanzienlijke onderlinge verschillen. In een aantal gevallen worden zelfs tegenstrijdige resultaten gevonden. Zo vinden Miller, Finley en McKinley (1990) en Richardson (1990) dat mannen vaker een betekenisgerichte leerstijl hanteren, terwijl Watkins en Hattie (1981) en Coles (1985) deze leerstijl juist vaker bij vrouwen tegenkomen. Een dergelijke uitkomst treffen we bijvoorbeeld ook aan bij een grootschalige meta-analyse op 100 onderzoeken naar seksspecifieke leerprestaties bij wiskunde. Gemiddeld genomen gaat het om (zeer) kleine verschillen, terwijl bijvoorbeeld binnen het hoger onderwijs duidelijke verschillen tussen vrouwen en mannen optreden in het voordeel van de laatsten (zie Hyde, Fennema & Lamon, 1990).

De uitkomst van de meta-analyse op onderzoeken naar leerstijlen/leerstrategieën kan op ten minste twee manieren worden geïnterpreteerd (zie ook Severiens & ten Dam, geaccepteerd). Allereerst kunnen sekseverschillen afhankelijk zijn van de onderwijscontext. Voor de uitgevoerde meta-analyse waren echter onvoldoende gegevens beschikbaar om na te gaan welke onderwijsfactoren (bijv. schooltype, vak, instructie-strategie) er toe doen.

Ten tweede zijn er vraagtekens te plaatsen bij de wijze waarop de variabele 'sekse' is geoperationaliseerd. Dit aspect staat centraal in het artikel. In alle ons bekende onderzoeken naar sekseverschillen in leren wordt een onderscheid gemaakt tussen *biologische* vrouwen en

mannen. Dit is opvallend aangezien vrijwel geen van de onderzoekers de gevonden sekseverschillen expliciet uit biologische eigenschappen of vaardigheden verklaart. In de introductie op de vraagstelling of in de conclusie wordt, in wisselende bewoordingen, vaak gesteld dat sekseverschillen in leerstijlen/leerstrategieën begrepen moeten worden als resultaat van socialisatieprocessen. Terwijl dus in de theoretische inkadering 'sekse' wordt benaderd op het niveau van 'vrouwelijke' en 'mannelijke' eigenschappen en gedragingen van de respondenten, wordt in de gehanteerde onderzoeksinstrumenten 'sekse' onderzocht op het niveau van geslacht (biologie).

Het denken in termen van socialisatie impliceert echter dat de biologische sekse en de culturele betekenis die daaraan wordt gegeven ('gender') zowel theoretisch als empirisch moet worden onderscheiden. In het onderzoek naar leerstijlen van vrouwen en mannen in het dag- en avondonderwijs voor volwassenen, dat in dit artikel wordt beschreven, hanteren we daarom niet alleen sekse als dichotome (biologische) variabele, maar tevens de mate waarin vrouwen en mannen zichzelf als vrouwelijk of mannelijk *percipiëren*.

2. SEKSE ALS SOCIAAL CONSTRUCT: SEKSE-IDENTITEIT ALS VARIABLE

In haar oratie 'Sekse verplicht' beschrijft Willemsen (1993) de betekenis van sekse als het resultaat van de op een bepaald moment binnen een bepaalde cultuur heersende opvatting over wat vrouwelijk is en wat mannelijk. Zij vat hiermee het denken binnen emancipatie- en vrouwenstudiesonderzoek over 'sekse' kort samen. Sinds het midden van de jaren tachtig worden de begrippen vrouwen/mannen en vrouwelijkheid/mannelijkheid in toenemende mate opgevat als een sociale constructie. De inhoud van deze begrippen wordt verondersteld te variëren naar plaats en tijd, en gezien als produkt van historische en culturele omstandigheden.

Betekenenissen van vrouwelijkheid en mannelijkheid komen veelal tot uiting in de vorm van associatieketens. In het onderwijs zijn de nieuwe vakken verzorging en techniek bijvoorbeeld al snel geassocieerd met de begrippen vrouwelijk en mannelijk (Robijns & Volman, 1991). Deze maatschappelijke betekenistoekenning werkt door op individueel niveau. Sociale verbanden, zoals het gezin, de peer-groep of de school dragen bij aan de identiteit van individuen. Onder andere op school leren meisjes en jongens op een bepaalde manier over zichzelf te denken, over vakken en vaardigheden en over leren en werken in het algemeen. De beelden die zij in het lesmateriaal van sekseverhoudingen krijgen voorgeschoteld en de manieren waarop leerlingen met elkaar omgaan, spelen daarbij een rol. Ook de (seksespecifieke) verwachtingen die leerkrachten hebben van de eigenschappen, vaardigheden en capaciteiten van leerlingen, vinden hun weg in de beleving en het gedrag van individuele meisjes en jongens. Maar meisjes en jongens nemen niet zomaar de beeldvorming ten aanzien van sekse over, ze geven zelf actief zin aan betekenissen die ze in hun omgeving tegenkomen (Henriques, Hollway, Urwin, Venn & Walkerdine, 1984; Hollway, 1989; ten Dam & Volman, 1995). Deze manier van denken vertoont overeenkomst met het constructivisme door het accent dat wordt gelegd op betekenisgeving (Duffy & Jonassen, 1991): kennen (over de wereld en de eigen plaats daarin) speelt zich af in een wisselwerking tussen persoon en omgeving.

Met het begrip sekse-identiteit kan recht worden gedaan aan het subjectieve karakter van de kennis die mensen over zichzelf als vrouw of man hebben. Sekse-identiteit duidt op de manier waarop individuen zichzelf als vrouwelijk of mannelijk waarnemen. Het begrip vormt een invalshoek om op individueel niveau handen en voeten te geven aan een opvatting van sekse als sociaal veranderlijk verschijnsel.

In dit artikel richten we ons op samenhangen tussen sekseverschillen en leerstijlen in termen van sekse als dichotome (biologische) variabele en de sekse-identiteit van leerlingen. De onderzoeksvraag luidt dan ook: Hoe ziet de samenhang eruit tussen leerstijlen, sekse en sekse-identiteit van leerlingen in het dag- en avondonderwijs voor volwassenen?

3. METHODE

Respondenten

In het onderzoek participeerden drie scholen voor dag- en avondonderwijs in de Randstad. Op de drie scholen zijn in verschillende vooreindexamenklassen havo en vwo vragenlijsten afgenomen bij vier verschillende vakken: wiskunde (12% van de respondenten), Nederlands (43%), biologie (16%) en geschiedenis (29%). Van alle respondenten zat 38% op het vwo, 17% op het havo en 45% in een gecombineerde havo-vwo (vierde) klas. De vragenlijsten zijn tijdens de les ingevuld.

Van de 103 respondenten die de vragenlijsten invulden, waren er 67 vrouw en 33 man, van 3 leerlingen was de sekse onbekend. Landelijk is de verhouding vrouwen/mannen in de volwasseneneducatie op het havo en vwo: 65% – 35% (bron: CBS, 1991). In onze steekproef is de verhouding 67% – 33%. De leeftijd van de respondenten varieerde van 17 tot 65 jaar: 58% zat in de leeftijdsgroep tot 21 jaar.

Instrumenten

In deze studie is gebruik gemaakt van twee gestandaardiseerde instrumenten. De *Inventaris Leerstijlen* voor het voortgezet onderwijs (ILS-VWO) (Roosendaal & Vermunt, 1993) is gebruikt om het begrip leerstijl te meten². In dit instrument wordt een brede definitie van leerstijlen gehanteerd: naast de leeraanpak (verwerking en sturing) wordt aandacht besteed aan mentale leermodellen en leeroriëntaties. De vragenlijst bestaat uit 120 vragen die op 5-puntschalen beantwoord worden. 20 schalen liggen ten grondslag aan de items. De betrouwbaarheden (Cronbach's alfa's) van de schalen van deze aangepaste ILS varieerden van .70 to .90.

Het begrip sekse-identiteit is gemeten met behulp van de *Nederlandse Sekse Rol Vragenlijst* (Willemsen & Fischer, 1992). De lijst beoogt de mate te meten waarin iemand zichzelf beschrijft in termen van vrouwelijke dan wel mannelijke eigenschappen en gedragingen. Het gaat om culturele opvattingen over mannelijkheid en vrouwelijkheid, dat wil zeggen om stereotiepe kenmerken en gedragingen die over het algemeen, in de Nederlandse samenleving, vaker aan één van de twee seksen toegeschreven worden. De vragenlijst bestaat uit twee delen. In het eerste deel worden respondenten gevraagd zichzelf te scoren op 30 eigenschappen. Dit deel van de lijst is gebaseerd op Bem's Sex Role Inventory (1974) en op Nederlands onderzoek naar seksestereotiepe eigenschappen. Het gaat om eigenschappen als afhankelijk, besluiteloos, ijverig, nieuwsgierig enz. In het tweede deel wordt gevraagd naar gedragingen, in zowel privé als publieke situaties. De gedragsitems betreffen vier gebieden: seks en relaties, emoties, sociale (publieke) rollen/interactie, en werk en hobby's. De items hebben bijvoorbeeld betrekking op wie op welke terreinen het initiatief neemt en welke (emotionele) activiteiten men ontplooit. Dit deel van de vragenlijst is deels gebaseerd op Orlofsky en O'Heron (1987). Elke respondent krijgt vier scores op deze lijst: één score voor de vrouwelijke eigenschappen-schaal, één voor de mannelijke eigenschappen-schaal, één voor de vrouwelijke gedrags-schaal en één voor de mannelijke gedrags-schaal. De betrouwbaarheden van de vier schalen in onze steekproef waren: vrouwelijke eigenschappen .73, mannelijke eigenschappen .77, vrouwelijk gedrag .68 en mannelijk gedrag .74.

Beide instrumenten zijn zelfrapportage instrumenten: respondenten geven aan in hoeverre ze bepaald gedrag vertonen of eigenschappen op zich van toepassing achten. Of respondenten die bijvoorbeeld zeggen de leerstof vaak kritisch te verwerken, dit ook daadwerkelijk doen blijft onduidelijk. In het geval van de ILS is een start gemaakt met de validering middels de hardopdenk methode, de resultaten zijn echter vooralsnog niet eenduidig te interpreteren (Prins, Busato, Hamaker & Visser, 1996). Het bovenstaande is inherent aan zelfrapportage instrumenten en bij het interpreteren van de resultaten dient hier dan ook rekening mee gehouden te worden.

Dataverwerking

In eerste instantie is nagegaan of de leerstijlen zoals Roosendaal en Vermunt (1993) ze observeren in het voortgezet onderwijs ook in de volwasseneneducatie onderscheiden kunnen worden. In feite betreft het hier het probleem van de externe validiteit van het leerstijlen construct. Qua inhoud is het reguliere onderwijs gelijk aan het onderwijs op scholen voor volwassenen en vanwege deze reden zou het logisch zijn dat ongeveer hetzelfde leerstijlenpatroon zou verschij-

nen. Maar de leerlingen in beide typen onderwijs verschillen wel van elkaar. Ten eerste is de gemiddelde leeftijd hoger, maar de motivatie van de leerlingen in het volwassenenonderwijs is, over het algemeen, ook anders dan die van de 'gewone' leerlingen in het voortgezet onderwijs. Vaak maken leerlingen in het volwassenenonderwijs een bewuste keuze voor het volgen van een bepaald vak en zijn de jongeren erg gemotiveerd om een diploma te halen zodat ze door kunnen stromen naar vervolgonderwijs. De verschillen tussen leerlingen uit beide typen onderwijs zouden een verschil in factorpatronen kunnen opleveren.

Om dit te onderzoeken zijn de schalen van de ILS in een principale componenten-analyse geanalyseerd. Daarbij zijn de eerste vier factoren geïnterpreteerd als een leerstijl. De steekproef in het huidige onderzoek is relatief klein, maar omdat op schaalniveau wordt geanalyseerd in plaats van op variabele niveau is het mogelijk een dergelijke analyse uit te voeren. Bovendien verwachten we op basis van eerder onderzoek met de ILS (VWO) een duidelijk factorpatroon met hoge ladingen. Guadagnoli en Velicer (1988) laten in een empirische studie naar stabiliteit van factorpatronen zien dat in het geval van vier of meer hoge ladingen (hoger dan .60) op een factor de grootte van de steekproef de oplossing nauwelijks beïnvloedt. Vooruitlopend op de resultaten, in ons geval hebben alle vier de factoren drie of meer ladingen van .60 of hoger. Dit geeft aan dat de analyse waarschijnlijk redelijk betrouwbare gegevens op zal leveren, hoewel natuurlijk toch enige voorzicht geboden is bij het generaliseren van de onderzoeksresultaten naar de volwasseneneducatie in het algemeen.

Om de onderzoeksvraag te beantwoorden, is geanalyseerd met de leerstijlen die de factoranalyse opleverde³. Middels *t*-toetsen zijn de verschillen tussen vrouwen en mannen bekeken op de vier leerstijlen. Door correlaties te berekenen tussen deze leerstijlen en de NSRV hebben we onderzocht in hoeverre er verbanden bestaan tussen leerstijlen en sekse-identiteit. Na deze exploraties is met behulp van regressie-analyses onderzocht in hoeverre sekse-identiteit een betere verklaring geeft van leerstijlen vergeleken met sekse.

4. RESULTATEN

Leerstijlen

Combinaties van schalen van de ILS, leveren vier onderscheiden leerstijlen op (Roosendaal & Vermunt, dit themanummer). De *betekenisgerichte* leerstijl kenmerkt zich door hoge scores op diepte-verwerking, zelfsturing, opbouw van kennis en persoonlijke interesse. De *reproductiegerichte* leerstijl kenmerkt zich door stapsgewijze verwerking, externe sturing, opname van kennis, certificaat en beroepsgerichte oriëntatie. Mensen met een *toepassingsgerichte* leerstijl scoren hoog op concrete verwerking, externe sturing (leerresultaten), gebruik van kennis en persoonlijke interesse. De *ongerichte* leerstijl ten slotte kenmerkt zich door stuurloosheid, ambivalentie en een voorkeur voor stimulerend onderwijs.

Om te controleren of dit leerstijlenpatroon ook in onze respondentengroep aanwezig is, verrichtten we een principale componenten-analyse met vier factoren. In tabel 1 worden de resultaten van deze analyse vermeld. Na bestudering van de factorladingen kan niet zonder meer geconcludeerd worden dat de vier leerstijlen gevonden door Roosendaal en Vermunt ook in deze groep onderscheiden kunnen worden. De eerste factor kan geïnterpreteerd worden als *betekenisgericht*: hoge ladingen verschijnen op diepte-verwerkingsstrategieën, analyseren en concreet verwerken, op zelfsturing en opbouw van kennis. De tweede factor duidt op een *ongerichte leerstijl*: stuurloos leergedrag, opname van kennis, verwachten dat het onderwijs je stimuleert, een voorkeur hebben voor samen studeren en een ambivalente leeroriëntatie, laden hoog op deze factor. De ladingen van opbouw van kennis, gebruik van kennis en certificaatgericht op deze tweede factor zijn minder goed te interpreteren. De derde factor laat verschillende componenten van de *reproductiegerichte leerstijl* zien: analyseren, externe sturing en testgericht zijn. Opname en (net als op de tweede factor) gebruik van kennis hebben een redelijke lading op deze factor. De laatste factor komt niet overeen met de vierde (*toepassingsgerichte*) factor die Roosendaal en Vermunt vonden. In onze groep kenmerkt deze factor zich door memoriseren en herhalen, niet persoonlijk geïnteresseerd zijn, certificaat- en beroepsgericht

zijn. Deze vierde factor zou uniek kunnen zijn voor de volwasseneneducatie: een relatief grote groep is immers gericht op het behalen van nog één of twee certificaten om zo snel mogelijk aan een vervolgopleiding of werk te kunnen beginnen. In de rest van dit artikel noemen we deze factor: de certificaatgerichte leerstijl.

Dat de eerste factor vooral uit strategieën bestaat en de tweede en vierde uit mentale leermodellen en leeroriëntaties, wordt ook in ander onderzoek met de ILS geconstateerd (zie Vermunt, 1995). Hij schrijft de soortgelijke bevindingen toe aan een ontwikkeling in leerstijlen tijdens het voortgezet onderwijs. In de eerste klassen blijken de relaties relatief zwak te zijn, terwijl in de vierde klas het gebruik van strategieën in iets grotere mate wordt geleid door een mentaal leermodel en leeroriëntatie. Pas in het hoger onderwijs is dit verband duidelijk aanwezig. Dat de ladingen soms moeilijk te interpreteren zijn, kan er ook op duiden dat leerstijlen in deze fase van het voortgezet onderwijs (4 havo) nog niet zijn uitgekristalliseerd.

Tabel 1. Factorladingen van de schalen van de ILS-VWO, analyse met 4 factoren en varimax rotatie.

	F1	F2	F3	F4
Verwerkingsstrategieën:				
Diepte-verwerking				
relateren & structureren	.79			
kritisch verwerken	.75			
Stapsgewijze verwerking				
memoriseren en herhalen				.46
analyseren	.48		.60	
Concrete verwerking	.74			
Regulatiestrategieën:				
Zelfsturing				
leerproces & -resultaten	.70			
leerinhoud	.74			
Externe sturing				
leerproces			.77	
leerresultaten			.73	
Stuurloos leergedrag		.59		
Mentale leermodellen:				
Ophouw van kennis	.60	.50		
Opname van kennis		.65	.37	
Gebruik van kennis	.42	.42	.38	
Stimulerend onderwijs		.73		
Samen studeren		.66		
Leeroriëntaties:				
Persoonlijk geïnteresseerd				-.77
Certificaatgericht		.38		.70
Testgericht		.43	.49	
Beroepsgericht				.66
Ambivalent		.57		
Eigenwaarde	4.6	3.6	1.8	1.7
% verklaarde variantie	23%	18%	9%	8%

Noot. Alleen ladingen >.35 en <-.35 worden vermeld.

Om antwoord te geven op de vraag of er verschillen bestaan op de ILS tussen mannen en vrouwen die deelnemen aan de volwasseneneducatie, is gewerkt met de vier factoren die hierboven beschreven zijn. Daarbij hebben we de volgende regels gehanteerd: schalen worden in een leerstijl opgenomen als de lading hoger dan .45 is op de betreffende factor en schalen mogen maar in één leerstijl opgenomen worden. De scores op de vier leerstijlen zijn een som van de schalen. In tabel 2 wordt afgebeeld hoe de vier factoren (leerstijlen) zijn samengesteld. Omdat we in eerste instantie geïnteresseerd zijn in de verbanden op 'leerstijl-niveau' in plaats van op 'schaalniveau' en om bovendien de kapitalisatie op kans te verminderen, het aantal *p*-waarden wordt immers drastisch vergroot bij analyses op schaalniveau, komen in de resultaten vooral de leerstijlen ter sprake.

In een betrouwbaarheidsanalyse komen de volgende Cronbach's alfa's naar voren: de betekenisgerichte leerstijl: .83, de ongerichte leerstijl: .72, de reproductiegerichte leerstijl: .66 en de certificaatgerichte leerstijl: .62.

Uit tabel 3 blijkt dat mannen op de betekenisgerichte leerstijl hoger scoren dan vrouwen. Gemiddeld genomen zeggen mannen vaker dan vrouwen gebruik te maken van de diepte-verwerkingsstrategieën, zoals Vermunt (1992) de verwerkingsstrategieën in deze leerstijl noemt. In vergelijking met vrouwen lijken mannen kritischer in hun verwerking van de leerstof, ze zeggen de stof vaker te structureren en de afzonderlijke onderdelen uit cursussen aan elkaar te relateren. Verder vertalen ze de stof vaker naar de dagelijkse werkelijkheid, ze geven vaker aan dan vrouwen een zelfgestuurde leerstrategie te gebruiken en zeggen leren vaker te zien als het opbouwen van kennis.

Op de overige leerstijlen zijn geen statistisch significante verschillen tussen vrouwen en mannen gevonden.

Het onderzoek levert daarmee andere resultaten op dan het onderzoek van Vermunt (1992) bij Ou-studenten en KUB-studenten. Vermunt vond bij Ou-studenten dat (p. 141) vrouwen

Tabel 2. De vier leerstijlen en hun schalen.

betekenisgerichte leerstijl:

- relateren en structureren
- kritisch verwerken
- concreet verwerken
- zelfsturing, leerproces & -resultaten
- zelfsturing, leerinhoud
- opbouw van kennis

de ongerichte leerstijl:

- stuurloos leergedrag
- opname van kennis
- stimulerend onderwijs
- samen studeren
- ambivalent

de reproductiegerichte leerstijl:

- analyseren
- externe sturing, leerproces
- externe sturing, leerresultaten
- testgericht

de certificaatgerichte leerstijl:

- memoriseren en herhalen
 - persoonlijk geïnteresseerd (omgecodeerd)
 - certificaatgericht
 - beroepsgericht
-

Tabel 3. Gemiddeldes, standaard deviaties en t 's op de vier leerstijlen.

	N	$M_{\text{♀♀}}$	SD	$M_{\text{♂♂}}$	SD	t	p
de betekenisgerichte leerstijl	94	2.68	.55	2.98	.67	-2.32	.02
de ongerichte leerstijl	94	2.87	.54	2.85	.54	0.14	.89
de reproductiegerichte leerstijl	95	3.11	.54	3.12	.70	-0.04	.97
de certificaatgerichte leerstijl	95	3.02	.69	3.21	.72	-1.22	.23

Noot. De N in deze tabel wijkt af van de totale N vanwege ontbrekende gegevens.

hoger scoren op de schalen van een reproductiegerichte leerstijl (externe regulatiestrategie, memoriserende verwerkingsstrategie, reproductief mentaal leermodel). Aangezien de onderzoeksgroepen (universitaire studenten en deelnemers aan het dag- en avondonderwijs voor volwassenen) in verscheidene opzichten van elkaar verschillen (bijv. vooropleiding, 'school-type'), is niet op voorhand te zeggen waar de verschillen tussen beide aan toe te schrijven zijn.

Sekse-identiteit

Vrouwen en mannen schrijven zichzelf vaak de eigenschappen en gedragingen toe die 'horen' bij hun sekse (zie tabel 4). Op drie van vier schalen in de NSRV wordt significant verschillend gescoord in de verwachte richtingen. Vrouwen scoren hoger op vrouwelijk gedrag, mannen scoren hoger op mannelijk gedrag en mannelijke eigenschappen. Op de schaal voor vrouwelijke eigenschappen scoren vrouwen wel iets hoger dan mannen maar dit verschil is niet significant.

Leerstijlen en sekse-identiteit

Respondenten die zichzelf meer beschrijven in termen van mannelijk gedrag scoren hoger op de betekenisgerichte leerstijl (zie tabel 5). Op zich is dit niet zo verwonderlijk als we zien dat mannen, die in vergelijking met vrouwen hoger scoren op deze leerstijl, ook significant hoger scoren op de schaal voor mannelijk gedrag (zie tabel 4). Het is wel opvallend dat de gevonden correlaties alleen betrekking hebben op mannelijk gedrag: mannelijke en vrouwelijke eigenschappen en vrouwelijk gedrag doen er niet in dezelfde mate toe als het gaat om sekseverschillen op de betekenisgerichte leerstijl. Net als bij verschillen tussen vrouwen en mannen blijkt hier dat de schalen voor verwerkingsstrategieën 'kritisch verwerken' en 'relateren en structureren' statistisch significante correlaties opleveren. De correlatie tussen 'kritisch verwerken' en de mannelijke gedrags-schaal is .39 ($p < .01$) en de correlatie tussen 'relateren en structureren' en de schaal voor mannelijk gedrag is .26 ($p < .05$). De correlaties verschijnen dus tussen twee 'gedrags-maten', het gaat om studie-activiteiten en typische mannelijke activiteiten. Als we naar de inhoud van de items kijken, dan liggen deze correlaties niet direct voor de hand: de

Tabel 4. Sekse en de scores op de NSRV.

	N	$M_{\text{♀♀}}$	SD	$M_{\text{♂♂}}$	SD	t	p
vrouwelijke eigenschappen	100	53.4	7.1	52.0	5.7	1.0	.30
vrouwelijk gedrag	100	40.5	7.0	34.8	5.4	4.1	.00
mannelijke eigenschappen	100	48.8	7.2	53.6	6.9	-3.2	.00
mannelijk gedrag	100	32.8	6.6	40.6	7.0	-5.5	.00

Noot. De N in deze tabel wijkt af van de totale N vanwege ontbrekende gegevens.

Tabel 5. Correlaties tussen de ILS-VWO en NSRV.

	vrouwelijke eigenschappen	mannelijk gedrag	mannelijke eigenschappen	vrouwelijk gedrag
de betekenisgerichte leerstijl	-.04 ($p=.71$)	-.06 ($p=.59$)	.15 ($p=.13$)	.29 ($p=.00$)
de ongerichte leerstijl	.23 ($p=.03$)	.04 ($p=.69$)	-.10 ($p=.31$)	-.11 ($p=.30$)
de reproductiegerichte leerstijl	-.04 ($p=.72$)	-.12 ($p=.23$)	-.04 ($p=.72$)	-.12 ($p=.23$)
de certificaatgerichte leerstijl	.16 ($p=.11$)	.03 ($p=.79$)	.09 ($p=.39$)	.01 ($p=.93$)

Noot. In alle gevallen is de N 97: afwijkend van de totale N vanwege ontbrekende gegevens.

gedragsitems van de NSRV betreffen andere gebieden dan de studie-activiteiten. De overige schalen in deze leerstijl correleren niet significant met deze schaal van de NSRV.

Tabel 5 laat tevens zien dat respondenten die hoger scoren op de vrouwelijke eigenschappen-schaal vaker aangeven een ongerichte leerstijl te hebben. Respondenten die hoger scoren op deze schaal lijken ook hoger te scoren op stuurloosheid, samenwerken, ambivalentie, stimulerend onderwijs en opname van kennis. De schalen van deze ongerichte leerstijl correleren overigens afzonderlijk niet significant met de schaal voor vrouwelijke eigenschappen⁴.

Verschillen in leerstijlen of leerstrategieën tussen studenten zijn onder andere interessant omdat ze binnen een bepaald type onderwijs tot verschillende prestaties kunnen leiden. Voor het hoger onderwijs vond Vermunt een aantal verbanden tussen leerstrategieën en tentamenresultaten. Met name relateren en structureren vormt een succesvolle verwerkingsstrategie, terwijl stuurloos leergedrag een niet-succesvolle regulatie'strategie' is (1992, 166). Deze strategieën vormen een aspect van respectievelijk de betekenisgerichte leerstijl en de ongerichte leerstijl. We willen opmerken dat het precies deze leerstijlen zijn die in onze steekproef van deelnemers aan het dag- en avondonderwijs voor volwassenen correleren met sekse-identiteit: mannelijk gedrag correleert met de betekenisgerichte leerstijl en vrouwelijke eigenschappen met de ongerichte leerstijl.

Sekse, sekse-identiteit en leerstijlen

De t -toetsen en correlatie-analyses wezen met name op de verbanden tussen twee leerstijlen (de betekenisgerichte en de ongerichte leerstijl) met sekse en sekse-identiteit. Als we een en ander in een regressiemodel vertalen, dan lijkt het erop dat sekse, maar ook sekse-identiteit voor een deel het gebruik van de betekenisgerichte leerstijl en de ongerichte leerstijl, kunnen voorspellen. Een regressie-analyse waarin we deze leerstijlen laten voorspellen door sekse als dummy-variabele en sekse-identiteit in termen van respectievelijk mannelijk gedrag en vrouwelijke eigenschappen, kan dit bevestigen. Tevens kan met deze methode worden getoetst of verschillen in leerstijlen beter door sekse dan wel door sekse-identiteit voorspeld kunnen worden.

De schaal voor mannelijk gedrag wordt als eerste variabele gekozen in een 'forward' regressie-analyse met de betekenisgerichte leerstijl als afhankelijke variabele. Sekse wordt vervolgens niet meer opgenomen in de vergelijking omdat deze variabele de grens van $p=.05$ niet meer haalt. Met andere woorden, als de variantie die mannelijk gedrag in deze leerstijl verklaart eruit is gehaald, verklaart de variabele sekse een te verwaarlozen hoeveelheid variantie. De multiële correlatie van de regressievergelijking waarin de betekenisgerichte leerstijl wordt voorspeld is .29. Omdat alleen deze variabele in de regressievergelijking is opgenomen is de multiële correlatie gelijk aan de 'gewone' correlatie in tabel 5. Het model verklaart dus 8% van de variantie in de afhankelijke variabele. Als sekse in de vergelijking wordt 'gedwongen' (met de methode 'enter') levert dit een multiële correlatie op van .32 (10% verklaarde variantie). Als sekse als enige variabele in de vergelijking wordt opgenomen, is de correlatie .19 en wordt dus 4% van de

variantie verklaard. Ook uit deze gegevens blijkt dat sekse minder variantie verklaart in leerstijlen, om precies te zijn 4% minder. Sekse-identiteit in termen van mannelijk gedrag is dus een betere voorspeller van de betekenisgerichte leerstijl dan sekse (als dichotome variabele).

Opvallend is dat we géén sekseverschillen constateren op de ongerichte leerstijl, maar wél sekse-identiteitsverschillen (als het gaat om de vrouwelijke eigenschappen-schaal). Vrouwen en mannen zeggen even vaak ongericht te leren, maar respondenten die zichzelf in hogere mate vrouwelijke eigenschappen toeschrijven, leren wel vaker ongericht. Het is verklaarbaar als we in ogenschouw nemen dat deze schaal van sekse-identiteit geen significante sekseverschillen opleverde in onze respondentengroep: vrouwen en mannen bleken even hoog te scoren op vrouwelijke eigenschappen. Ter vergelijking: de multiële correlatie (in feite de 'gewone' correlatie, zie tabel 5) van een regressievergelijking waarin de ongerichte leerstijl alleen door de vrouwelijke eigenschappen-schaal wordt voorspeld, is .23 (5% verklaarde variantie).

5. DISCUSSIE

Het gebruik van de variabele sekse-identiteit in onderzoek naar sekseverschillen in leerstijlen is een uitvloeisel van de theoretisch en empirisch gefundeerde opvatting dat de sekse van een persoon niet los van de cultuur vorm krijgt. Onder invloed van primaire en secundaire socialisatieprocessen gaan mensen zichzelf als vrouwelijk of mannelijk waarnemen. Dit proces werd in de zeventiger jaren als eenduidig en lineair begrepen: doordat meisjes en jongens in het gezin en op school verschillend worden behandeld, internaliseren ze de bij hun sekse passende waarden, normen, attitudes, verwachtingen enzovoorts en gaan ze zich op een bij hun sekse passende manier gedragen (Bekker, 1983, 79). Het beschouwen van mensen als gevormd door hun omgeving maakt het in zekere zin onnodig de laag van sekse-identiteit in onderzoek te onderscheiden. Immers, de biologische sekse van iemand bepaalt hoe hij of zij ingevoegd wordt in de cultuur. In de tweede helft van de jaren tachtig kwam meer aandacht voor de actieve rol die individuen in socialisatieprocessen spelen, waardoor de uitkomst soms anders is dan 'beoogd'. Bovendien werd beargumenteerd dat de sekse-identiteit van mensen niet stabiel is: al naar gelang de maatschappelijke en historische context treden veranderingen op. Het 'zelf' of de identiteit van een subject wordt steeds opnieuw gevormd door reflecties op en verhalen over gebeurtenissen en ervaringen (van Drenth, 1994, 146). De subjectieve beleving van mensen vormt dus als het ware een laag tussen biologische sekse en sociale sekse-verhoudingen. Eén en ander vormde voor ons de reden een maat voor sekse-identiteit mee te nemen in empirisch onderzoek naar sekseverschillen in leerstijlen.

In dit artikel is een onderzoek naar de leerstijlen van vrouwen en mannen in het dag- en avondonderwijs voor volwassenen beschreven. Een principale componenten-analyse op de schalen van de Inventaris Leerstijlen voor het VWO liet zien dat de betekenisgerichte, de ongerichte en reproductiegerichte leerstijl als verschillende leerstijlen in dit type onderwijs onderscheiden kunnen worden. De vierde leerstijl in onze respondentengroep leek op een 'zo snel mogelijk je certificaat halen' leerstijl. Het memoriseren en herhalen laadde niet op de reproductiegerichte leerstijl, maar op deze vierde leerstijl. Verder werd deze vierde leerstijl gekenmerkt door het niet persoonlijk geïnteresseerd zijn en wel certificaat- en beroepsgericht zijn. In een vervolgonderzoek zal moeten blijken of deze factor inderdaad een consistente factor is in het dag- en avondonderwijs voor volwassenen.

De gevonden verschillen in leerstijlen zijn in samenhang gebracht met de (biologische) sekse van de deelnemers en hun sekse-identiteit. Mannen scoren gemiddeld hoger op de betekenisgerichte leerstijl, op de overige leerstijlen vonden we geen sekseverschillen.

Sekse-identiteit, zoals gemeten met behulp van de *Nederlandse Sekse Rol Vragenlijst*, wordt gekarakteriseerd door de perceptie van mensen van zowel de eigenschappen die men heeft, als het gedrag dat men vertoont. In dit onderzoek blijken verschillen in leerstijlen tussen mensen deels door hun seksspecifieke gedrag en eigenschappen voorspeld te kunnen worden. De mate waarin iemand naar eigen zeggen mannelijk gedrag vertoont, hangt samen met de mate waarin

degene de betekenisgerichte leerstijl zegt te gebruiken. De mate waarin iemand zichzelf vrouwelijke eigenschappen toeschrijft hangt samen met de mate waarin iemand aangeeft de ongerichte leerstijl te gebruiken.

De relaties tussen sekse, sekse-identiteit en de betekenisgerichte leerstijl, zijn met behulp van regressie-analyses onderzocht en getoetst. Sekse-identiteit in termen van mannelijk gedrag bleek een betere voorspeller van de betekenisgerichte leerstijl te zijn dan sekse (geoperationaliseerd als dichotome variabele). De variabele sekse verklaart minder variantie in deze leerstijl. Als het gaat om de ongerichte leerstijl bleek sekse geen voorspeller te zijn en is sekse-identiteit dus in ieder geval een betere voorspeller.

Leerstijlen kunnen worden opgevat als het resultaat van de interactie tussen persoons- en omgevingsgebonden invloeden (Vermunt, 1992, 21). Bij wijze van afsluiting interpreteren we de onderzoeksresultaten vanuit de rol van het onderwijs als omgevingsfactor. Een aanwijzing voor de veronderstelling dat het onderwijs een bijdrage levert aan sekseverschillen in leerstijlen, levert het onderzoek van Dippelhofer-Stiem (1989). In een longitudinale studie uitgevoerd in vijf Europese landen naar een 'research oriented' leerstijl (een leerstijl die kritisch denken, autonomie, diepte-verwerking, relateren en intrinsieke motivatie omvat), laat zij zien hoe deze leerstijl zich in de loop van een studie in het hoger onderwijs met name bij mannen ontwikkelt. In het eerste semester zijn de scores van mannen en vrouwen op deze leerstijl heterogeen, dat wil zeggen sekseverschillen variëren al naar gelang het land en de discipline. Aan het eind van hun studie laten mannen echter, onafhankelijk van het land of de discipline, een meer ontwikkelde 'research oriented' leerstijl zien dan vrouwen. In het onderzoek van Dippelhofer-Stiem is geen aandacht besteed aan de onderwijsfactoren die voor deze seksspecifieke ontwikkeling van belang zijn.

Deelnemers aan het dag- en avondonderwijs voor volwassenen hebben alle een (gebroken) schoolloopbaan achter de rug waarin hun leerstijlen en leerstrategieën zijn ontwikkeld. Materiaal voor een interpretatie van de gevonden sekseverschillen waarin het onderwijs een rol speelt, levert het in de afgelopen jaren verrichte onderzoek naar sekse-ongelijkheid in het onderwijs. Räsänen (1992) relateert de geringere voorkennis van meisjes van het vak natuurkunde aan de leeractiviteiten die zij bij dit vak ontwikkelen. Daarnaast kan het handelen van de docent een factor van betekenis zijn. Uit verschillende onderzoeken is naar voren gekomen dat het handelen van de docent (als belangrijk aspect van instructie) niet in alle facetten sekse-neutraal is. Met name de eigenschappen en vaardigheden die docenten hun leerlingen toedichten, worden gestuurd door opvattingen over sekse. In bijvoorbeeld interviews die Clark (1989) hield met leerkrachten in het basisonderwijs komt naar voren dat deze docenten meisjes zien als minder intelligent en interessant dan jongens, terwijl meisjes zich juist als 'modelleerlingen' gedragen. Meisjes zijn ijverig, maar ook voorspelbaar, jongens zijn een uitdaging. Walkerdine (1989) laat zien hoe docenten de prestaties van meisjes interpreteerden als het resultaat van hard werken en het opvolgen van regels, bij jongens zagen zij creativiteit en inzicht. (Zie voor de Nederlandse situatie ten Dam, van Eck & Volman, 1992).

Op basis van het in dit artikel gehanteerde theoretisch kader en de genoemde onderzoeksresultaten met betrekking tot de rol van het onderwijs kan verondersteld worden dat de opvattingen van docenten over mannelijke en vrouwelijke leerlingen samenhangen met de sekse-identiteit van leerlingen en met de leeractiviteiten die zij ontplooiën. Wie als kritisch denkend wordt aangesproken en wordt uitgedaagd inzicht te ontplooiën, zal de eigen identiteit daaraan ontleunen en deze vaardigheden gaandeweg beter ontwikkelen. Het zinvolle van het benaderen van het probleem in termen van sekse-identiteit in plaats van sekse als dichotome variabele, is dat het inzichtelijk kan maken dat docenten niet steeds op dezelfde wijze op leerlingen van het vrouwelijke danwel het mannelijke geslacht reageren. Onze bevinding dat er geen sekseverschillen waren op de ongerichte leerstijl, vergeleken bij de correlatie tussen de vrouwelijke identiteit met deze leerstijl, laat zien dat er van een eenduidig lineair verband tussen sekse, sekse-identiteit en leerstijlen geen sprake is. De verbanden tussen kritische denkers en de mannelijke identiteit enerzijds en tussen stuurloosheid en de vrouwelijke identiteit anderzijds lieten in het eerste geval wel een verband met sekse (in termen van geslacht) zien, maar in het tweede geval niet.

Dat de samenhang tussen leerstijlen en sekse-identiteit evenwel sterker bleken te zijn dan de samenhang met sekse betekent een nuancering van en daarmee een verdieping voor het onderzoeksdomein sekse en leren.

DANKBETUIGING

We willen Gert Rijlaarsdam en Jan Vermunt hartelijk bedanken voor hun constructieve kritiek op een eerdere versie van dit artikel.

NOTEN

1. In de leerpsychologie zijn verschillende termen in omloop om iets te zeggen over de wijze waarop mensen leren kennen en kennis verwerven (bijv. leeraanpak, leerstrategie, leerstijl). De verschillen in terminologie verwijzen met name naar de – veronderstelde – invloed van het onderwijs/leertaak op het leerproces. In de meeste onderzoeken worden 'leerstijlen' bijvoorbeeld gedefinieerd als min of meer stabiele persoonlijkheidskenmerken, onafhankelijk van onderwijskundige factoren (bijv. Riding & Cheema, 1991). In Nederland definieert Vermunt (1992) in navolging van Schmeck, Geisler-Brenstein en Cercy (1991) leerstijlen echter als het resultaat van de interactie tussen persoons- en contextfactoren.
2. In de inleiding van de vragenlijst is aan de respondenten gevraagd de ILS in te vullen vanuit één bepaald vak. In de vragenlijst zelf werden de respondenten hieraan herinnerd door in een aantal items het vak te noemen. Een voorbeeld hiervan is 'Ik doe meer dan dat me gevraagd wordt bij geschiedenis.'. De les waarin de vragenlijsten werden ingevuld, gold als uitgangspunt. Het achterliggende idee voor dit vak-specifieke invullen was dat leren gebonden is c.q. kan zijn aan de context van het vak (zie Meyer, Dunne & Richardson, 1994, voor sekseverschillen in 'gecontextualiseerd' studiegedrag). In dit artikel gaan we op de mogelijke vakafhankelijkheid van leerstijlen niet in. We rapporteren alleen verbanden tussen sekse, sekse-identiteit en leerstijlen en laten de vakafhankelijkheid buiten beschouwing.
3. Daarbij moet de kanttekening geplaatst worden dat als de leerstijlen uit andere schalen bestaan dan de leerstijlen uit de studie van Roosendaal en Vermunt een eventuele directe vergelijking niet meer mogelijk is. In dat geval zijn alleen de resultaten op schaalniveau te vergelijken.
4. Het feit dat vrouwelijke eigenschappen en ongericht leren positief correleren én dat mannelijk gedrag en betekenisgericht leren positief correleren kan een spurieus verband betekenen. Als vrouwelijke eigenschappen en mannelijk gedrag negatief gecorreleerd zijn én de betekenisgericht en ongericht ook negatief gecorreleerd zijn, dan zou men net zo goed kunnen concluderen dat die verbanden zijn gemeten. Maar dat is niet het geval: de mannelijke en vrouwelijke identiteitsschalen correleren niet met elkaar. Dit bevestigt overigens de assumptie dat vrouwelijkheid en mannelijkheid niet de uiteinden van één dimensie vormen. Betekenisgericht leren en ongericht leren correleren ook niet met elkaar. We mogen dus aannemen dat de gevonden correlaties inderdaad de bedoelde schalen betreffen.

LITERATUUR

- Bekker, M. (1983). Socialisatie. In S. Poldervaart. *Vrouwenstudies. Een inleiding* (pp.78-91). Nijmegen: SUN.
- Bem, S. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42, 155-162.
- CBS (1991). Volwasseneneducatie. In *Zakboek voor onderwijsstatistieken* (pp. 54-69). Den Haag: SDU-uitgeverij/CBS-publicaties.
- Clark, M. (1989). Anastasia is a normal developer because she is unique. *Oxford Review of Education*, 15, 243-255.
- Coles, C.R. (1985). Differences between conventional and problem-based curricula in their students' approaches to studying. *Medical Education*, 19, 308-309.
- Dam, G. ten & Severiens, S.E. (1995). Leerstijlen, leerstrategieën en leermotivatie. Een meta-analyse van het leren van vrouwen en mannen. *Pedagogische Studiën*, 72, 206-220.
- Dam, G. ten, Eck, E. van, & Volman, M. (1992). *Onderwijs en sekse. Een verkenning van researchprogramma's*. 's Gravenhage: DCE/STEO Trendrapport 4: 's-Gravenhage.

- Dam, G. ten & Volman, M. (1995). Feminist research and educational policy. *Journal of Education Policy*, 10, 209-220.
- Dippelhofer-Stiem, B. (1989). The development of research-oriented learning in five European countries. *European Journal of Psychology of Education*, 4, 489-503.
- Drenth, A. van (1994). Zorg-en-de-macht. Over de sociale constructie van gender identiteit. *Psychologie en Maatschappij*, 18, 141-155.
- Duffy, T.M. & Jonassen, D.H. (1991). Constructivism: New implications for instructional technology? *Educational Technology*, 30, 7-12.
- Entwistle, N.J. (1981). *Styles of learning and teaching*. New York: John Wiley & Sons.
- Guadagnoli, E. & Velicer, W.F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103, 265-275.
- Henriques, J., Hollway, W., Urwin, C., Venn, C. & Walkerdine, V. (1984). *Changing the subject: Psychology, social relations and subjectivity*. London: Methuen.
- Hollway, W. (1989). *Subjectivity and method in psychology. Gender, meaning and science*. London, Newbury Park, New Delhi: Sage Publications.
- Hyde, J., Fennema, E., & Lamon, S. (1990). Gender differences in mathematics performance: a meta-analysis. *Psychological Bulletin*, 107, 139-155.
- Kolb, D.A. (1976). *Learning Style Inventory: Technical manual*. (rev. edn. 1978). Boston: McBerand Company.
- Meyer, J.H.F., Dunne, T.T., & Richardson, J.T.E. (1994). A gender comparison of contextualised study behaviour in higher education. *Higher Education*, 4, 469-486.
- Miller, C.D., Finley, J., & McKinley, D.L. (1990). Learning approaches and motives: Male and female differences and implications for learning assistance programs. *Journal of College Student Development*, 31, 147-154.
- Orlofsky, J.L. & O'Heron, C.A. (1987). Development of a short form sex role behavior scale. *Journal of Personality Assessment*, 51, 267-277.
- Prins, F.J., Busato, N.V., Hamaker, C., & Visser, K.H. (1996). Een bijdrage tot de validatie van het (meta)cognitieve deel van de Inventaris Leerstijlen. *Pedagogische Studiën*, 73, 108-122.
- Räsänen, L. (1992). The gender gap in learning physics concepts. In *Ten years GASAT activities in a changing Europe. Contributions to the conference October 25-29 1992*, 159-168.
- Richardson, J.T.E. (1990). Reliability and replicability of the approaches to studying questionnaire. *Studies in Higher Education*, 15, 155-168.
- Richardson, J.T. & King, E. (1991). Gender differences in the experience of higher education: quantitative and qualitative approaches. *Educational Psychology*, 11, 363-382.
- Riding, R. & Cheema, I. (1991). Cognitive styles – an overview and integration. *Educational Psychology*, 11, 193-215.
- Robijns, M. & Volman, M. (1991). *Verzorging, een vak apart? Het leergebied Verzorging in de eerste fase voortgezet onderwijs*. Amsterdam/Lisse: Swets & Zeitlinger.
- Roosendaal, A. & Vermunt, J.D.H.M. (1993). *Congruenties en fricties tussen leer- en instructiestrategieën*. Paper gepresenteerd op de ORD, Maastricht.
- Severiens, S.E. & ten Dam, G.T.M. (1994). Gender differences in learning styles: a narrative review and a quantitative meta-analysis. *Higher Education*, 27, 487-501.
- Severiens, S.E. & ten Dam, G.T.M. (geaccepteerd). Gender and gender-identity differences in learning styles. *Educational Psychology*.
- Schmeck, R.R., Geisler-Brenstein, E., & Cerci, S.P. (1991). Self-concept and learning: the revised inventory of learning processes. *Educational Psychology*, 11, 343-362.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in hoger onderwijs. Naar procesgerichte instructie in zelfstandig denken*. Amsterdam/Lisse: Swets & Zeitlinger.
- Vermunt, J.D.H.M. (1995). Leerstijlen: een overzicht en recente onderzoeksgegevens. In H.C. Schouwenburg & J.T. Groenewoud (Red.). *Studievaardigheid en leerstijlen Proceedings van de 14e Landelijke Dag Studievaardigheden te Groningen* (pp. 51-72). Groningen: Wolters-Noordhoff.
- Walkerdine, V. (1989). *Counting girls out*. London: Virago.
- Watkins, D. & Hattie, J. (1981). The learning processes of Australian university students: Investigations of contextual and personological factors. *British Journal of Educational Psychology* 51, 384-393.
- Willemsen, T.M. & Fischer, A.H. (1992). [M / V] * * [of iets er tussenin?]. *Psychologie*, 13.
- Willemsen, T.M. (1993). *Sekse verplicht*. Tilburg: University Press.

Leeromgevingspercepties en leerstrategieën van eerstejaars studenten sociale wetenschappen

Ronny F.A. Wierstra en Eva P.M. Beerends

Universiteit Utrecht, Vakgroep Onderwijskunde en Bureau Smits & Beerends, Training en Begeleiding, Utrecht.

ABSTRACT

A conceptual model is presented for research on the interaction between learning style and learning environment. Students' perceptions of process aspects of the learning environment play a prominent part in this model. Relationships between components of the model are investigated with 210 students of social science at Utrecht University, in their first year ('propedeuse'). Learning style was measured by means of Vermunt's Inventory of Learning Styles (ILS). In order to measure learning environment factors a new questionnaire was developed: the Inventory of Process Aspects of the Learning environment (IPAL). The results of the study give support to the validity of both ILS and IPAL. A typology of learning environment perceptions was found which agrees remarkably well with the learning style typology. Students could be classified in three learning style clusters which differ according to their positions on two independent dimensions 'externally regulated reproductive learning' and 'internally regulated meaning directed learning'. Also, with regard to the learning environment perceptions a classification in three clusters was found. These clusters differ on the dimensions 'activating instruction' and 'knowledge transmission instruction'. The interactions found between learning style and learning environment variables are described.

1. INLEIDING

In dit artikel wordt verslag gedaan van een onderzoek dat verricht is tussen oktober '93 en februari '94 bij eerstejaars studenten sociale wetenschappen aan de Universiteit Utrecht naar hun leerstrategieën en naar hun percepties van de leeromgeving. Dit is een deelonderzoek uit een onderzoeksprogramma van de Utrechtse Vakgroep Onderwijskunde, waarin nagegaan wordt hoe het 'activerende' karakter van het onderwijs verhoogd kan worden, dat wil zeggen hoe een vorm van onderwijs ontwikkeld kan worden dat leidt tot een actiever leren van studenten: dieptegericht en (mede) door de student gestuurd leren (zie ook Vermunt, 1992; Boekaerts & Simons, 1993; Duffy & Jonasson, 1991). De moderne leer- en onderwijspsychologie, onder andere de constructivistische stroming, maakt duidelijk dat leren een actief proces is, waarin de lerende in hoge mate zelf verantwoordelijk is voor zijn leerproces. Dit geldt zeker voor een universitaire studie, waar het hanteren van cognitieve strategieën door de student niet alleen een middel is om kennis te verwerven, maar ook doel van het onderwijs (Van Hout-Wolters, 1992; De Klerk, 1992; Ten Brinke, 1994). De indruk bestaat dat veel studenten dit doel niet bereiken en met ongewenste leerstrategieën de leertaken tegemoet treden. Naast leerstijlen van studenten kunnen ook onderwijsstijlen hieraan debet zijn. Om deze reden willen we in ons onderzoek niet alleen leerstijlen van studenten in kaart brengen, maar ook de leeromgeving en de interactie tussen de beide variabelen nagaan. In paragraaf 2 beschrijven we nu eerst het conceptuele model van ons onderzoeksprogramma, waarna in de rest van dit artikel de probleemstelling en resultaten uitgewerkt worden van een deelonderzoek binnen dit programma.

Adressen: R.F.A. Wierstra, Universiteit Utrecht, Vakgroep Onderwijskunde, Heidelberglaan 2, 3584 CS Utrecht.

Eva Beerends, Bureau Smits & Beerends, Training en Begeleiding, Regentesselaan 16, 3571 CE Utrecht.

2. EEN INTERACTIONEEL MODEL

Een centrale vraag in ons onderzoek is hoe in een bepaalde onderwijsleereenheid (bijvoorbeeld een moduul) een bepaalde leerstijl en een bepaald type leeromgeving samen de actuele leerstrategie van de student bepalen. Die interactie bestuderen we vanuit het model dat in Figuur 1 is afgebeeld.

Het model is interactioneel in de zin dat een leerstijl (B) niet opgevat wordt als een over tijdstippen en situaties consistente manier van leren, maar als een persoonsdispositie binnen een (organisch) interactie-model (Magnusson & Endler, 1977; Pervin, 1977). Kenmerkend voor een *organisch* interactiemodel is een *wederzijdse* afhankelijkheid van persoonsfactoren en situatiekenmerken, waarbij de wijze waarop de persoon de situatie percipieert, de subjectieve betekenis die hij aan de situatie geeft, van cruciaal belang is voor het gedrag in die situatie. In het conceptuele model komt een en ander tot uiting door onderscheid te maken tussen de habituele leerstijl (B) van een student (de *dispositie* voor een bepaalde leerstrategie) en het in een bepaalde onderwijsleereenheid daadwerkelijk gerealiseerde leergedrag (de *geactualiseerde* leerstrategie) (F) en daartussen subjectieve perceptie- en waarderingsprocessen te positioneren als medierende processen. Het model veronderstelt dat een leerstijl een op een bepaald moment aanwezige dispositie is, die zich – ook op de universiteit nog – kan ontwikkelen en dat een student in de ene onderwijsleereenheid in principe op een andere wijze kan leren dan in de andere. Als er inderdaad – zoals het model stelt – invloed blijkt te zijn van de (aard van de) leeromgeving op de gevolgde leerstrategie, dan geeft dit enige hoop dat (ongewenste) leerstijlen in een bepaalde mate te beïnvloeden zijn door veranderingen in de leeromgeving en door trainingen in leerstrategieën (zie ook Entwistle et al., 1989; Volet et al., 1994).

Alle variabelen uit het model zijn gedefinieerd op het niveau van een specifieke onderwijsleereenheid, behalve de habituele leerstijl. In het model gaat het om de leerstijl waarmee de student het onderwijsonderdeel binnen komt, zoals die bijvoorbeeld gemeten kan worden met de ILS (Inventaris LeerStijlen) van Vermunt. Het onderzoek naar leeromgevingsfactoren en naar de gevolgde leerstrategie willen wij bewust niet beperken tot het niveau van grotere cursusgehelen (bijvoorbeeld een heel propedeusejaar), waar de ILS zich op richt. Wij willen ook de wijze van leren op kleinere aggregatieniveaus meten: op moduulniveau, zoals in dit artikel, en

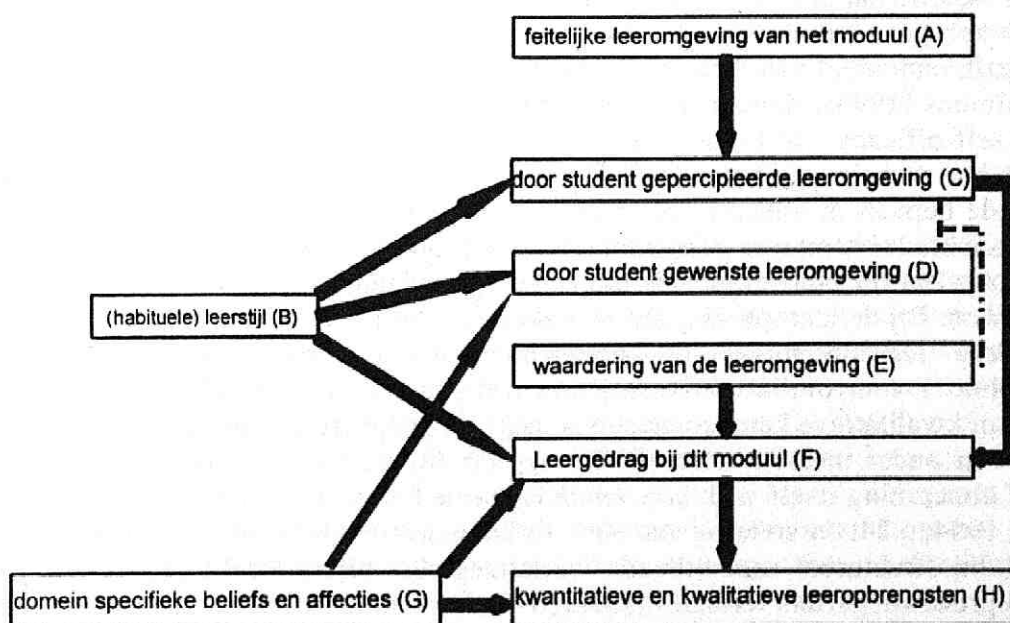


Fig. 1. Conceptueel model voor de interactie van leerstijl en leeromgeving binnen een onderwijsleereenheid.

nog kleinere eenheden, zoals in het onderzoek van Beerends (1994): hoe een student een artikel bestudeert of de stof van een college verwerkt. Daarmee wordt tevens een koppeling gelegd met het oudere leerstijlonderzoek eind jaren 70 (o.a. Marton & Säljö, 1976a,b) dat zich vooral op het niveau van kleine leereenheden afspeelde, veelal het bestuderen van teksten.

Het interactionele moment in het model betreft het totale traject tussen B en F, inclusief de indirecte paden. Wij veronderstellen dat dit traject verloopt via de in het model afgebeelde processen, dus onder andere ook via de perceptie en de waardering van de modulaire leeromgeving. Bij het complexe traject van B naar F kan ook de relatie van B naar C een rol spelen. Een voorbeeld van een B-C relatie doet zich voor wanneer de opdracht 'bestudeer voor het tentamen hoofdstuk 5' door reproductiegerichte studenten geïnterpreteerd wordt als: 'hoofdstuk 5 tot in details met letterlijke definities kunnen reproduceren', en door anderen (met een meer betekenisgerichte leerstijl) als: 'in eigen woorden het betoog van de auteur weer kunnen geven en kritisch kunnen bespreken' (Vermunt, 1992; Ramsden, 1992; Nuy & Gijselaers, 1993).

In het model komen diverse typen leeromgevingsvariabelen aan de orde. Allereerst de *feitelijke* leeromgeving (A) van het moduul, zoals die door de onderwijsinstelling, docent en onderwijsleersituatie (leerstof, werkvormen en leermiddelen) is vorm gegeven. De feitelijke leeromgeving komt overeen met wat Murray (1938) de objectieve ecologische aspecten van de omgeving noemt ('alpha press') en Brunswik (1955) 'distale' gebeurtenissen. De andere leeromgevingsvariabelen maken deel uit van wat Boekaerts (o.a. Boekaerts, 1983, 1991, 1992; Boekaerts & Simons, 1993) 'situatiebeleving' (in het engels 'appraisals') noemt. Bij C gaat het er om hoe deze leeromgeving door de student *gepercipieerd* wordt (Murray spreekt over 'beta press' en Brunswik over 'proximale' gebeurtenissen). Bij D gaat het om de *gewenste* leeromgeving en bij E om de *waardering* van de leeromgeving, dat wil zeggen om de discrepantie tussen C en D (aangegeven door de gestippelde lijnen). Van de vier hier genoemde leeromgevingsfactoren hebben volgens het model de leeromgevingsperceptie en de leeromgevingswaardering invloed op de actuele leerstrategie (het leergedrag).

We gaan nu nader in op de variabelen F, G en H. Onder het *leergedrag* (F) verstaan wij een typerende combinatie van leeractiviteiten die de student bij een bepaalde leertaak volgt om zijn leerdoel te bereiken (Shuell, 1988). Bij G gaat het om *domeinspecifieke opvattingen* ('beliefs') en *affecties* (McLeod, 1990; Schoenfeld, 1988). Als voorbeeld nemen wij een methodologiemoduul waarbij wij aspecten van G gemeten hebben (Beerends, 1994). Het gaat hier om denkbeelden (inclusief misconcepties) en gevoelens als: 'kennis van statistiek lijkt me belangrijk voor een sociale wetenschapper'; een goede statisticus lost elk statistiekprobleem in een paar minuten op'; 'ik heb geen enkele aanleg voor statistiek'. Zoals de laatste uitspraak aangeeft, wordt een belangrijk onderdeel van de domeinspecifieke beliefs en affecties gevormd door wat Boekaerts & Simons (1993) 'domeinspecifiek zelfvertrouwen' noemen en Wood & Locke (1987) 'academic self-efficacy'. In Figuur 1 is een pijl getrokken van G naar F. Dat wil bijvoorbeeld zeggen dat het domeinspecifieke zelfvertrouwen van een student met een betekenisgerichte leerstijl mede bepaalt in hoeverre bij deze student tijdens het moduul ook daadwerkelijk een betekenisgerichte leerstrategie geactualiseerd wordt en niet een meer reproductiegerichte strategie. De kans dat dit laatste gebeurt is groter bij studenten met een gering domeinspecifiek zelfvertrouwen. Bij de leeropbrengsten H maken we een onderscheid tussen "*kwalitatieve*" en "*kwantitatieve*" leeropbrengsten. We knopen hier aan bij de terminologie van Dahlgren (1984). Dahlgren pleit er voor om in de toetsing van wat een student geleerd heeft meer aandacht te schenken aan kwalitatieve kennisconcepties naast de gangbare kwantitatieve concepties. In het voetspoor van onder meer Bartlett (1932) typeert hij de kwalitatieve benadering als volgt: 'Instead of concerning itself with *how much* is learned, it seeks to investigate *what* is learned' (Dahlgren, 1984, p.24; cursivering van ons). Er kunnen structureel verschillende leeruitkomsten voorkomen bij structureel verschillende leerstrategieën: bijvoorbeeld kennis van geïsoleerde feiten bij reproductiegericht lerende studenten versus kennis van betekenisvolle relaties, of van strategieën bij betekenisgericht lerende studenten (Marton & Säljö, 1976a,b; Van Rossum & Schenk, 1984; Trigwell & Prosser, 1991; Beerends, 1994). Beerends (1994) laat in haar onderzoek naar F-G-H relaties zien hoe informatie over de *aard* van de leeropbrengsten wel aan het

licht komt bij een kwalitatieve meting, maar niet bij een meer conventionele toets. Zij vraagt de studenten in een essay kort uiteen te zetten "waar een college over testbetrouwbaarheid over ging" en analyseert dit op kwalitatieve wijze (de fenomenografische methode, zie Marton, 1986).

3. PROBLEEMSTELLING EN OPZET VAN EEN DEELONDERZOEK NAAR LEERSTIJL – LEEROMGEVING RELATIES

Het deelonderzoek dat in dit artikel beschreven wordt exploreert de relaties B t/m F. Een belangrijk doel is daarnaast om een aanzet te leveren tot de constructie en validatie van een instrument om procesaspecten van de leeromgeving te meten die rechtstreeks te maken hebben (zie paragraaf 3.2.2) met de cognitieve verwerkingsactiviteiten uit de ILS. Dit instrument, de IPAL (Inventaris van ProcesAspecten van de Leeromgeving) is een vragenlijst voor meting van de variabelen C, D en E.

3.1. Onderzoeksvragen

De volgende onderzoeksvragen zijn geformuleerd:

- 1a. Welke samenhangen zijn er in het begin van het propedeusejaar tussen de verwerkings- en regulatiestrategieën die studenten sociale wetenschappen gebruiken, hun mentale leermodellen en hun leeroriëntaties?
- b. Welke leerstijlen treffen we in het begin van de propedeuse aan bij studenten sociale wetenschappen?
2. In hoeverre is het mogelijk met een leeromgevingsvragenlijst (de IPAL) student-percepties van de leeromgeving te meten, die corresponderen met de cognitieve verwerkingsactiviteiten uit de ILS? Is het bijvoorbeeld mogelijk een leeromgeving door middel van de IPAL als reproductiegericht te typeren? Is het mogelijk met de IPAL tot een inhoudelijk goed interpreteerbare typologie van leeromgevingspercepties te komen?
3. Zijn er onderlinge relaties tussen de volgende typen van variabelen: de leerstijl van een student (voorafgaande aan een moduul gemeten via de ILS), zijn beleving van de leeromgeving in het moduul en zijn leergedrag tijdens het moduul?

De onderzoeksvragen in dit deelonderzoek hebben dus geen betrekking op de variabelen G en H (zie daarvoor Wierstra & Van der Linden, 1996). Bij de onderzoeksvragen 1 en 2 merken we op dat het ons meer gaat om de detectie van typologieën en dimensies dan om precieze kwantificeringen hoe vaak bepaalde leerstijlen of leeromgevingsstijlen voorkomen.

3.2. Methode

Het onderzoek is uitgevoerd bij een moduul (begrote studietijd, inclusief onderwijs 160 uur) uit de algemene propedeuse voor studenten sociale wetenschappen aan de Universiteit van Utrecht. Het moduul "Inleiding in de veldoriëntatie" geeft voor alle studenten in een van de zes studierichtingen een inleiding in de studie van een bepaald maatschappelijk veld (waarop meerdere sociaalwetenschappelijke disciplines werkzaam zijn). Aan het onderzoek namen studenten uit drie velden deel: 'onderwijs en opvoeding', 'gezondheidszorg' en 'leefvormen en socialisatie'. Er namen in totaal 210 studenten deel aan het onderzoek in een of meerdere vragenlijstrondes. Deelname van de studenten aan het onderzoek gebeurde op basis van vrijwilligheid. Om hen tot deelname te motiveren werd het belang van het onderzoek voor herziening van het moduul benadrukt. Ook werd de vertrouwelijkheid van de informatie gegarandeerd. De studenten hoefden niet hun naam op het antwoordformulier te schrijven, maar wel hun studentnummer om koppeling van bestanden in het onderzoek mogelijk te maken. De betrokken docenten kregen geen inzage in de individuele gegevens van studenten.

3.2.1. De leerstijlmeting

Een van de verschillen met de oudere definities van leerstijl (en cognitieve stijl, zie bijvoorbeeld Bakker, 1991) is dat een leerstijl in ons onderzoek niet als een onveranderlijke dimensie gezien wordt (zie paragraaf 2). Een ander verschil is dat we niet van een bipolaire definitie uitgaan. Riding & Cheema (1991) zien een leerstijl als een prototypisch verschijnsel waarvan een student meer of minder kenmerken kan vertonen. Karakteristiek voor de in ons onderzoek gehanteerde definitie van leerstijl is voorts dat deze breder is dan gebruikelijk. Wij sluiten ons aan bij de opvatting van Vermunt (1992), welke zich ongeveer als volgt laat verwoorden: een leerstijl is een samenhangend geheel van voor de student (in een bepaalde periode) karakteristieke *verwerkingsactiviteiten*, een karakteristieke wijze van *metacognitieve regulatie* (activiteiten als oriënteren, plannen, toetsen, bijstellen), een karakteristieke *visie op leren en onderwijzen* (mentale modellen) en karakteristieke *studiemotieven* (leeroriëntaties).

De breedheid van Vermunts leerstijlbegrip zit onder andere hierin dat men zich meestal in de definitie van leerstijlen beperkt tot de verwerkingsactiviteiten en eventueel de mentale modellen. De studiemotieven vallen er gewoonlijk niet onder. De vier componenten zijn in de ILS gerepresenteerd in 120 items die zijn ondergebracht in 20 schalen. Dit stelt ons bij het uitvoeren van de analyses voor het probleem hoe we met de 20 schalen om moeten gaan: uitgaan van 20 scores of een of meer samengestelde maten construeren? In paragraaf 4.1 gaan we daar nader op in.

3.2.2. De leeromgevingsmeting

Een manier om (aspecten van) de feitelijke leeromgeving A vast te stellen is om deze door onafhankelijke observatoren te laten scoren op de onderscheiden aspecten. In enkele deelonderzoeken hebben we dit gedaan (zie bijvoorbeeld Visser, 1994). In het onderhavige onderzoek waren we primair in de belevingen van de studenten geïnteresseerd en hebben daarom van de leeromgeving alleen de subjectieve variabelen C, D en E expliciet gemeten. Daartoe is een vragenlijst geconstrueerd, de IPAL: Inventarisatie van ProcesAspecten van de Leeromgeving. De IPAL richt zich op de cognitieve verwerkingsactiviteiten en metacognitieve regulatieactiviteiten, die in de ILS aan bod komen. De IPAL bestaat uit 120 items van vier verschillende typen (zie tabel 1). Er zijn 40 items over het beeld dat de student heeft van wat er van hem verwacht wordt aan studiegedrag (type item 38), voorts 28 perceptie-items over het beeld dat de student heeft van de onderwijssituatie (type item 50a) en 12 items over gedrag waartoe de student 'uitgenodigd' werd (type 78a). Bij de perceptie-items en de uitnodigings-items werden parallelle wens-items geconstrueerd (type 50b en 78b).

Tabel 1. Voorbeelditems (5-puntsschalen) uit ILS en IPAL op het gebied van concretiseren.

ILS

52. Ik denk aan zaken die ik ken uit eigen ervaring die samenhangen met de studiestof die ik bestudeer. (*leerstijl-item*)

IPAL

38. In dit moduul wordt van de student verwacht dat hij bij het bestuderen van de studiestof denkt aan dingen uit eigen ervaring die hiermee samenhangen. (*verwachtings-item*)
- 50a. In dit moduul wordt de aangeboden kennis gekoppeld aan gebeurtenissen uit de dagelijkse werkelijkheid (*perceptie-item*)
- b. Ik zou voor mezelf willen dat dit in het moduul aandacht kreeg. (*wens-item*)
- 78a. In dit moduul wordt je uitgenodigd de studiestof te koppelen aan eigen ervaringen. (*uitnodigings-item*)
- b. Ik zou voor mijzelf willen dat dit in het moduul aandacht kreeg. (*wens-item*)

Combinatie van deze vraagtypen werd geacht een indicatie te geven in hoeverre de leeromgeving – in de terminologie van Vermunt (1992, hoofdstuk 2) – “activerend”, “substituerend” en “capitaliserend” is met betrekking tot de cognitieve activiteit concretiseren. Om een enigszins gesimplificeerd voorbeeld te geven: een positieve beantwoording van items van het type 78a verwijst naar een leeromgeving die *activerend* is met betrekking tot concretiseren; een positieve beantwoording van items van het type 50a in combinatie met een negatieve beantwoording van items van het type 38 en 78a indiceert een leeromgeving die *substituerend* is; een positieve beantwoording van items van het type 38 en een negatieve beantwoording van de andere itemtypen gaat in de richting van een leeromgeving die *capitaliseert* met betrekking tot concretiseren. Een negatieve beantwoording van al de genoemde items wijst op een leeromgeving die de instructie-leer-functie ‘concretiseren’ *negeert*.

Voor de ‘verwachtings’items (itemtype 38) zijn geen wensitems (b-items) geconstrueerd, omdat een proefafname uitwees dat dit voor de student een te complex item op zou leveren. Bij de overige itemtypen bleken de b-items goed te voldoen. Ook in eerdere onderzoeken met soortgelijke leeromgevingsschalen bleken b-items aan hun doel te voldoen en goed te worden onderscheiden van de a-items (Fraser, 1986; Wierstra, Jörg & Wubbels, 1987).

3.2.3. De meting van het leergedrag

We vroegen de studenten naar hun leergedrag door middel van een vragenlijst van 40 items: de ASA (vragenlijst voor geActualiseerde Studie- Activiteiten: actualisatie in de zin waarin die term voorkomt in een organisch interactiemodel, zie par. 2). Evenals de IPAL-items hebben deze items uitsluitend betrekking op het moduul en zijn ze eveneens geïnspireerd door de ILS. Een voorbeeld is het volgende 5-puntsitem: ‘De stof uit de reader bij dit moduul bracht ik tijdens het studeren in verband met dingen die ik zelf ervaren heb’.

Afname instrumenten

De ILS is in het eerste en tweede college van het moduul aan de studenten meegegeven om thuis in te vullen. De IPAL werd na driekwart van het moduul aan de studenten afgenomen. De ASA werd aan het eind van het moduul afgenomen. Door omstandigheden was er helaas slechts bij één van de drie docenten gelegenheid om de ASA aan de studenten voor te leggen.

4. RESULTATEN

4.1. Analyses met betrekking tot onderzoeksvraag 1

Factoranalyse ILS-schalen

We voerden een hoofdcomponentenanalyse uit op achttien van de twintig schalen (van de schalen zelfsturing en externe sturing namen we de integrale ongedeelde schalen op). Er werden zes componenten gevonden met een eigenwaarde groter dan 1. Varimax rotatie over verschillende aantallen factoren (tussen 2 en 6) gaf de duidelijkste oplossing voor vier factoren. Deze oplossing, die analoog is aan de oplossing van Vermunt (1992, p. 116,117), is weergegeven in Tabel 2. De eerste en de derde factor zijn naar ons idee het duidelijkst te interpreteren. De achtereenvolgende factoren kunnen benoemd worden als: 1 *betekenisgericht* (= actief zelfsturend) *leren*, 2 *een ongerichte wijze van leren*, 3 *reproductiegericht* (= passief, extern gestuurd) *leren* en 4 *toepassingsgericht* (*beroepsgericht*) *leren*. Deze factoren verklaren tezamen 55.4% van de totale variantie op de schalen.

De hier verrichte (R-)factoranalyse heeft betrekking op een groepering van *variabelen* (schalen). Het is dan ook niet zonder meer duidelijk in hoeverre met deze factoren (bijvoorbeeld met de factoren ‘betekenisgericht leren’ en ‘reproductiegericht leren’) verschillende groepen *personen* corresponderen. Ten behoeve van onderzoeksvraag 1b hebben we dan ook een clusteranalyse op personen uitgevoerd. Om de clusters te kunnen interpreteren, onderwerpen we deze tenslotte aan een discriminantanalyse.

Tabel 2. Factormatrix ILS-schalen bij varimax rotatie over 4 factoren.

	Factor 1	Factor 2	Factor 3	Factor 4
Relateren & structureren	.85			
Kritisch verwerken	.69			
Concrete verwerking	.73			.33
Memoriseren & herhalen		.71		
Analyseren	.36		.70	
Stuurloos		.68		
Zelfsturing	.77			
Externe sturing			.49	.52
Opname van kennis		.39	.66	
Opbouw van kennis	.54			.34
Gebruik van kennis				.79
Stimulerend onderwijs		.59		
Samen studeren		.56		
Certificaatgericht		.31	.55	
Beroepsgericht				.76
Testgericht		.60		
Persoonlijk geïnteresseerd	.54			-.31
Ambivalent	-.30	.65		

Noot. Ladingen < .25 of > -.25 zijn weggelaten.

Clusteranalyse en discriminantanalyse ILS-schalen

Op de achttien ILS-schalen uit tabel 2 hebben we een hiërarchische clusteranalyse uitgevoerd via de samenvoegmethode. De studenten werden op basis van de vorm van het profiel van achttien schaalesscores (gestandaardiseerde scores) in stappen bij elkaar gevoegd tot clusters. Voor het samenvoegen van clusters kwamen op basis van de literatuur (o.a. Everitt, 1980; Milligan & Cooper, 1987) twee methoden in aanmerking: 'complete linkage' (furthest neighbour) en 'between groups average'. We hebben beide methoden toegepast en verkregen vrijwel identieke resultaten. In de weergave van de resultaten zullen we ons daarom tot één methode beperken: de between group average methode.

In eerste instantie hebben we ons gericht op een oplossing met vier clusters, omdat Vermunt (1992) op grond van fenomenografische analyses voor universitaire studenten vier leerstijltypen suggereert. We verkregen vier clusters met respectievelijk 30%, 15%, 25% en 30% van de studenten. Om de clusters te kunnen interpreteren voerden we vervolgens een discriminantanalyse uit. Dit betekent dat we nagaan welke van de achttien schalen het grootste gewicht hebben in het discrimineren tussen de clusters. Deze gewichten zijn de structuurcoëfficiënten in de discriminantfuncties. Het bleek dat twee van de vier clusters studenten nauwelijks van elkaar te onderscheiden waren, zodat het meer voor de hand lag om uit te gaan van drie clusters en daarop een discriminantanalyse uit te voeren. De twee discriminantfuncties (eigenwaarden 2.10 en 1.64) voor de drie-cluster oplossing bleken voor 97% van de studenten te kunnen voorspellen in welk clusters de student was ingedeeld en ze waren goed interpreteerbaar. Voor de interpretatie wendden we ons tot de structuurcoëfficiënten (correlaties tussen een discriminantfunctie en een schaal). Deze zijn weergegeven in Tabel 3. Zoals een factor in de factoranalyse een kunstmatige (latente) variabele is die een aantal manifeste variabelen overkoepelt, zo is ook een discriminantfunctie een latente variabele. De structuurcoëfficiënten zijn dan ook te vergelijken met factorladingen en helpen ons de discriminantfuncties te interpreteren.

Op grond van Tabel 3 veronderstellen we dat de eerste discriminantfunctie genterpreteerd kan worden als '(mate van) extern gestuurd reproductief leren' en de tweede als '(mate van) zelfgestuurd betekenisgericht leren'. Kijken we naar het scatterplot van de twee dimensies (Figuur 2)

Tabel 3. Structuurcoëfficiënten voor de ILS-discriminantanalyse bij drie clusters.

ILS Schaal	Discriminantfunctie 1	Discriminantfunctie 2
Externe sturing	.45*	-.07
Analyseren	.37*	.23
Opname van kennis	.35*	-.18
Memoriseren & herhalen	.35*	.23
Stimulerend onderwijs	.34*	-.09
Gebruik van kennis	.32*	.09
Beroepsgericht	.25*	.03
Samen studeren	.24*	.00
Certificaatgericht	.20*	-.13
Testgericht	.11*	.05
Relateren & structureren	-.11	.62*
Kritisch verwerken	.00	.51*
Concrete verwerking	.07	.41*
Zelfsturing	-.03	.39*
Persoonlijk geïnteresseerd	-.05	.38*
Ambivalent	.06	-.31*
Opbouw van kennis	.11	.20*
Stuurloos	.15	-.17*

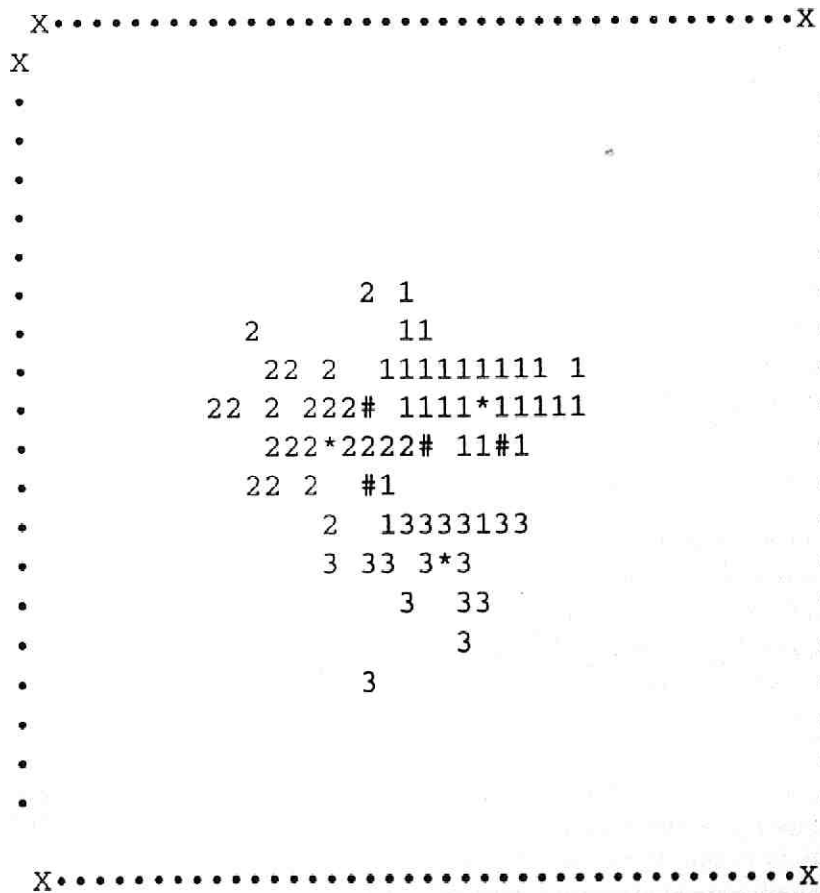


Fig. 2. Plot ILS-discriminantanalyse voor drie clusters. De horizontale as is discriminant functie 1 ('extern gestuurd reproductief leren') en de verticale as is de discriminant functie 2 ('zelfgestuurd betekenisgericht leren').

dan blijkt dat de clusters 2 en 3 in zekere zin polair tegenover elkaar staan op beide dimensies. De positie van de vrij grote groep 1 onderstreept echter dat we – zoals in paragraaf 3.2.1. betoogd is – leerstijlen niet als polaire tegenstellingen moeten zien. In die zin is groep 1 op te vatten als een groep met een flexibele, veelzijdige leerstijl.

De twee dimensies vertonen overeenkomst met de bekende onderscheiding oppervlaktegerichte versus dieptegegerichte manier van leren (o.a. Marton & Saljö, 1976a,b; Biggs & Rihn, 1984), maar voegen daar het onderscheid interne versus externe sturing aan toe. Dit resultaat bevestigt de zienswijze van Vermunt (1992) om in de leerstijldefinitie een centrale rol toe te kennen aan sturingsaspecten (metacognitieve regulatie, zie paragraaf 3.2.1). De orthogonaliteit van de twee dimensies (discriminantfuncties) en de positie van leerstijlcluster 1 maakt duidelijk dat het onderscheid tussen de twee wijzen van leren niet – zoals vaak gedacht wordt – als een continuüm moet worden opgevat, maar dat het om twee onderscheiden dimensies gaat.

Samenhangen tussen factoren en discriminantfuncties

Bovengenoemde bidimensionaliteit wordt ook al door de factoranalyse resultaten (Tabel 2) gesuggereerd, gezien de orthogonaliteit van de factoren 1 en 3. Een vergelijking van de Tabellen 2 en 3 geeft dan ook de indruk dat er verwantschap is tussen de factoren en de discriminantfuncties. Dit wordt bevestigd als we correlaties berekenen. De tweede discriminantfunctie overlapt grotendeels met de eerste factor, zoals blijkt uit de Pearson product-moment correlatiecoëfficiënt ($r=.85$). De eerste discriminantfunctie hangt sterk samen met de factoren 3, 4 en 2 (de multiële correlatie bedraagt .85 met betagewichten .54, .52 en .35).

ILS PROFIEL

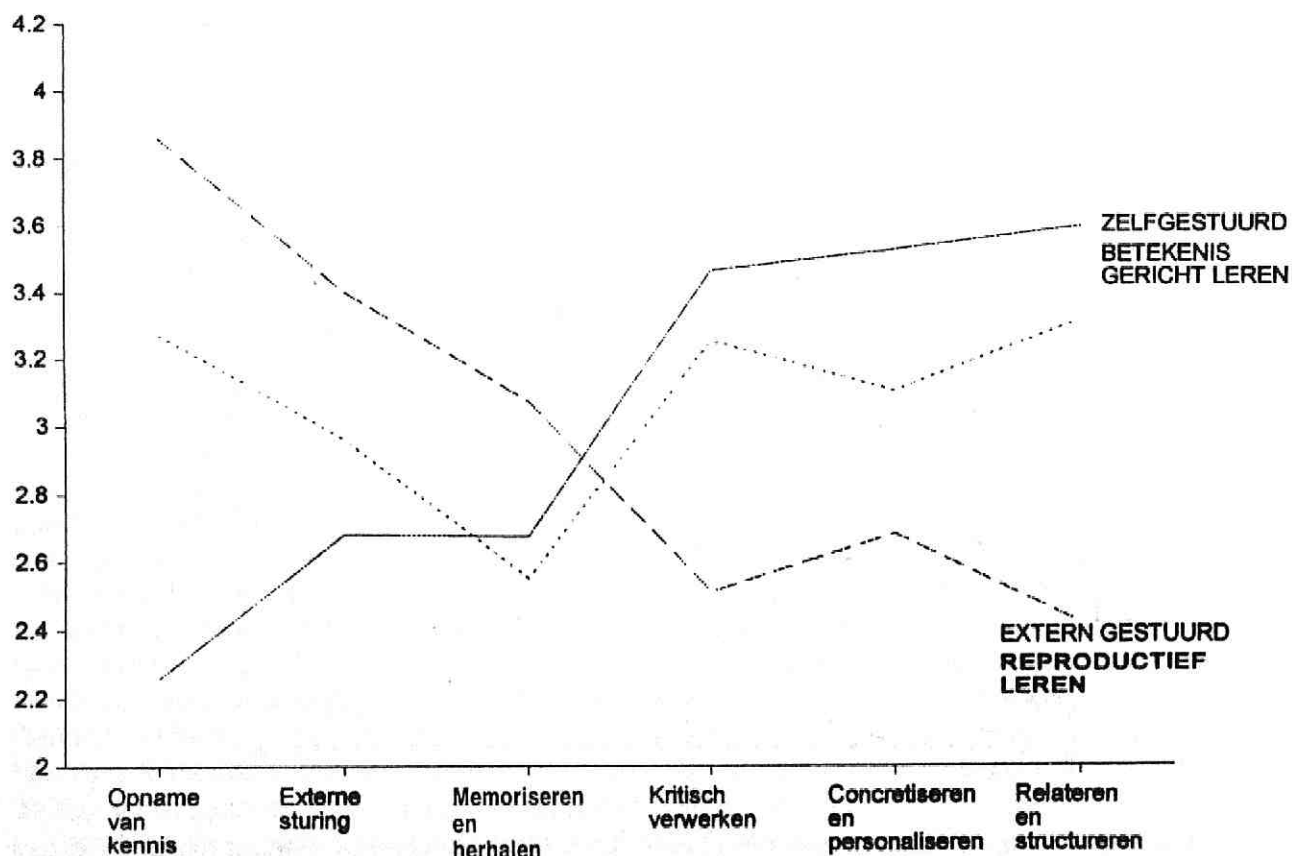


Fig. 3. Profielen van drie ILS-clusters.

ILS clusterprofielen

Ter illustratie geven we in Figuur 3 de profielen van de drie groepen op een zestal schalen die in de factor- en discriminantanalyses grote ladingen hebben, factor- en discriminant-'zuiver' zijn en uit variantie-analyses als sterk differentiërend naar voren komen. Op de verticale as is voor ieder cluster de gemiddelde score per item weergegeven.

Conclusie met betrekking tot onderzoeksvraag 1

Het is mogelijk de studenten in te delen in drie leerstijlclusters die verschillen al naar gelang hun positie op de variabelen 'extern gestuurd reproductief leren' en 'zelfgestuurd betekenisgericht leren': twee clusters die elkaars tegenpolen zijn en een cluster met een flexibele leerstijl. Er werd geen aanwijzing gevonden dat 'toepassingsgericht leren' als aparte variabele een belangrijke typologie-dimensie is voor deze groep studenten. In vervolgonderzoek moet blijken hoe representatief deze stijlen zijn voor alle propedeusestudenten. De factoranalyse geeft nageenog dezelfde samenhangen te zien als door Vermunt (1992) gerapporteerd zijn; de in factor- en clusteranalyse gevonden vervlechting van cognitieve en affectieve variabelen geeft ondersteuning van de gedachte (o.m. Biggs & Rihn, 1984; Tait & Entwistle, 1996) dat motivaties en intenties van de student van grote invloed zijn op de (meta)cognitieve strategie die hij inzet. De resultaten van factor- en clusteranalyse ondersteunen voorts dat het bij de twee leerstijldimensies om onafhankelijke variabelen gaat.

4.2 Analyses met betrekking tot onderzoeksvraag 2*Homogeniteit IPAL-schalen voor gepercipieerde en gewenste leeromgeving*

We onderwierpen de perceptie-items aan een factoranalyse en berekenden Cronbach alpha-coëfficiënten voor deelverzamelingen van factorzuivere items. Dit leidde tot vorming van de volgende zeven schalen voor de perceptie van de leeromgeving: reproductiegericht, externe sturing, samenhanggericht, praktijkgericht, elaboratiegericht, kritisch gericht en brongericht. Zie Tabel 4 voor de betekenis van de schalen, de Cronbach alpha's, en enkele voorbeelditems.

De 'uitnodigingsitems' (zie Tabel 1) kwamen vrijwel alle terecht in de schaal 'elaboratiegerichte' leeromgeving, een schaal die aangeeft in welke mate de leeromgeving diepteverwerking en concrete verwerking van de leerstof activeert. De overige schalen bevatten steeds de beide andere itemtypen uit Tabel 1. Het bleek niet mogelijk een factorzuivere meting van een substituerende of capitaliserende leeromgeving te realiseren. Nadat de perceptieschalen gevormd waren, vormden we vijf schalen voor de gewenste leeromgeving, op basis van dezelfde items (het met het a-item corresponderende b-item)³. Het aantal items en de alpha-coëfficiënten van deze schalen zijn eveneens in tabel 4 aangegeven.

Factoranalyse IPAL perceptie-schalen

Een hoofdcomponentenanalyse van de zeven perceptieschalen leverde twee componenten op met een eigenwaarde groter dan 1. Varimax rotatie over twee factoren leverde een tweetal goed interpreteerbare factoren op, die tezamen 53.2% van de variantie verklaren. Zoals uit Tabel 5 blijkt kan de eerste factor geïnterpreteerd worden als het (kennis)activeringsaspect van de leeromgeving en de tweede factor, gerepresenteerd door de schalen 1 en 2, als het (kennis)-overdrachtsaspect van de leeromgeving.

Clusteranalyse en discriminantanalyse IPAL perceptieschalen

Analoog aan de procedures voor de ILS voerden we een clusteranalyse uit op de IPAL-schalen, gevolgd door een discriminantanalyse voor de driecluster-oplossing. Op basis van de twee discriminantfuncties (eigenwaarden 2.65 en .64) kon 95 % van de studenten in het juiste cluster worden ingedeeld. De eerste (en meest belangrijke) discriminantfunctie lijkt op basis van de discriminantmatrix het meest duidelijk interpreteerbaar. Er is een grote overeenkomst met IPAL-factor 1, waarmee deze discriminantfunctie .90 correleert. Discriminantfunctie 2 is minder eenduidig en correleert .42 met factor 2. Gezien de discriminantmatrix en de correlaties tussen

Tabel 4. IPAL-schalen: namen, voorbeelditems en Cronbach Alpha.

Schaal	Aantal items	Voorbeelditem	Cronbach alpha
1a. Reproductiegerichte leeromgeving	14	In dit moduul wordt van de student verwacht dat hij rijtjes met kenmerken uit zijn hoofd kent	.82
b. Wens naar reproductiegerichte leeromgeving	6	Ik zou voor mijzelf willen dat het inprenten van begrippen in je hoofd in dit moduul aandacht krijgt	.77
2a. Sturende leeromgeving	9	In dit moduul wordt een strikte volgorde gehanteerd bij de behandeling van de stof	.65
b. Wens naar externe sturing	9	Ik zou voor mijzelf willen dat de stof precies zo behandeld werd als in het boek	.77
3a. Samenhanggerichte leeromgeving	13	In dit moduul werden de afzonderlijke onderwerpen samengebracht tot een geheel	.80
b. Wens naar samenhanggerichte leeromgeving	7	Ik zou voor mijzelf willen dat conclusies uit verschillende hoofdstukken in dit moduul met elkaar vergeleken worden	.76
4a. Praktijkgerichte leeromgeving	10	In dit moduul werden aan de hand van de theorie oplossingen gezocht voor praktische problemen	.72
b. Wens naar praktijkgerichte leeromgeving	4	Ik zou voor mijzelf willen dat de stof in dit moduul zodanig wordt aangeboden dat ik die kan gebruiken bij activiteiten buiten de studie	.72
5a. Elaboratiegerichte leeromgeving	9	In dit moduul word je uitgenodigd de studiestof te koppelen aan eigen ervaringen	.76
b. Wens naar elaboratiegerichte leeromgeving	9	Ik zou voor mijzelf willen dat je in dit moduul uitgenodigd wordt eigen conclusies te trekken	.76
6. Verwijzing van leeromgeving naar <i>andere bronnen</i>	4	In dit moduul wordt andere, niet verplichte, literatuur aanbevolen	.73
7. Leeromgeving die zich richt op <i>kritisch verwerken</i>	4	In dit moduul wordt van de student verwacht dat hij de interpretaties van deskundigen kritisch benadert	.55

discriminantfuncties en factoren, kan de eerste discriminantfunctie geïnterpreteerd worden als 'activerend onderwijs' en de tweede discriminantfunctie als 'kennisoverdragend onderwijs' (waarbij wel aandacht is voor bronverwijzing, maar met als doel primair kennisuitbreiding en niet zozeer verdieping).

IPAL clusterprofielen

Figuur 4 toont een profielschets voor de drie clusters. Hierin is de schaal externe sturing buiten beschouwing gelaten, omdat variantie-analyse uitwees dat de clusters hierop niet significant

Tabel 5. Factormatrix IPAL-perceptie schalen bij varimax rotatie over 2 factoren en Discriminant matrix bij twee discriminantfuncties.

	Factormatrix		Discriminantmatrix	
	Factor 1	Factor 2	Discriminantfunctie 1	Discriminantfunctie 2
Kritisch verwerken	.66	.17	.55*	.03
Samenhanggericht	.71	.26	.49*	.04
Praktijkgericht	.77	-.14	.43*	.09
Elaboratiegericht	.72	-.30	.40*	-.25
Brongericht	.53	.09	.13	.43*
Reproductiegericht	-.08	.77	-.01	.62*
Extern sturend	.16	.69	-.02	-.15*

verschillen. De twee bovenste profielen zijn van clusters die zich tamelijk tegengesteld gedragen op de dimensies activerend en kennisoverdragend onderwijs.

Conclusie met betrekking tot onderzoeksvraag 2

Met de IPAL kunnen aspecten van de leeromgeving gemeten worden, die betrekking hebben op de cognitieve verwerkingsactiviteiten die in de ILS gemeten worden. Er werden drie typen leeromgevingspercepties gevonden. Het bleek mogelijk om de drie typen te karakteriseren

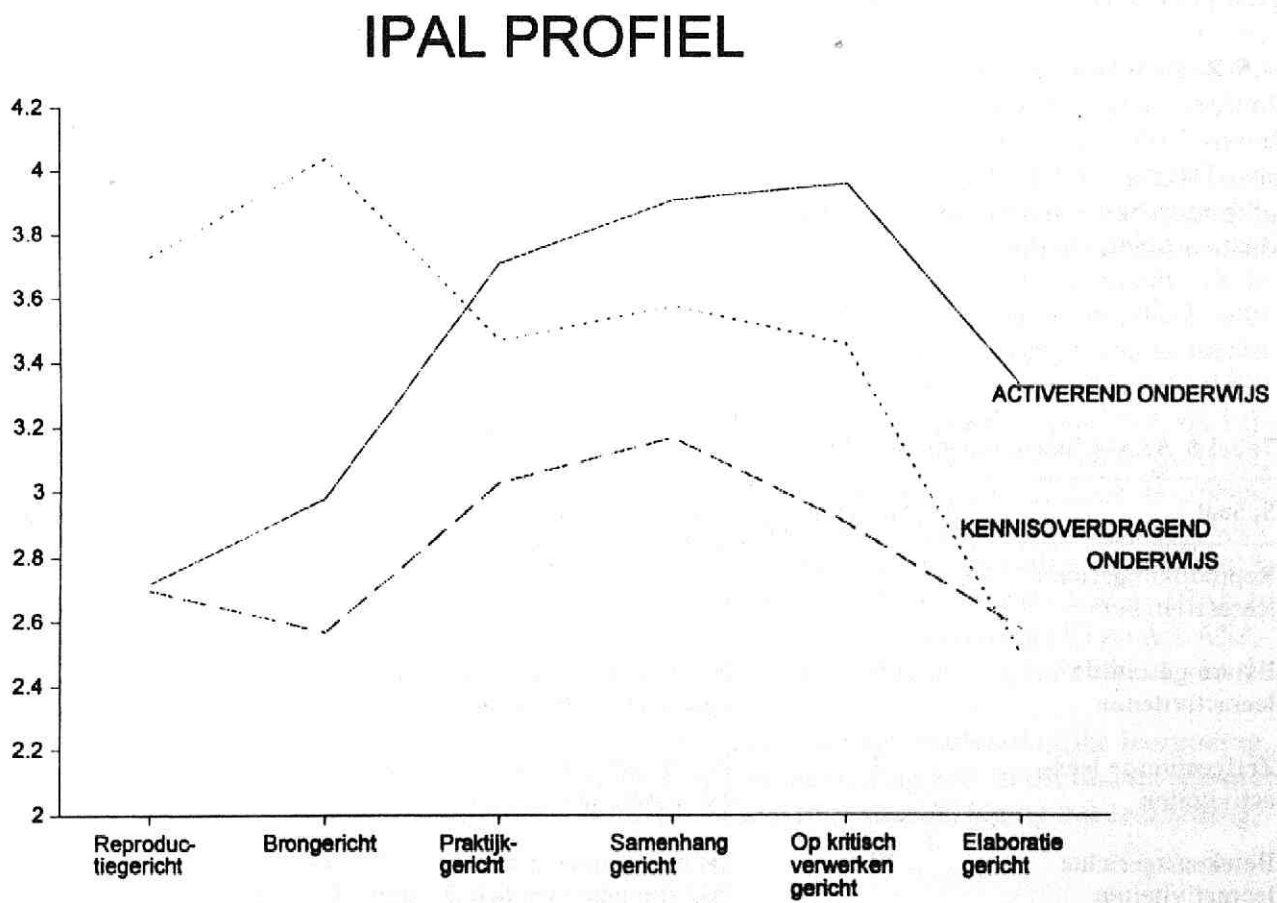


Fig. 4. Profielen van drie IPAL-clusters.

vanuit dimensies die met actief zelfgestuurd leren dan wel passief extern gestuurd leren te maken hebben. Ook hier gaat het evenals bij de leerstijlschalen om onafhankelijke dimensies.

4.3 De meting van de leeromgevingswaardering

Een methodologische voorwaarde voor het valide meten van de leeromgevingswaardering door middel van het (absolute) verschil van perceptie- en wens-score is dat de perceptie- en wensmeting als onafhankelijke metingen kunnen worden opgevat. Dit betekent onder meer dat de leeromgevingspercepties niet negatief mogen correleren met de leeromgevingswensen, omdat dit een indicatie is dat er een invloed uitgaat van wat de persoon wenst op wat hij waarneemt (Wierstra, Jörg & Wubbels, 1987). Zo zal een negatieve correlatie tussen de behoefte aan externe sturing van een student en de door hem gepercipieerde sturing als volgt genterpreteerd kunnen worden: *meer* behoefte aan externe sturing maakt dat je *minder* externe sturing ervaart (of althans verwoordt dat je dat minder ervaart). De perceptie- en de wensschalen zouden dan niet onafhankelijk van elkaar gemeten zijn. In ons onderzoek is dat niet het geval: er worden geen negatieve correlaties gevonden tussen wens- en perceptieschalen, maar een enkele zwakke positieve correlatie. Doordat er nauwelijks (positieve of negatieve) correlaties gevonden worden, doet zich ook een ander methodologisch probleem niet voor, namelijk het probleem dat de betrouwbaarheid van een waarderingsscore (als verschilscore van corresponderende wens- en perceptiescore) laag is wanneer de correlatie tussen de afzonderlijke schalen hoog is in vergelijking met de betrouwbaarheid van de beide schalen (Drenth & Sijtsma, 1990).

4.4 De vorming van schalen voor de geActualiseerde Studie Activiteiten (ASA)

Op basis van een clusteranalyse op items, gevolgd door berekening van Cronbach alpha coëfficiënten konden we de volgende vier schalen vormen voor het leergedrag: reproductiegerichte leeractiviteiten, extern gestuurde leeractiviteiten, zelfgestuurde leeractiviteiten en betekenisgerichte leeractiviteiten. Zie voor de Cronbach alpha's en voor voorbeelditems Tabel 6.

4.5 Resultaten met betrekking tot onderzoeksvraag 3

In deze onderzoeksvraag gaat het om een aantal verbanden tussen ILS-, IPAL-, en ASA-variabelen. Deze verbanden zijn onderzocht via Pearson correlatie- en (multiple) regressie-analyses. Daarbij zijn de ILS- en de IPAL- variabelen op drie niveaus gescoord: als scores op de afzonderlijke schalen, en als scores op de in de voorgaande paragrafen behandelde factoren en discriminantfuncties.

Tabel 6. ASA-schalen: namen, voorbeelditems en Cronbach Alpha.

Schaal	Aantal items	Voorbeelditem	Cronbach alpha
Reproductiegerichte leeractiviteiten	7	De studiestof uit de reader heb ik zo letterlijk mogelijk geleerd	.93
Extern gestuurde leeractiviteiten	6	Ik heb de opgave uit de reader gemaakt tijdens het bestuderen van de stof	.93
Zelfgestuurde leeractiviteiten	7	Bij de verplichte studiestof heb ik zelf extra materiaal gezocht	.78
Betekenisgerichte leeractiviteiten	10	De opvattingen die in dit moduul aanbod zijn gekomen heb ik vergeleken om overeenkomsten en verschillen te ontdekken	.55

Verbanden tussen leerstijl- en leeromgevingsvariabelen

Tussen de leerstijlen en de leeromgevingspercepties blijken gematigde correlaties te bestaan (zowel op het niveau van de afzonderlijke subschalen als op het niveau van factoren en discriminantfuncties). Het meest opvallend was het positieve verband tussen de IPAL schaal waargenomen 'Samenhanggerichtheid' van de leeromgeving en de Betekenisgerichte leerstijlvariabelen, namelijk de ILS variabelen: Relateren & structureren (.40), Kritisch verwerken (.32), Concrete verwerking (.29), Zelfsturing (.31), ILS- factor 1 (.36) en ILS- discriminantfunctie 2 (.33). Alle correlaties hebben een tweezijdige overschrijdingskans van .001.

De relaties tussen leerstijl en leeromgevingswensen, zoals in figuur 1 gepostuleerd, bleken minder sterk aanwezig te zijn dan wij dachten. Het belangrijkste resultaat was dat de behoefte aan een leeromgeving met een sterk externe sturing vooral voorspeld wordt ($r=.41$, $p=.01$) door de score van de student op ILS discriminantfunctie 1, die een aanpak van extern gestuurd reproductief leren representeert. Voorts is vermeldenswaard en goed interpreteerbaar de correlatie .40 tussen de opvatting dat leren in de eerste plaats opbouw van kennis is (en niet zozeer opname van kennis) en de behoefte aan een elaboratiegerichte leeromgeving. Wellicht zouden wij een sterkere samenhang gevonden hebben tussen de leerstijl en de behoefte aan een bepaalde leeromgeving, als we gevraagd hadden naar de wensen ten aanzien van de leeromgeving *in het algemeen* en niet zoals in dit onderzoek naar de *specifieke* leeromgeving in dit moduul. De wensen ten aanzien van deze specifieke leeromgeving worden wellicht sterker bepaald door domeinspecifieke affecties (pijl van G naar D in Figuur 1).

Relaties die ILS- en IPAL-scores hebben met ASA-scores

We vinden hoge en significante correlaties ($p<.01$) tussen de overeenkomstige schalen van de ILS en de ASA. Zo correleert de ASA-schaal Reproductiegerichte leeractiviteiten respectievelijk .71 en .73 met de ILS-schaal Memoriseren & herhalen en Analyseren en .65 met ILS-factor 3, die een tendens tot reproductiegericht leren representeert. De ASA-schaal Betekenisgerichte leeractiviteiten correleert .65 met ILS-schaal Concrete verwerking en .62 met Kritische verwerking. De ASA schaal Zelfgestuurde leeractiviteiten correleert .70 met de ILS schaal Zelfsturing en eveneens .70 met ILS-factor 1, Betekenisgericht leren. Deze resultaten pleiten voor de validiteit van zowel de ILS (meting op cursusniveau) als de ASA (meting op moduulniveau). Correlaties tussen de IPAL-perceptie variabelen en de ASA-variabelen werden daarentegen veel minder aangetroffen. Dit kan veroorzaakt zijn doordat de afname van de ASA slechts plaats vond bij één onderwijsgroep, waardoor de leeromgevingsperceptie-scores van de studenten een lagere spreiding hadden dan wanneer er over alle onderwijsgroepen zou zijn gemeten. Er is daarnaast nog een andere verklaring denkbaar. Anders dan wij dachten blijkt de IPAL niet zozeer activerende, substituerende, en capitaliserende aspecten van de leeromgeving te meten (met uitzondering van activerende elaboratie, maar dit aspect wordt weer niet door de ASA gemeten), maar meer een soort totaalklimaat, namelijk in hoevere bepaalde aspecten als bijvoorbeeld relateren en structureren überhaupt in het moduul aan bod komen. Een hoge IPAL score op samenhanggerichtheid kan dus betekenen dat de samenhang vooral door de *docent* wordt aangegeven, hetgeen tot gevolg kan hebben dat de *student* er niet zelf actief naar op zoek gaat. Nader onderzoek met de IPAL en de ASA, waarbij ook observatoren worden ingeschakeld, is nodig om hier uitsluitsel over te krijgen. Wel vonden wij bij de ene docent waar de IPAL is afgenomen gematigde correlaties ($p<.05$) tussen de leeromgevingswaardering (E) en het ASA-leergedrag (F). Met name blijkt een grotere waardering van de samenhanggerichtheid van de leeromgeving te leiden tot sterker betekenisgericht leren.

Samenvattend: er is een duidelijke invloed van de leerstijl op het daadwerkelijke leergedrag. De directe invloed van de leeromgevingsperceptie op het leergedrag kon in dit onderzoek niet goed worden aangetoond, wel de invloed van de leeromgevingswaardering op het leergedrag.

5. CONCLUSIE EN DISCUSSIE

Met betrekking tot *onderzoeksvraag 1* wijst het onderzoek uit dat de verschillende ILS componenten, waaronder cognitieve en motivationele componenten nauw met elkaar samenhangen in een richting zoals door de literatuur gesuggereerd wordt (Biggs & Rihn, 1984; Tait & Entwistle, 1996). Dit betekent bijvoorbeeld ook dat pogingen om (ongewenste) leerstijlen van studenten te veranderen ook op motivationele aspecten moeten inspelen en niet alleen, zoals zo vaak gebeurt, op cognitieve factoren. Qua leerstijlen bleken de propedeuse-studenten in dit onderzoek te kunnen worden ingedeeld in drie groepen die verschillen al naar gelang hun neiging tot 'extern gestuurd reproductief leren' en 'zelfgestuurd betekenisgericht leren'. Deze twee typen van leren vertonen een grote overeenkomst met diepte- en oppervlaktegericht leren, zoals dat in internationaal onderzoek met geheel andere onderzoeksmethoden is gevonden (in kwantitatief onderzoek met andere vragenlijsten of in kwalitatief, fenomenografisch onderzoek). Onze resultaten met betrekking tot deze twee leertypen ondersteunen voorts de bevindingen van Biggs & Rihn (1984, p. 283): "Contrary to what may seem to be implied in the original distinction these are not either end of a single continuum. Rather, they suggest quite independent ways in which students may become involved in learning: Students may adopt one, the other, or both, to varying extents". Misschien is het vinden van een goede balans tussen beide leertypen wel waar het om gaat en beantwoordt groep 1 in Figuur 2 daaraan. Trainingen gericht op verbetering van leerstrategie zouden dan vooral van belang zijn voor de leerstijlgroepen 2 en 3 (en niet uitsluitend voor groep 3).

Met betrekking tot *onderzoeksvraag 2* vonden we dat leeromgevingen gekarakteriseerd kunnen worden vanuit dimensies die met actief zelfgestuurd leren, dan wel passief extern gestuurd leren te maken hebben. In dit opzicht sluit de IPAL goed aan bij de ILS. Er werd een goed interpreteerbare typologie van leeromgevingspercepties gevonden. De percepties bleken in drie clusters te kunnen worden ingedeeld. Deze clusters onderscheiden zich van elkaar op de variabelen 'activerend onderwijs' en 'kennisoverdragend onderwijs'. Ook hier gaat het om twee onafhankelijke dimensies. Met betrekking tot *onderzoeksvraag 3* vonden we een sterke invloed van de leerstijl op het daadwerkelijke leergedrag. Verder blijkt dat de invloed van de leerstijl op de waarneming van de leeromgeving niet te verwaarlozen is. Het is goed voor docenten om zich te realiseren dat hun feitelijk gedrag in de collegezaal voor studenten op meerdere wijze interpreteerbaar is en dat deze interpretatie mede van de leerstijl van de student afhangt. De leerstijl aan het begin van het moduul heeft niet alleen invloed op hoe men de leeromgeving wenst maar ook hoe men die waarneemt. De relatie tussen de leerstijl en de leeromgevingswensen bleek minder sterk dan gevonden werd in het onderzoek van Wierstra et al. (1995) bij Nederlandse en buitenlandse studenten. In dat onderzoek werd echter gevraagd naar de leeromgevingswensen in het algemeen en niet met betrekking tot een specifiek moduul.

Het onderzoek heeft voorts ook nieuwe instrumenten opgeleverd. De sterke relatie tussen de ILS en de ASA pleit voor de validiteit van beide instrumenten. Ook levert het onderzoek positieve resultaten op voor de betrouwbaarheid en validiteit van de IPAL. Op deze instrumenten kan worden voortgebouwd in vervolgonderzoek. De IPAL leent zich er ook uitstekend voor om -zoals zo vaak gebeurt met leeromgevingsinstrumenten, zie bijvoorbeeld Fraser (1986)- als instrument voor zelfdiagnose door docenten te worden gehanteerd.

In vervolgstudies verdient het aanbeveling onderzoek te doen naar de gecombineerde invloed van de leeromgevingsperceptie (C) en de leerstofspecifieke beleving (G) op het daadwerkelijke leergedrag. Daarbij is het wenselijk sterk variërende leeromgevingen in het design op te nemen. Met name nemen we ons voor om de verschillende instrumenten en verdere versies daarvan af te nemen in leeromgevingen die variëren wat betreft de mate van externe sturing, dat wil zeggen in een substituerende, een activerende en een capitaliserende leeromgeving (Ver-munt, 1992).

DANKBETUIGING

Onze dank gaat uit naar Gellof Kanselaar, Jos van der Linden en twee anonieme beoordelaars voor hun commentaar bij een eerdere versie van dit artikel.

NOOT

1. Het gaat bij de 'wens'schalen altijd om een kleiner aantal items dan bij de perceptie schalen, omdat er van de 'verwachtings'items geen wensschalen gevormd zijn (zie 3.2.2). Dit is ook de belangrijkste reden naast factoronzuiverheid waarom er van de schalen 6 en 7 geen wensschalen zijn.

LITERATUUR

- Bakker, A.J.J. (1991). Leerstijlen: verborgen thema's in opleidingen. In J.W.M. Kessels & C.A. Smit (red.). *Handboek opleiders in organisaties* (pp. 23-50). Deventer: Kluwer Bedrijfswetenschappen.
- Bartlett, F.C. (1932). *Remembering. A study in experimental and social psychology*. Cambridge: University Press.
- Beerends, E.P.M. (1994). *Leerstrategie, leeromgeving en leeropbrengsten*, doctoraalscriptie. Utrecht, Universiteit Utrecht, Vakgroep Onderwijskunde.
- Biggs, J.B. & Rihn, B.A. (1984). The effects of intervention on deep and surface approaches to learning. In J.R. Kirby (ed.). *Cognitive strategies and educational performance* (pp. 279-293). Londen: Academic Press.
- Boekaerts, M. (1983). Motivatie en onderwijs: Theorieën en modellen op een rijtje gezet. In S. Dijkstra & J.R. Takens (red.). *Psychologie en onderwijs*. Lisse: Swets & Zeitlinger.
- Boekaerts, M. (1991). Subjective competence, appraisals, and self-assessment. *Learning and Instruction*, 1, 1-17.
- Boekaerts, M. (1992). The adaptable learning process: Initiating and maintaining behavioural change. *Applied Psychology: An International Review* (special issue), 41, 377-397.
- Boekaerts, M. & Simons, P.R.J. (1993). *Leren en instructie*. Assen: Dekker & Van de Vegt.
- Brinke, S. ten (1994). *Samen kennis maken-over de didactiek van de academische vorming*. Afscheidscollege. Universiteit Utrecht.
- Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193-217.
- Dahlgren, L.O. (1984). Outcomes of learning. In F. Marton, D. Hounsell & N. Entwistle (eds.). *The experience of learning* (pp. 19-35). Edinburgh: Scottish Academic Press.
- Drenth, P.J.D. & Sijtsma, K. (1990). *Testtheorie, inleiding in de theorie van de psychologische test en zijn toepassingen*. Houten: Bohn, Stafleu & Van Loghum.
- Duffy, T.M. & Jonasson, D.H. (1991). Constructivism: New implications for instructional technology. *Educational Technology*, 31, 5, 7-12.
- Entwistle, N., Kozeki, B., & Tait, H. (1989). Pupils' perception of schools and teachers. Relationships with motivation and approaches to learning. *British Journal of Educational Psychology*, 59, 340-350.
- Everitt, B.S. (1980). *Cluster analysis*. Londen: Heineman.
- Fraser, B.J. (1986). Two decades of research on perceptions of classroom environment. In B.J. Fraser (ed.). *The study of learning environments*. Oregon: Assessment Research.
- Hout-Wolters, B.H.A.M. van (1992). *Cognitieve strategieën als onderwijsdoel*. Inaugurele oratie. Groningen: Wolters-Noordhoff.
- Klerk, L.F.W. de (1992). *De moderne academicus*. Rede ter gelegenheid van de viering van de 65^e dies natalis van de KUB. Tilburg University Press.
- Magnusson, D. & Endler, N.S. (1977). Interactional psychology: Present status and future prospects. In D. Magnusson & N.S. Endler (Eds.). *Personality at the crossroads* (pp. 3-31). Hillsdale, N.J.: Lawrence Erlbaum.
- Marton, F. (1986). Phenomenography- A research approach to investigating different understandings of reality. *Journal of Thought*, 21, 28-49.
- Marton, F. & Säljö, R. (1976a). On qualitative differences in learning: 1. Outcome and process. *British Journal of Educational Psychology*, 46, 4-11.
- Marton, F. & Säljö, R. (1976b). On qualitative differences in learning: 2. Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology*, 46, 115-127.

- McLeod, D.B. (1990). Information processing theories and mathematics learning: The role of affect. *International Journal of Educational Research*, 14, 13-29.
- Milligan, G.W. & Cooper, M.C. (1987). Methodology review: Clustering methods. *Applied Psychological Measurement*, 11, 329-354.
- Murray, H.A. (1938). *Exploration in personality*. New York: Oxford University Press.
- Nuy, H.J.P. & Gijselaers, W.H. (1993). Studenten-oordelen over onderwijskwaliteit: de rol van studie-oriëntatie. In H. van Berkel (red.). *Onderwijsonderzoek in Nederland en Vlaanderen* (pp. 205-206). Houten: Bohn, Stafleu van Loghum.
- Pervin, L.A. (1977). The representative design of person-situation research. In D. Magnusson & N.S. Endler (eds.). *Personality at the crossroads* (pp. 371-384). Hillsdale, N.J.: Lawrence Erlbaum.
- Ramsden, P. (1992). *Learning to teach in higher education*. New York: Routledge.
- Riding, R. & Cheema, I. (1991). Cognitive styles - an overview and integration. *Educational Psychology*, 11, 193-215.
- Rossum, E.J. van & Schenk, S.M. (1984). The relationship between learning conception, study strategy and learning outcome. *British Journal of Educational Psychology*, 54, 73-83.
- Schoenfeld, A.H. (1988). When good teaching leads to bad results: the disasters of 'well-taught' mathematics courses. *Educational Psychologist*, 23, 145-166.
- Shuell, T.J. (1988). The role of the student in learning from instruction. *Contemporary Educational Psychology*, 13, 276-295.
- Tait, H. & Entwistle, N. (1996). Identifying students at risk through ineffective study strategies. *Higher Education*, 31, 97-116.
- Trigwell, K. & Prosser, M. (1991). Relating approaches to study and quality of learning outcomes at the course level. *British Journal of Educational Psychology*, 61, 265-275.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs - Naar procesgerichte instructie in zelfstandig denken*. Amsterdam/Lisse: Swets & Zeitlinger.
- Visser (1994) *Leerstijlen en doceergedrag*, doctoraalscriptie. Utrecht: Universiteit Utrecht, Vakgroep Onderwijskunde Utrecht.
- Volet, S.E., Renshaw, P.D., & Tietzel, K. (1994). A short term longitudinal investigation of cross-cultural differences in study approaches using Biggs' SPQ questionnaire. *British Journal of Educational Psychology*, 64, 301-318.
- Wierstra, R.F.A., Jörg, A.G.D., & Wubbels, Th. (1987). Contextual and individually perceived learning environment in curriculum evaluation. In B.J. Fraser, *The study of learning environments*, 2. Perth: Curtin University of Technology.
- Wierstra, R.F.A., Vermunt, J.D.H.M., Kanselaar, G., Linden, J.L. van der, & Lodewijks, J.G.L.C. (1995). *Studying abroad in Europe - Cultural embeddedness of learning styles and learning environments*. Nijmegen, EARLI congrespaper.
- Wierstra, R.F.A. & Linden, J.L. van der (1996). *Leerstijlen en domeinspecifieke cognitieve en affectieve processen*. Intern verslag. Utrecht: Universiteit Utrecht, Vakgroep Onderwijskunde Utrecht.
- Wood, R.E. & Locke, E.A. (1987). The relation of self-efficacy and grade goals to academic performance. *Educational and Psychological Measurement*, 47, 1013-1024.

Manuscript ontvangen 1-5-1996

Definitief aanvaard 21-6-1996

Leerstijlen en het opzoeken van informatie in tekst

Evelien T. Stoutjesdijk en Jos J. Beishuizen
Rijksuniversiteit Leiden, FSW, Onderwijspsychologie

ABSTRACT

Locating information in texts, as well as studying expository texts consist of specification of a reading goal, selection of text units based on the reading goal, processing of text units, and evaluation of the reading result. In two experiments we focussed on text selection. First-year psychology students used a hypertext version of an introductory textbook chapter in order to find a text unit of which a paraphrase had been presented. This search task can be seen as a very restrictive study task. Action logfiles of interactions with the text were analysed for global and local selection decisions and related to several measures from the Inventory of Learning Styles. Deep processors acted mainly departing from an overview during their search, whereas surface processors made more local level decisions. Both groups though, succeeded to the same extent in locating the text unit. Success in a selection task like this seems not to be related to mere learning styles, but to the optimal combination of a processing style with a certain regulational style.

Het opzoeken van informatie in tekstdocumenten is een veel voorkomende activiteit in het dagelijks leven, bijvoorbeeld als onderdeel van leren of werken (Dreher & Guthrie, 1990). Voor studenten in het hoger onderwijs geldt dit evenzeer; om werkboekvragen te beantwoorden of om een essay te schrijven, moeten zij informatie uit tekst kunnen halen. Maar ook het bestuderen van lange studieteksten voor een tentamen vergt opzoeken en selecteren van informatie.

Theorieën op het gebied van lezen (bijv. Van Dijk & Kintsch, 1983; Just & Carpenter, 1987; Brown & Palinscar, 1989) beschrijven en verklaren lezen in de betekenis van tekstbegrip. Zo heeft de invloed van syntactische en semantische eigenschappen van de tekst op het begrijpen en onthouden van de informatie veel aandacht gekregen. Hoewel deze theorieën zeker hebben bijgedragen aan het begrijpen van het fenomeen lezen als tekstbegrip, verklaren ze onvoldoende hoe studenten lange studieteksten doorwerken met doelen als college- of tentamenvoorbereiding of het schrijven van een essay. Ook wordt nauwelijks rekening gehouden met individuele verschillen in strategiekeuze. Aangezien het bestuderen van lange informatieve teksten een belangrijk middel is ter uitbreiding van kennis, is meer inzicht in de achterliggende processen en in de effectiviteit van gehanteerde strategieën nodig om studenten te kunnen onderwijzen in het lezen en leren van lange teksten. Uitbreiding van dit inzicht is het doel van het project waarvan het hier gerapporteerde onderzoek deel uitmaakt. De resultaten zouden kunnen bijdragen aan het ondersteunen van studenten die moeite hebben met het bestuderen van grote hoeveelheden tekst.

Selectie en verwerking van tekst

In aansluiting op algemene modellen voor informatieverwerking (Neisser, 1976), probleem oplossen (Polya, 1957; Newell & Simon, 1972) en zoeken in het geheugen (Norman & Bobrow, 1979) en op meer specifieke modellen voor tekstbestudering (Kaldeway & Oost, 1990) en zoeken in tekst (Guthrie, 1988), gaan we uit van een recursief cyclisch proces van selectie en verwerking van tekst. Dit proces bestaat uit vier (overlappende) stadia: (1) specificatie van een lees- en/of leerdoel, gebaseerd op beschikbare informatie over de taak; (2) selectie van teksteenheden op basis van het doel; (3) verwerking van de teksteenheden; en (4) evaluatie van het leerresultaat. Deze stadia vinden niet noodzakelijkerwijs na elkaar plaats (overlap), herhalen

zich steeds (cyclisch) en op verschillende niveaus (recursief). Zo kan evaluatie leiden tot het besluit te stoppen met studeren omdat de doelen bereikt zijn, of door te gaan met dezelfde of een aangepaste, verfijnde of uitgebreide doelomschrijving. En het stadium van verwerking kan bijvoorbeeld zelf weer uit een cyclus van de vier stadia bestaan.

Hoewel er verschillen zijn tussen studeren en opzoeken van informatie, is bovenstaand model op beide soorten taken toepasbaar. Guthrie en Kirsch (1987) noemen als grootste verschil tussen de twee taken de invulling van het eerste stadium: de doelspecificatie. Ze stellen dat het doel bij studeren, het onthouden van de hele tekst, wezenlijk anders is dan het doel bij het opzoeken van specifieke informatie. Een kenmerk van een zoektaak, vergeleken met een meer algemene studeeropdracht, is de restrictiviteit ervan. De twee typen doelen liggen echter minder ver uit elkaar dan op het eerste gezicht lijkt. Net als in een zoektaak moet bij het schrijven van een essay, maar ook bij studeren voor een tentamen, geselecteerd worden. Met name onder tijdsdruk is selectie van belang. En het is juist in de selectie van tekst waar grote individuele strategische verschillen kunnen optreden. Restricties in de opdracht of met betrekking tot de randvoorwaarden voor uitvoering (bijvoorbeeld een beperkte tijd) noodzaken tot strategisch handelen. In onderzoek waarin we het gedrag tijdens een zoek- en een studeertaak vergeleken, vonden we meer variatie in gedrag en een duidelijkere relatie met leerstijlen bij de zoektaak dan bij een studeertaak (Beishuizen, Stoutjesdijk & Van Putten, 1994).

In de analyse van de hier gerapporteerde experimenten werd de aandacht gericht op individuele verschillen in het stadium van tekstselectie, gebaseerd op de doelspecificatie. Omdat dit stadium meer nadruk heeft in een zoektaak (Guthrie, 1988) dan in een studeertaak en vanwege de verwachte grotere variatie in gedrag, hebben we voor een zoektaak gekozen.

Tijdens de *doelspecificatie* (stadium 1) worden door de student zelf doelen gesteld en gerepresenteerd in een doelomschrijving met criteria voor selectie van relevante teksteenheden. De doelomschrijving omvat een representatie van de taak en mogelijk andere, individuele doelen. Als iemand voor haar plezier leest, richt zij zich op andere aspecten van de tekst dan bij het zoeken van informatie om een essay te schrijven. Studeren voor een tentamen vraagt weer om een ander soort lezen. Een ander onderscheid dat de doelomschrijving betreft, is of de lezer een tekst voor de eerste maal ziet of deze herleest. Het laatste wordt verondersteld bij de voorbereiding voor een tentamen, wanneer men tot doel heeft specifieke kennis te verwerven in plaats van de tekst globaal door te nemen. Studenten zijn zich er in het algemeen van bewust dat een bepaalde situatie een bepaald type lezen vereist (Lorch, Lorch & Klusewitz, 1993).

De doelen van een student kunnen vooral gebaseerd zijn op extern aangereikte opdrachten of met name intern gestuurd worden. Sommige lezers geven er de voorkeur aan hun eigen doelen te stellen, terwijl andere meer de neiging hebben op externe aanwijzingen te steunen. Deze verschillen in regulatie van het leerproces zijn door diverse auteurs beschreven (bijv. Entwistle, 1988; Vermunt, 1992).

Op basis van criteria in de doelomschrijving vindt *selectie van teksteenheden* (stadium 2) om het verwerkingsproces mee te vervolgen plaats. Lezers kunnen een inhoudsopgave, andere in de tekst opgenomen overzichten of zelf opgebouwde schema's gebruiken om teksteenheden te kiezen die relevant lijken voor het huidige leesdoel. Beishuizen (1986) vond dat selectiebeslissingen ook gebaseerd kunnen zijn op subdoelen, die tijdens het lezen opkomen en de volgende stap bepalen. Als de eenheid die men leest (bijvoorbeeld een alinea) relevant is voor het doel en/of nieuwe vragen oproept, kan men besluiten verder te gaan met een direct hieraan gerelateerde eenheid (de volgende alinea). Zo kunnen dus tijdens het zoekproces globale en lokale selectiebeslissingen genomen worden. De globale beslissingen behoren tot wat in literatuur over zoeken en probleemoplossen (Norman & Bobrow, 1972; Pask, 1976) de verbredende of holistische strategie wordt genoemd. Hierbij worden alle relevant geachte aspecten van het probleem vanaf het begin in het zoeken naar een oplossing betrokken. Lokale beslissingen behoren eerder tot de verdiepende of seriële strategie. Hierbij wordt vanuit één relevant aspect van het probleem op lokaal niveau verder gewerkt en pas als deze weg dood loopt, kiest men een ander aspect als startpunt.

Het ligt voor de hand om uitgesproken voorkeuren voor globale of lokale beslissingen te zien in het licht van leerstijlen. Diverse auteurs (Vermunt, 1992; Schmeck, 1983; Entwistle, 1988) maken onderscheid tussen diepte-georiënteerde en oppervlakte-georiënteerde (of stapsgewijze) verwerkingsstrategieën. Hun beschrijving van deze strategieën doet veronderstellen dat de stapsgewijze verwerker in het algemeen de voorkeur geeft aan lokale informatie voor het stellen van subdoelen en de selectie van de volgende teksteenheid, waar de diepte-verwerker uitgaat van een overzicht over het geheel om te opereren op zowel het globale als het lokale niveau.

De taak en het materiaal waren zo gemaakt dat selectiebeslissingen op globaal of lokaal niveau genomen konden worden. De proefpersonen moesten bepaalde informatie opzoeken in een lange studietekst. Om door de tekst heen te gaan, gebruikten zij een hypertext-versie van het hoofdstuk: elektronisch gepresenteerde tekst met diverse hulpmiddelen om te zoeken. Men kon middels een hiërarchie van inhoudskaarten de diverse bladzijden opzoeken, lineair het hoofdstuk doorlezen of van een bladzijde naar allerlei gerelateerde bladzijden springen. Ook de keuze voor een hypertext was er op gericht de gedragsvariatie te vergroten. Het lezen van computerscherm en het gebruik van hypertext kost, vergeleken met het lezen van papier en het gebruik van lineaire tekst, wat meer tijd, maar verandert niet wezenlijk de opname van informatie als daarvoor geen bijzondere hulp is ingebouwd (Reinking, 1988; McGoldrick, Martin, Bergering & Symons, 1992; Alexander, Kulikowich & Jetton, 1994).

Leerstijlen en zoekgedrag

In het voorgaande is al een aantal malen over individuele voorkeuren, leerstijlen en strategieën gesproken. Wij houden ons in ons onderzoek bezig met samenhangen tussen zelf-gerapporteerde leerstijl-gegevens en waargenomen strategieën tijdens de uitvoering van diverse studietaken. Een leerstrategie is een serie leer- en denkactiviteiten die in samenhang worden gehanteerd om een leerdoel te bereiken. Een leerstijl is meer dan dat: het is een samenhangend geheel van geprefereerde leerstrategieën (Schmeck, 1983), maar omvat in bredere zin ook de studiemotieven en de ideeën over leren die de student heeft (Vermunt, 1992). Om leerstijlen in kaart te brengen, gebruiken we de *Inventaris Leerstijlen (ILS)* van Vermunt en Van Rijswijk (1987). Zoals de lezer wellicht bekend, bestaat de ILS uit vier soorten items: over verwerkingsactiviteiten, regulatieactiviteiten en -problemen, studiemotieven en leerconcepties, aansluitend bij de definitie die Vermunt geeft van een leerstijl als een bepaalde combinatie van deze vier elementen, een patroon. Hij noemt vier specifieke patronen: de betekenisgerichte, de reproductiegerichte, de toepassingsgerichte en de ongerichte leerstijl. Als we kijken waaruit deze patronen volgens factoranalyses op de items van de ILS bestaan (Vermunt, 1992), dan blijken de eerste twee leerstijlen vooral bepaald te worden door items die de frequentie van het verrichten van diverse soorten activiteiten bevragen (verwerkings- en regulatieactiviteiten). De andere twee leerstijlen worden in grotere mate – in een replicatiestudie van Busato en Prins (1994) zelfs *alleen* – bepaald door items over leermotieven en -oriëntaties en door items over regulatieproblemen en niet of nauwelijks door verrichte activiteiten.

De meest directe link tussen leerstijlen en waargenomen strategieën moet gezocht worden in de *leeractiviteiten*: zelf-gerapporteerde frequenties van activiteiten enerzijds en waargenomen patronen van activiteiten anderzijds. Wat betreft de ILS betrekken wij daarom alleen de schalen voor regulatie- en verwerkingsactiviteiten in ons onderzoek en beperken wij ons dus tot de onderdelen van de betekenisgerichte en de reproductiegerichte leerstijl. Dit is overeenkomstig de 'nauwe' opvatting van het begrip leerstijlen (Schmeck, 1983). Om verwarring met het begrip leerstijl in brede zin (zoals door andere auteurs gebruikt in dit nummer) te voorkomen, zullen we in in dit artikel spreken over verwerkings- en regulatiestijlen.

Diep verwerken van de studiestof wordt vaak – maar niet altijd – geprefereerd door studenten die zelf hun leren willen reguleren (deze combinatie is een van de kenmerken van Vermunts betekenisgerichte leerstijl). Stapsgewijs verwerken heeft vaak – maar niet altijd – de voorkeur van studenten met een externe regulatiestijl (de kenmerkende combinatie binnen de reproductiegerichte leerstijl van Vermunt). Bij deze combinaties spreken wij van *concordantie* tussen verwerking en regulatie. Vermunt (1992) heeft ook de combinaties van externe regulatie met

diepte-verwerking en zelfregulatie met stapsgewijze verwerking beschreven. In factoranalyses om leerstijlen te identificeren werden deze echter niet terug gevonden, omdat ze 'kennelijk weinig voorkomen' (Vermunt, 1992; p. 128). Deze combinaties zijn dan ook niet binnen zijn leerstijlen onder te brengen. We noemen het *discordante* combinaties van verwerking en regulatie, omdat ze minder gunstig lijken te zijn voor het effectief en efficiënt verwerken van informatie. Hierop wordt hieronder nader ingegaan.

Verwachte uitkomsten van het onderzoek

Gebaseerd op de ILS werden diep en stapsgewijs verwerkende eerstejaars studenten psychologie geselecteerd. Ook waren hun scores voor zelf- en externe regulatie uit de ILS bekend. We verwachtten dat *diepte-verwerking* zich zou manifesteren in het werken vanuit een overzicht van het geheel: overwegend gebruikmaking van de hiërarchische inhoudskaarten om teksteenheden te selecteren, en in verdiepende zoektochtjes binnen enkele relevante onderwerpen: het gebruik van directe relaties tussen teksten. Dit zou een breedte-diepte-strategie genoemd kunnen worden: afwisselend streven naar overzicht en inzicht, uitgaande van de inhoudelijke en formele structuur van het materiaal. Deze werkwijze veronderstelt veel keuzemomenten en -mogelijkheden en dus veel selectiebeslissingen. De criteria voor selectie worden gevonden in het verwerkingsproces zelf. Opstelling en gebruik van deze criteria eist een zekere mate van zelf-regulatie. Een sterke neiging tot zelf-regulatie (concordantie) zou de kans op gebruik van deze strategie bij diepte-verwerkers dan ook vergroten evenals de kans op succes ervan. Een uitgesproken voorkeur voor externe regulatie daarentegen (discordantie), zou diepte-verwerkers in de verleiding kunnen brengen hun strategie steeds te onderbreken op grond van lokaal opgeroepen associaties en in te gaan op alle door het materiaal aangereikte suggesties, zodat een minder duidelijke strategie en minder succes in de zoektaak het gevolg zijn. *Stapsgewijze verwerking* zou herkenbaar moeten zijn aan het stap voor stap doornemen van de stof om het gezochte fragment te vinden, dus aan veel lokale selectiebeslissingen. Concreet betekent dit dat men in het programma een weg van teksteenheden naar teksteenheden zou moeten zoeken via de directe tekstrelaties of langs de lineaire route van het oorspronkelijke hoofdstuk. We verwachtten dat stapsgewijze verwerkers dit suggesties-volgende gedrag vooral zouden vertonen als ze in de ILS hadden aangegeven extern gestuurd te zijn (concordantie). Externe regulatie lijkt bovendien een voorwaarde voor succesvol gebruik van deze strategie; een sterke neiging tot zelfregulatie (discordantie) zou kunnen resulteren in het voortdurend onderbreken van de volgstategie voor eigen uitstapjes, zodat de student de draad kwijtraakt. Weer oppakken van de gevolgde lijn is moeilijk aangezien hij niet vanuit een overzicht werkt.

Kortom, zowel de verwerkingsstijl zelf (diep of stapsgewijs) als de interactie tussen verwerkingsstijl en mate van zelf/externe regulatie werden verondersteld de strategiekeuze te beïnvloeden. In de lijn van bovenstaande redenering is er geen aanleiding te veronderstellen dat een van de twee verwerkingsstijlen tot betere resultaten zou leiden dan de andere; beide groepen zouden met succes hun strategieën kunnen uitvoeren, mits ze op de juiste manier reguleren. Op het succes van de gebruikte strategie verwachtten we daarom alleen een effect van de interactie tussen verwerking en regulatie.

Van de twee hier gerapporteerde experimenten is eerder verslag gedaan in de context van andere onderzoeksvragen (Stoutjesdijk, Verheij & Beishuizen, 1993; Beishuizen, Stoutjesdijk & Van Putten, 1994; Verheij, Stoutjesdijk & Beishuizen, 1996). De data van beide studies zijn samengevoegd en onderworpen aan nieuwe analyses. In de methode-sectie beperken we onze presentatie tot het voor bovenstaande vragen relevante deel van de procedure, dat in beide onderzoeken gelijk was, en tot de resultaten van de overall analyse op de data.

METHODE

Proefpersonen

In twee achtereenvolgende jaren werd eerstejaars studenten psychologie aan de Rijksuniversiteit Leiden verzocht de ILS (Vermunt & Van Rijswijk, 1987) in te vullen. Hieraan gaven

respectievelijk 143 en 142 studenten gehoor. Hun scores op de schalen voor diepte-verwerking, stapsgewijze verwerking, zelfregulatie en externe regulatie werden uitgerekend. Elk van deze schalen bestond uit 11 items over studiegewoonten, beantwoord op een 5-puntsschaal voor frequentie van een activiteit. Voor beide experimenten werden diepte- en stapsgewijze verwerkers geselecteerd. Een diepte-verwerker (D) werd geoperationaliseerd als een student met een score op diepte-verwerking boven de mediaan van de schaal voor diepte-verwerking en een score op stapsgewijze verwerking onder de mediaan van stapsgewijze verwerking. Voor stapsgewijze verwerkers (S) gold het omgekeerde. In totaal voldeden 128 studenten aan de voorwaarden voor een van de groepen (D: 69, S: 59). Een andere voorwaarde voor deelname aan de experimenten was dat het eerste tentamen *Inleiding in de Psychologie*, waarvan het te gebruiken hoofdstuk deel uitmaakte, al moest zijn afgelegd, zodat zinvol zoeken mogelijk was (Guthrie, 1988). Tachtig studenten voldeden aan alle voorwaarden en wilden, tegen betaling, aan het onderzoek meedoen (D: 43, S: 37). Voor hen werd een regulatiescore berekend: de ILS-score voor externe regulatie werd afgetrokken van die voor zelfregulatie. (Omdat het om twee onafhankelijke factoren gaat, gemeten op dezelfde schaal, kan het verschil berekend worden.) Dit verschil werd getransformeerd tot een afwijkingsscore van het overall gemiddelde ($N = 285$), zodat een positieve regulatiescore stond voor de neiging tot zelfregulatie (vergeleken met externe regulatie en vergeleken met de gemiddelde student) en een negatieve score voor de neiging tot externe regulatie. De gemiddelde regulatiescore was voor groep D: 6.3 ($SD\ 9.0$; $N = 43$) en voor groep S: -6.6 ($SD = 8.0$; $N = 37$). Selectie op zowel een duidelijke verwerkings- als een duidelijke regulatiestijl bleek te kleine aantallen proefpersonen op te leveren. Om in deze situatie zo weinig mogelijk informatie te verliezen is gekozen voor behandeling van de regulatiescore als continue variabele.

Materiaal

In beide experimenten werd een hypertext-versie van het hoofdstuk over het geheugen (ca. 20.000 woorden) uit Gleitmans *Psychology* (1990) gebruikt. De structuur van deze hypertext was gebaseerd op een inhoudsanalyse van het hoofdstuk, waarin elk van 150 teksteenheden (ongeveer overeenkomend met de alinea's) was getypeerd als de weergave van een fenomeen, een theorie of empirische evidentie. Vervolgens werden binnen elk van deze drie groepen circa vijf hoofdonderwerpen onderscheiden, waaronder de eenheden konden worden ondergebracht. Voor elk hoofdonderwerp werd de meest algemene, inleidende teksteenheid gekozen als de top van een hiërarchische structuur van subonderwerpen. De resterende eenheden werden in de hiërarchie geplaatst op basis van hun specificiteit, tot maximaal drie niveaus onder de top. Tenslotte werden kris-kras tussen alle onderwerpen van de drie groepen, op alle niveaus, relaties opgespoord en benoemd (bijv.: verklaring, evidentie, voorbeeld).

De hypertext werd gemaakt in HyperCard op een Macintosh Iici met een scherm van A4-formaat en bestond uit een inhoudskaart met de hoofdonderwerpen van de drie groepen (de 'hoofdkaart'), andere inhoudskaarten met subonderwerpen van de hoofdonderwerpen en van verder onderverdeelde subonderwerpen ('subkaarten') en tekstkaarten, elk met een teksteenheid. De navigatiemogelijkheden tussen inhoudskaarten onderling, tussen inhouds- en tekstkaarten en tussen tekstkaarten onderling stonden steeds op het scherm aangegeven en konden door een klik met de muis gerealiseerd worden. De gebruiker zag eerst de hoofdkaart en kon van daaruit een teksteenheid of een subkaart kiezen bij een van de weergegeven onderwerpen. Vanuit de subkaarten had men dezelfde twee opties en kon men bovendien in een keer terug naar de hoofdkaart. Ook vanuit de tekstkaarten was die sprong naar de hoofdkaart altijd mogelijk. Verder kon men (1) vanaf een tekst naar de subkaart die het laatst in beeld was geweest of de kaart die de subonderwerpen van de tekst liet zien (indien van toepassing); (2) naar een tekstkaart die direct inhoudelijk gerelateerd was en in een rijtje van maximaal vier tekstrelaties onder de tekst stond; of (3) naar de volgende teksteenheid volgens Gleitman. Elke tekstkaart was bereikbaar via elk van deze kanalen. Het programma registreerde alle navigatie-handelingen met het tijdstip waarop ze plaatsvonden in een zogenaamd protocol.

De taak bestond uit een tekstfragment, uitgereikt op papier, met de instructie te zoeken naar een tekstkaart met dezelfde inhoud, maar in andere bewoordingen, als het fragment. Het fragment was gekozen met het doel lange handelingsprotocollen te verkrijgen; opsporing binnen enkele minuten was niet waarschijnlijk. Een ander fragment werd gebruikt om te oefenen met het programma en de opdracht.

Procedure

De proefpersonen werkten individueel in sessies van twee tot drie uur aan diverse studeeropdrachten. De eerste 65 minuten werden in beide experimenten besteed aan instructie en de hier beschreven zoektaak. De instructie bestond uit uitleg over de structuur van de tekst en een demonstratie van het programma (20 minuten), de gelegenheid voor de proefpersoon de mogelijkheden van het programma te exploreren en onduidelijkheden door vragen op te helderen (10 minuten) en een niet-geregistreerde oefen-zoektaak (15 minuten), waarbij de proefleider eventuele vragen alsnog beantwoordde. Uiteindelijk werd de experimentele zoektaak uitgevoerd. De proefpersoon kreeg het fragment, las het en werd daarna voor maximaal 20 minuten alleen gelaten om de juiste teksteenheid te lokaliseren, waarbij het programma een handelingsprotocol bijhield. Het fragment bleek inderdaad moeilijk op te sporen; veel studenten maakten de tijd vol zonder het te vinden (zie de sectie Resultaten). De gemiddelde tijd die men aan de taak besteedde was 16.3 minuten ($SD = 3.1$).

Data-analyse

De protocollen werden geclassificeerd naar het type acties dat de overhand had. Als meer dan de helft van de acties in een protocol bestond in navigatie tussen inhoudskaarten onderling en tussen inhoudskaarten en tekstkaarten, werd het protocol getypeerd als *Overzicht-strategie* (OVs). Overwegend voorkomen van tekstrelaties werd getypeerd als de *Directe relatie-strategie* (DRs) en overwegend opvragen van tekstkaarten in de oorspronkelijke boekvolgorde als de *Lineaire strategie* (LNs). Protocollen die geen van deze uitgesproken strategieën vertegenwoordigden tenslotte, werden geclassificeerd als *Gemengde strategieën* (GMs). Bovendien werd uit de protocollen geconcludeerd of het fragment al dan niet binnen de beschikbare tijd gevonden was.

Strategie en Succes waren de afhankelijke variabelen in Chi^2 -analyses van associatie met Verwerkingsstijl (D of S) om onze verwachting te toetsen dat Verwerkingsstijl wel Strategie zou beïnvloeden, maar niet noodzakelijkerwijs Succes. Samenhang tussen Strategie en Succes onderling werd ook onderzocht met Chi^2 -analyses op celfrequenties voor de hele groep en voor D en S afzonderlijk. Tenslotte werden de effecten van Regulatie en van de interactie tussen Regulatie en Verwerking onderzocht met logistische regressie analyses op Strategie (te voorspellen uit Verwerking en Regulatie) en Succes (te voorspellen uit Verwerking, Regulatie en Strategie).

RESULTATEN

Stijl, Strategie en Succes

Eerst worden hier de analyses met de categorische variabelen gepresenteerd; in het volgende deel zullen de effecten van Regulatie in de analyses worden betrokken. De variabele verwerkingsstijl (D of S) wordt bij de bespreking van de resultaten kortweg Stijl genoemd.

Effecten van Stijl op Strategie en Succes

De Overzicht-strategie (OVs) werd door 40 studenten gehanteerd, de Directe relatie-strategie (DRs) door 25, de Lineaire strategie (LNs) door 10 en de Gemengde strategie (GMs) door 5 studenten. De frequentieverdeling hiervan verschilde voor diepte- en stapsgewijze verwerkers (Tabel 1): D gebruikte zeer vaak OVs, terwijl S meer DRs gebruikte dan de andere strategieën en dan D. Het fragment werd door 29 van de 80 proefpersonen gevonden binnen 20 minuten. Er was geen samenhang tussen Stijl en Succes.

Tabel 1. Frequenties van Strategiegebruik en Succes door diepe en stapsgewijze verwerkers.

	Stijl		
	Diep (N=43)	Stapsgewijs (N=37)	Allen (N=80)
Strategie:			
OVs	28	12	40
DRs	8	17	25
LNs	5	5	10
GMs	2	3	5
Succes:			
Gevonden	16	13	29
Niet gevonden	27	24	51

Noot. OVs: Overzicht-strategie, DRs: Directe relatie-strategie, LNs: Lineaire strategie, GMs: Gemengde strategie. Strategie: $\chi^2(3) = 9.44, p = .024$; Succes: geen samenhang met Stijl.

Effecten van Strategie op Succes. Het gebruik van bepaalde strategieën hing samen met het succesvol uitvoeren van de taak (zie Tabel 2). Voor de groep in zijn geheel waren de Overzicht- en de Lineaire strategie het minst succesvol en was de Directe relatie-strategie (en de Gemengde strategie) het meest succesvol. Het zelfde patroon is te zien bij de diepte-verwerkers afzonderlijk. Slechts 10 van hen gebruikten DRs of GMs, maar in die gevallen leidde het zoeken vrijwel altijd (bij 8 van de 10) tot succes. Stapsgewijze verwerkers gebruikten DRs weliswaar frequenter, maar voor hen leidde geen van de strategieën significant vaker tot succes dan een van de andere.

Verwerking en Regulatie, Strategie en Succes

Wie van groep S gebruikte de Overzicht-strategie? Welke diepte-verwerkers gaven de voorkeur aan directe tekstrelaties? Welke stapsgewijze verwerkers waren het meest succesvol? Was succes met de Overzicht-strategie voor groep D afhankelijk van Regulatie? Deze vragen hebben betrekking op de veronderstelde interactie tussen Stijl (D of S) en Regulatie (zelf – extern dimensie) en op mogelijke interacties met Strategie. Aangezien Regulatie als een continue variabele gedefinieerd is, terwijl Stijl, Strategie en Succes categorisch zijn, was het niet moge-

Tabel 2. Succesfrequenties per Strategie voor alle proefpersonen, alsmede opgesplitst naar Stijl.

	OVs	DRs	LNs	GMs	Totaal
Succes allen:	(N=40)	(N=25)	(N=10)	(N=5)	(N=80)
Gevonden	10	15	1	3	29
Niet gevonden	30	10	9	2	51
Succes D:	(N=28)	(N=8)	(N=5)	(N=2)	(N=43)
Gevonden	7	6	1	2	16
Niet gevonden	21	2	4	0	27
Succes S:	(N=12)	(N=17)	(N=5)	(N=3)	(N=37)
Gevonden	3	9	0	1	13
Niet gevonden	9	8	5	2	24

Noot. OVs: Overzicht-strategie, DRs: Directe relatie-strategie, LNs: Lineaire strategie, GMs: Gemengde strategie. D: diepe verwerking, S: stapsgewijze verwerking). Allen: $\chi^2(3) = 12.49, p = .006$; D: $\chi^2(3) = 10.69, p = .014$; S: $\chi^2(3) = 5.62, p = .132$.

lijk om variantie analyse te doen, noch om de interacties te tonen in het gebruikelijke type grafieken. In plaats daarvan hebben we grafiekjes geconstrueerd met de gemiddelde regulatiescore van groep D en S als ‘afhankelijke variabele’ (op de y-as) van Strategie (Figuur 1) of Succes (Figuur 2).

Effecten van Stijl en Regulatie op Strategie. De twee bovenste grafieken van Figuur 1 laten concordantie van verwerkingsstijl en regulatie zien voor de gebruikers van de Overzicht- en de Directe relatie-strategie (OVs en DRs). De diepte-verwerkers die deze strategieën hanteerden,

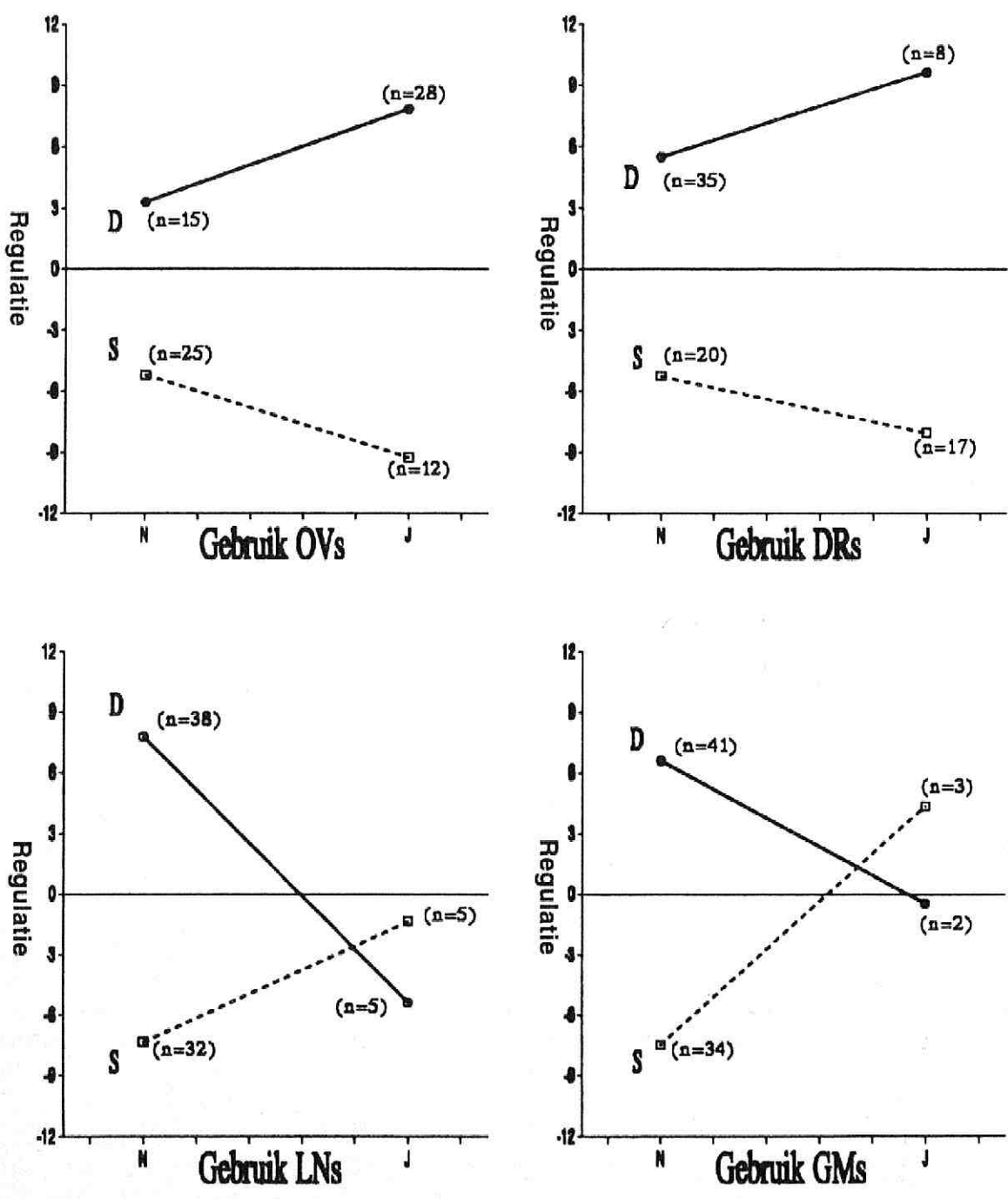


Fig. 1. Gemiddelde regulatiescores van diepte- (D) en stapsgewijze (S) verwerkers afgezet tegen gebruik van een bepaalde strategie (J: de betreffende strategie is wel gebruikt; N: de betreffende strategie is niet gebruikt). Voor elk van de vier strategieën vertegenwoordigt een grafiek de hele groep (N=80).

hadden een hoge regulatiescore (relatief zelfgestuurd) vergeleken met de diepte-verwerkers die niet van deze strategieën gebruik maakten. Stapsgewijze verwerkers hadden bij OV's en DR's juist een lage regulatiescore (relatief extern gestuurd) vergeleken met non-gebruikers van deze strategieën. Discordantie werd daarentegen waargenomen bij de gebruikers van de veel minder vaak gekozen Lineaire en Gemengde strategie (LNs en GMs) in de onderste twee plaatjes van Figuur 1. Hier hadden de gebruikers de meer uitzonderlijke combinaties van Stijl en Regulatie vergeleken met de non-gebruikers (D: relatief extern gestuurd; S: relatief zelfgestuurd).

Om te zien of en in welke mate het gebruik van de vier strategieën kon worden voorspeld uit Stijl en Regulatie werden logistische regressie analyses uitgevoerd (Tabel 3A). Stijl speelde een rol in de keuzen voor de Overzicht- en de Directe relatie-strategie (OV's en DR's). Diepte-verwerking vergrootte, zoals eerder bleek, de kans op OV's (positieve *B*) en verkleinde de kans op DR's (negatieve *B*). Omgekeerd werkte stapsgewijze verwerking dus bevorderlijk voor DR's en remmend voor OV's. Regulatie leverde geen zelfstandige bijdrage aan de voorspelling van beide strategieën. Wel waren er, zoals verwacht, interactie-effecten van Stijl en Regulatie voor sommige strategieën. Op de Overzicht-strategie (OV's) werd een licht positief effect gevonden van de genoemde interactie; op de Lineaire en de Gemengde strategie (LNs en GMs) een licht negatief effect. Een positief effect wil hier zeggen dat de kans op gebruik van de betreffende strategie toenam door concordantie en afnam door discordantie. Dus binnen groep D werd de neiging OV's te hanteren groter naarmate men hogere regulatiescores had (positief: zelfgestuurd) en kleiner bij lagere regulatiescores (negatief: extern gestuurd). Binnen groep S vergrootte juist externe regulatie (negatieve/lage scores) de kans op OV's iets en verkleinde zelfregulatie die kans. We verwachtten hetzelfde patroon te vinden voor DR's, met name binnen groep S meer kans op DR's naarmate men meer extern gestuurd was. Maar hoewel het teken van de coëfficiënt wel positief was, had de interactie tussen Stijl en Regulatie geen significant effect op DR's. De negatieve interactie-effecten op de Lineaire en Gemengde strategie (LNs en GMs) staan voor vergroting van de kans op het gebruik van deze strategieën door discordantie (D – extern; S – zelf) en verkleining door concordantie van verwerkingsstijl en regulatie.

Tabel 3. Statistieken van de significante predictoren in logistische regressie analyses.

Afhankelijke variabele	Predictoren	<i>B</i>	<i>Wald</i>	<i>df</i>	<i>p</i>	<i>R</i>	Perc. juist voorspeld
A. Strategie voorspeld uit Stijl en Regulatie							68 – 94%
OV's	Stijl	.74	8.73	1	.003	.25	
	Stijl × Reg.	.07	4.64	1	.031	.15	
DR's	Stijl	-.66	6.57	1	.010	-.21	
LNs	Stijl × Reg.	-.14	9.13	1	.003	-.34	
GMs	Stijl × Reg.	-.14	5.06	1	.025	-.29	
B. Succes voorspeld uit Stijl en Regulatie							63%
Succes	Stijl × Reg.	.07	4.93	1	.026	.17	
C. Succes voorspeld uit Stijl, Regulatie en Strategie							78%
Succes	Reg.	.17	9.26	1	.002	.27	
	Stijl × OV's	-1.66	6.20	1	.013	-.20	
	Stijl × Reg. × DR's	.26	11.62	1	.001	.30	

A. Elk van de vier strategieën (OV's: Overzicht-strategie, DR's: Directe relatie-strategie, LNs: Lineaire strategie, GMs: Gemengde strategie); onafhankelijk: Stijl (D = 1, S = -1), Regulatie, Stijl × Regulatie. B. Succes; onafhankelijk: Stijl, Regulatie, Stijl × Regulatie. C. Succes; onafhankelijk: Stijl, Regulatie, Strategie, alle hogere orde interacties. Strategie werd in deze analyse opgenomen in de vorm van dummyvariabelen voor OV's, DR's and LNs, elk gecontrasteerd met de gemiddelde kansen op succes.

Noot. N=80.

Effecten van Stijl, Regulatie en Strategie op Succes. De grafieken van Figuur 2 tonen de regulatiescores van groep D en S, weergegeven als ‘afhankelijke’ variabele van Succes. In de eerste grafiek zijn gemiddelden over alle proefpersonen gerepresenteerd, onafhankelijk van welke strategie tot het resultaat had geleid ($N=80$); de andere twee grafiekjes omvatten alleen de proefpersonen die de meest gebruikte strategieën hanteerden: respectievelijk OV's ($N=40$) en DR's ($N=25$).

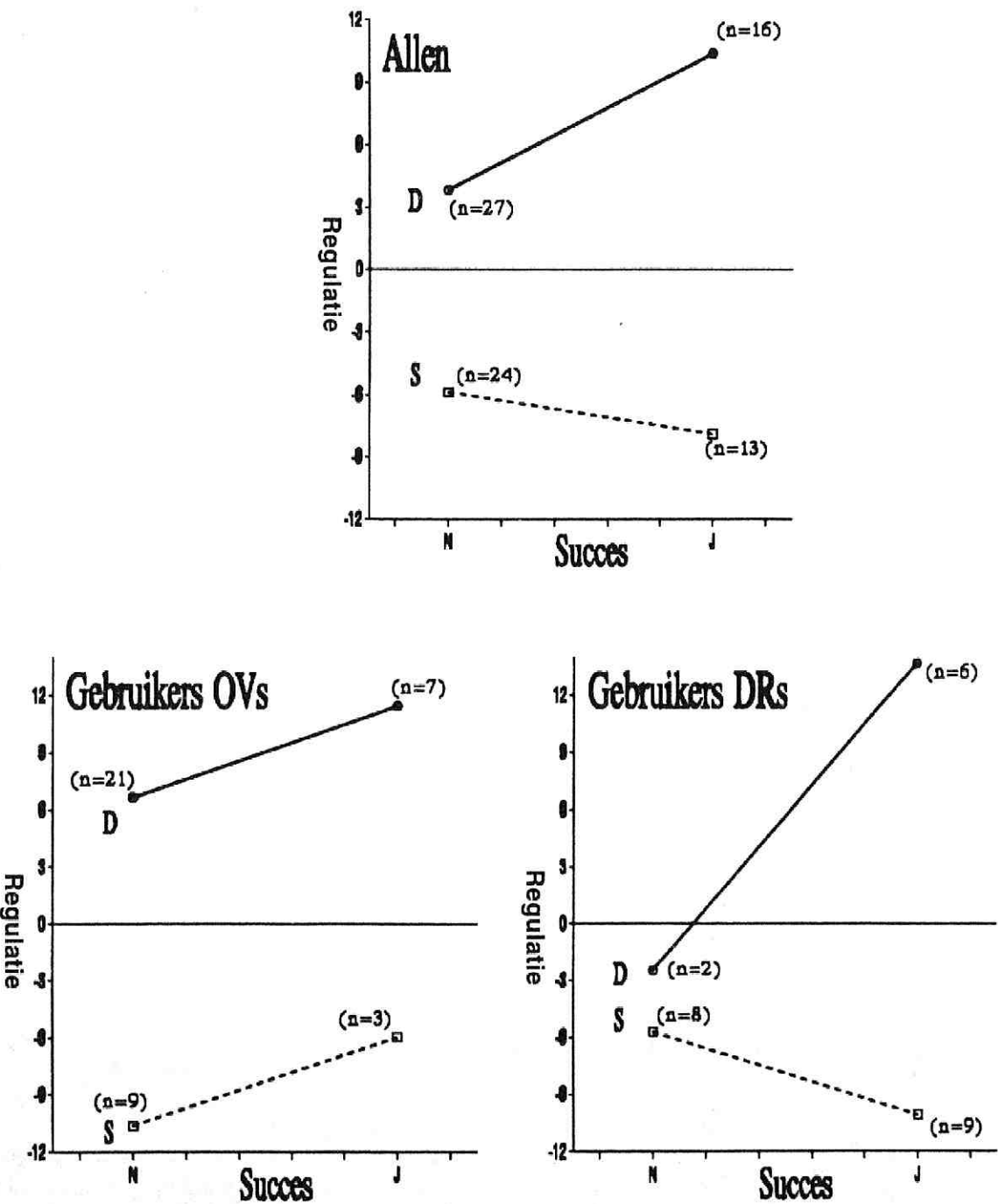


Fig. 2. Gemiddelde regulatiescores van diepte- (D) en stapsgewijze (S) verwerkers afgezet tegen succes bij het vinden van het fragment (J: gevonden, N: niet gevonden). Voor de hele groep ($N=80$), de gebruikers van OV's ($N=40$) en de gebruikers van DR's ($N=25$).

Over het geheel lijken concordante combinaties van verwerkingsstijl en regulatie meer succesvol te zijn geweest dan discordante combinaties, te zien aan het uit elkaar lopen van de lijnen voor D en S in het eerste grafiekje. Dit werd bevestigd door een positief interactie-effect van Stijl en Regulatie op Succes in een logistische regressie (Tabel 3B). Stijl had hierin geen zelfstandig effect op Succes, maar wel in combinatie met Regulatie: vergroting van de kans op succes binnen groep D door zelfregulatie en binnen S door externe regulatie, verkleining binnen groep D door externe regulatie en binnen S door zelfregulatie. Hetzelfde geldt voor de gebruikers van de Directe relatie-strategie apart (Figuur 2), vooral voor groep D, waar de succesvolle proefpersonen veel meer zelfregulerend waren dan degenen die het fragment niet vonden met deze strategie. Voor gebruikers van de Overzicht-strategie wordt uit de grafiek echter een ander patroon zichtbaar. Zowel binnen groep D als binnen groep S waren de succesvolle gebruikers wat meer zelfgestuurd dan de niet-succesvolle gebruikers van OV's. Er lijkt dus sprake te zijn van een tweede orde interactie effect.

In Tabel 3C staan de significante resultaten van een laatste regressie analyse, waarin naast Stijl en Regulatie ook Strategie is opgenomen in de voorspelling van Succes. Uit deze analyse met alle variabelen blijkt dat het hierboven beschreven concordantie effect vooral is veroorzaakt door de gebruikers van DR's. Bovendien werd een klein hoofdeffect van Regulatie gevonden; zelfregulatie leidde over het geheel genomen tot grotere succesansen. Dit verklaart de parallelle lijnen in de grafiek voor OV's van Figuur 2. Nu zouden we kunnen veronderstellen dat groep D dus succesvoller geweest moet zijn dan groep S door een grotere mate van zelfsturing, maar dat was, zoals we eerder hebben gezien, niet het geval. Dit wordt verklaard door compensatie voor de gebruikers van OV's door een negatief effect van $\text{Stijl} \times \text{OV's}$. Kort samengevat: zelfregulatie helpt enigszins bij het vinden van het fragment, maar concordantie van regulatie met verwerkingsstijl heeft een groter positief effect, vooral voor de gebruikers van de Directe relatie-strategie.

DISCUSSIE

Diepte-verwerkers geven er de voorkeur aan informatie in een studietekst op te zoeken vanuit overzichten van het geheel (de Overzicht-strategie), terwijl stapsgewijze verwerkers hun selectiebeslissingen veeleer op het lokale niveau van de tekst nemen (de Directe tekstrelatie-strategie). Studenten in het hoger onderwijs selecteren voortdurend en op alle niveaus te verwerken informatie. Ze worden geacht veel te lezen; veel van de aangeboden informatie heeft de vorm van lange informatieve teksten, zoals artikelen en boeken. Tijdens het voorbereiden van een tentamen vinden, als tijdens ieder informatieverwerkingsproces, talloze momenten van selectie plaats. Bovendien bestaat studeren behalve uit tentamenvoorbereiding, natuurlijk uit vele andere op lange teksten gerichte activiteiten: het schrijven van een paper, het opzoeken van iets wat je vergeten bent of niet begrijpt, enzovoort. Volgens het cyclische model voor de verwerking van tekst moeten voor iedere informatie-eenheid beslissingen genomen worden met betrekking tot het al dan niet verwerken ervan, de mate van verwerking en de manier waarop; de student kan niet alles even diep verwerken. Juiste selectiebeslissingen spelen dus een belangrijke, maar vaak onderbelichte rol bij het studeren. Studenten slagen er echter veelal niet in gevraagde informatie op te zoeken in lange teksten (Kirsch & Jungeblut, 1986; Dreher & Guthrie, 1990; Symons & Pressley, 1993).

In dit artikel werd verslag gedaan van *on-line* metingen van selectieprocessen tijdens een zoektaak en verbanden met leerstijlen. Het bleek dat strategiegebruik bij zo'n taak deels voorspeld kon worden uit verwerkings- en regulatiestijl en dat niet elke strategie even succesvol was. Bovendien bleek dat de kans op succes met een bepaalde strategie deels afhing van de leerstijl van de student. Het meest interessante resultaat is dat stapsgewijze verwerkers even vaak de taak met succes volbrachten als diepte-verwerkers. Voor beide groepen geldt dat de mate waarin iemands regulatievoorkeuren passen bij zijn verwerkingsstijl een betere voorspeller is voor succes dan verwerkingsstijl op zichzelf. Diepte-verwerkers zijn beter af als ze zelfge-

stuurd zijn, stapsgewijze verwerkers kunnen zich beter meer extern laten sturen. Dit noemden we concordantie tussen verwerking en regulatie.

Concordantie blijkt de strategievoorkeur te versterken. Voor diepte-verwerkers bevordert zelfregulatie het gebruik van overzichten om op globaal niveau selectiebeslissingen te nemen. Zij worden door zelfregulatie als het ware pas 'echte' diepte-verwerkers. In termen van Dreher en Guthrie (1990) besteden zij relatief veel aandacht aan het eerste deel van de zoekcyclus: het bepalen van de manier en plaats van zoeken door overzichten te raadplegen. Dreher en Guthrie (1990) vonden dat juist deze extra aandacht een kenmerk is van efficiënt zoeken. Uit ons onderzoek blijkt echter dat andere strategieën evenzeer tot succes kunnen leiden. Stapsgewijze verwerkers geven de voorkeur aan het gebruik van lokale tekstrelaties boven de overzichten om te selecteren, en zijn uiteindelijk even succesvol als de diepte-verwerkers, met name naarmate ze meer extern gestuurd zijn.

Aangezien alleen het gedrag bij een zoektaak is onderzocht, moet natuurlijk voorzichtigheid betracht worden in de generalisatie van de resultaten naar studeren in het algemeen. Het opzoeken en selecteren van informatie is naar ons idee echter wel een zodanig belangrijke activiteit in verschillende studietaken, dat we de resultaten in een wat breder perspectief willen plaatsen.

De bevindingen zijn enigszins in tegenspraak met de heersende visie onder docenten, studie-begeleiders en onderzoekers op het gebied van studievvaardigheden, dat diepte-verwerking en zelfsturing verreweg de voorkeur verdienen boven stapsgewijze verwerking en externe sturing (bijv. Anderson & Roth, 1989; Vermunt, 1992). Training van studenten in zulke studiemethoden slaagt er vaak niet in stabiele gedragsveranderingen te bereiken (Beishuizen & Jongenotter, 1995). Een van de oorzaken hiervan is ongetwijfeld dat de 'diepe' methode niet altijd beloond wordt in het bestaande onderwijssysteem. Uiteindelijk zouden studenten dan ook geen experts moeten worden in één best geachte studeer- en denkwijze, maar moeten ze een breed repertoire verwerven en leren de juiste aanpak voor elke taak te kiezen. Een zoektaak bijvoorbeeld kan een andere aanpak vergen dan de voorbereiding op een toets; een multiple choice tentamen vereist weer een andere voorbereiding dan open vragen.

De resultaten van ons onderzoek suggereren dat we er misschien niet naar moeten streven dat studenten binnen korte tijd helemaal over gaan op andere strategieën dan ze gewend zijn, maar dat we hen eerst zouden moeten helpen te werken vanuit hun eigen stijl, bijvoorbeeld door het aanbieden van zowel overzichten als sturingsfaciliteiten op tekstniveau. Verder zouden discordanties opgeheven kunnen worden door stapsgewijze verwerkers te trainen in het gebruik van externe aanwijzingen en diepte-verwerkers in zelfsturing. Vanuit een situatie waarin dat wat de student uit zichzelf doet ook daadwerkelijk succesvol verloopt, kan flexibilisering van studiemethoden nagestreefd worden.

LITERATUUR

- Alexander, P.A., Kulikowoch, J.M., & Jetton, T.L. (1994). The role of subject-matter knowledge and interest in the processing of linear and nonlinear texts. *Review of Educational Research*, 64, 201-252.
- Anderson, C.W. & Roth, K.J. (1989). Teaching for meaningful and self-regulated learning of science. In J. Brophy (ed.), *Advances in research on teaching*, Vol. 1 (pp. 265-309). London: JAI Press.
- Beishuizen, J.J. (1986). *Leren opzoeken van informatie*. Dissertatie. 's Gravenhage: SVO.
- Beishuizen, J.J., Stoutjesdijk, E.T. & Van Putten, C.M. (1994). Learning styles and regulation during studying and search. *Learning and Instruction*, 4, 151-174.
- Brown, A.L. & Palinscar, A.S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (ed.), *Knowing, learning, and instruction. Essays in honor of Robert Glaser* (pp. 393-451). Hillsdale, NJ: Lawrence Erlbaum.
- Busato, V.V. & Prins, F.J. (1994). *Replicatie- en validatieonderzoek ILS*. Paper gepresenteerd tijdens bijeenkomst van de HogCog, Amsterdam, juni 1994.
- Dreher, M.J. & Guthrie, J.T. (1990). Cognitive processes in textbook chapter search tasks. *Reading Research Quarterly*, 25, 323-339.

- Entwistle, N. (1988). Motivational factors in students' approaches to learning. In R.R. Schmeck (ed.), *Learning strategies and learning styles* (pp. 21-51). New York: Plenum Press.
- Gleitman, H. (1990). *Psychology*. New York: Norton and Company.
- Guthrie, J.T. (1988). Locating information in documents: Examination of a cognitive model. *Reading Research Quarterly*, 28, 178-199.
- Guthrie, J.T. & Kirsch, I.S. (1987). Distinctions between reading comprehension and locating information in text. *Journal of Educational Psychology*, 79, 220-227.
- Beishuizen, J. & Jongenotter, I. (1995). Trainen van studievoordigheden: leren schematiseren of leren reflecteren. In H. C. Schouwenburg & J.T. Groenewoud (eds.), *Studievoordigheid en leerstijlen: Proceedings van de Landelijke Dag Studievoordigheden 1994*. Groningen: Wolters-Noordhoff.
- Just, M.A. & Carpenter, P.A. (1987). *The psychology of reading and language comprehension*. Newton, MA: Allyn and Bacon.
- Kaldeway, J. & Oost, H. (1990). Een proces-produkt model van het studeren. In F. J. Slobbe, J. Blankestijn, O.S.C. Zeeuwen & T.J.M. Willemsen (eds.), *Terugblik op 10 jaar studievoordigheden. Toekomstige ontwikkelingen: Proceedings van de Landelijke Dag Studievoordigheden 1990*. Eindhoven.
- Kirsch, I.S. & Jungeblut, A. (1986). *Literacy: Profiles of America's young adults*. Princeton, NJ: Educational Testing Service.
- Lorch, R.F., Lorch, E.P., & Klusewitz, M.A. (1993). College students' conditional knowledge about reading. *Journal of Educational Psychology*, 85, 239-252.
- McGoldrick, J.A., Martin, J., Bergering, A.J., & Symons, S. (1992). Locating discrete information in texts: Effects of computer presentation and menu formatting. *Journal of Reading Behavior*, 24, 1-20.
- Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: Freeman.
- Newell, A. & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Norman, D.A. & Bobrow, D.G. (1979). Descriptions: An intermediate stage in memory retrieval. *Cognitive Psychology*, 11, 107-123.
- Pask, G. (1976). Styles and strategies of learning. *British Journal of Educational Psychology*, 46, 128-149.
- Polya, G. (1957). *How to solve it*. New York: Doubleday Anchor Books.
- Reinking, D. (1988). Computer-mediated text and comprehension differences: The role of reading time, reader preference, and estimation of learning. *Reading Research Quarterly*, 23, 484-499.
- Schmeck, R.R. (1983). Learning styles of college students. In R.F. Dillon & R. Schmeck (eds.), *Individual differences in cognition, Vol.1* (pp. 233-278). New York: Academic Press.
- Stoutjesdijk, E., Verheij, J. & Beishuizen, J. (1993). Leerstijlen en de bestudering van lange studieteksten. In K. Brantsma S. Koolstra, E. Masset, D. Out, W. Wijzen, & A. Uhlenbeck (eds.), *De student als professional. Proceedings van de Landelijke Dag Studievoordigheden 1993*. Leiden: R.U.L.
- Symons, S. & Pressley, M. (1993). Prior knowledge affects text search success and extraction of information. *Reading Research Quarterly*, 28, 250-261.
- Van Dijk, T. & Kintsch, W. (1983). *Strategies for discourse comprehension*. New York: Academic Press.
- Verheij, J., Stoutjesdijk, E. & Beishuizen, J. (1996). Search and study strategies in hypertext. *Computers in Human Behavior*, 12, 1-15.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs: naar procesgerichte instructie in zelfstandig denken*. Lisse: Swets & Zeitlinger.
- Vermunt, J.D.H.M. & Van Rijswijk, F.A.W.M. (1987). *Inventaris leerstijlen voor het hoger onderwijs*. Tilburg: Katholieke Universiteit Brabant.

Manuscript ontvangen 8-5-1996

Definitief aanvaard 21-6-1996

Leerstijlen en zelfstandig leren in het voorportaal van het studiehuis

Alex Roosendaal en Jan Vermunt

Digitools B.V., Tilburg, en ICLON, Rijksuniversiteit Leiden¹

ABSTRACT

This study addresses the following questions: What learning styles can be discerned with students from the upper phase of secondary school? What relations exist among the processing and regulation strategies these students employ, their mental models of learning and their learning orientations?; how are learning styles associated with test achievements, especially for Natural Sciences and Economy? The Inventory of Learning Styles for Higher Education (Vermunt, 1992) was adapted for use within this secondary school context. This adapted version was administered to 107 students from a secondary school, in the context of a larger experiment. The results indicated that four qualitatively different learning styles could be discerned, and that the application-directed learning style is later in its development than the other three styles: undirected, reproduction directed and meaning directed. The results also indicated that self-regulation of learning is associated with almost all processing strategies, but that external regulation of learning is only associated with a stepwise, detailed analytic processing strategy. Moderate relations of deep processing strategies and self-regulation of learning with test achievements were found. The results are interpreted within recent developments in learning theory and educational innovation.

INLEIDING

Onderzoeken die de laatste jaren zijn verricht naar de leerstijlen van studenten in het hoger onderwijs leveren over het algemeen een redelijk consistent beeld op. Vermunt (1992) vat een leerstijl op als een coherent geheel van verwerkings- en regulatiestrategieën die studenten hanteren bij hun studie, van hun mentale leermodel en hun leeroriëntatie, een geheel dat kenmerkend is voor een student in een bepaalde periode. Hij vond in twee afzonderlijke populaties, studenten van de open universiteit en van een reguliere universiteit, en in verschillende deelonderzoeken, vier kwalitatief verschillende leerstijlen: de ongerichte, reproductiegerichte, betekenisgerichte en toepassingsgerichte leerstijl. De ongerichte leerstijl wordt gekenmerkt door stuurloos leergedrag, een ambivalente leeroriëntatie en een mentaal leermodel waarin veel waarde wordt gehecht aan stimulerend onderwijs en aan samenwerking met medestudenten. De reproductiegerichte leerstijl wordt gekenmerkt door het hanteren van een stapsgewijze verwerkingsstrategie (memoriseren, analyseren), en een extern gestuurde regulatiestrategie, een mentaal leermodel waarin het opnemen van kennis van anderen centraal staat, en een leeroriëntatie gericht op het uittesten van de eigen capaciteiten en het behalen van certificaten. De betekenisgerichte leerstijl wordt gekenmerkt door het gebruik van een diepte-verwerkingsstrategie (relateren, structureren, kritisch verwerken) en een zelfgestuurde regulatiestrategie, een mentaal leermodel waarin het opbouwen van eigen kennis en inzichten centraal staat, en een persoonlijk geïnteresseerde leeroriëntatie. De toepassingsgerichte leerstijl tenslotte wordt gekenmerkt door het gebruik van een concrete verwerkingsstrategie (concretiseren, toepassen), een mentaal leermodel waarin het gebruik van kennis wordt benadrukt, en een beroepsgerichte leeroriëntatie. Wat

Adressen: Alex Roosendaal, Digitools B.V., Postbus 1104, 5004 BC Tilburg.

Jan Vermunt, Rijksuniversiteit Leiden, ICLON, Postbus 9555, 2300 RB Leiden.

Dit is het laatste artikel behorenden bij het themanummer.

betreft de regulatie van het leren komen bij deze leerstijl zowel meer zelfgestuurde als meer extern gestuurde vormen voor. Onderzoeken van Busato, Prins, Hamaker en Visser (1995), Schouwenburg (1996), en Lonka en Lindblom-Ylänne (1996), bij andere groepen studenten in het wetenschappelijk onderwijs, bevestigen dit beeld vrij nauwkeurig.

Het onderzoek van Vermunt (1992) laat ook zien, dat bij studenten in het hoger onderwijs zelfsturing van leerprocessen samengaat met het hanteren van *alle* verwerkingsstrategieën, terwijl externe sturing van leerprocessen alleen samenhangt met het gebruik van stapsgewijze en met name analytische, gedetailleerde verwerkingsstrategieën.

Het bovenbeschreven onderzoek naar leerstijlen en leerstrategieën is mede geïnspireerd door het verschijnsel 'leerfunctie' (Vermunt, 1992; Boekaerts & Simons, 1993). Leerfuncties zijn psychologische functies die vervuld moeten worden wil er sprake zijn van volwaardig leren. Wie deze functies vervult, de leerling of iemand anders (docent, leerboek, studiewijzer, computer, e.d.) kan variëren, als ze maar worden vervuld. Vanuit de optiek van het bevorderen van het zelfstandig leervermogen is het echter ook belangrijk dat leerlingen en studenten niet alleen de vakinhoud leren, maar ook geleidelijk leren zelf leerfuncties te vervullen die tot dan toe door een externe bron werden vervuld. Zo wordt de leerfunctie 'relateren' vervuld door de docent wanneer deze uitlegt hoe twee begrippen, verschijnselen, gebeurtenissen e.d. samenhangen. Op deze wijze kunnen leerlingen tot het zien van die samenhangen worden gebracht. 'Zoeken naar samenhangen' is echter ook een belangrijke leer- en denkstrategie, die leerlingen dienen te beheersen wanneer zij zich zelfstandig in nieuwe materie verdiepen. Leren zelfstandig te leren betekent dan dat leerlingen op school en zeker ook in het studiehuis niet alleen kennis van de vakinhoud verwerven, maar ook de strategieën leren waarmee ze zelf hun kennis van soortgelijke vakinhouden kunnen verwerven, bijstellen en gebruiken om bijvoorbeeld problemen op te lossen. Dit betekent een geleidelijke, systematische overheveling van leerfuncties van de docent naar de leerling. Onderzoek naar leerstijlen en leerstrategieën van studenten in het hoger onderwijs heeft onder meer uitgewezen in welke mate en op welke wijze studenten de verschillende leerfuncties zelfstandig vervullen bij hun studie (bijv. Vermunt, 1996). Voor het invoeren van meer procesgerichte onderwijsvormen is deze kennis onontbeerlijk.

Over de leerstijlen van leerlingen in het voortgezet onderwijs is echter nog maar weinig bekend. Om verschillende redenen achten wij het noodzakelijk dat hierover meer gegevens beschikbaar komen. Vanuit maatschappelijk oogpunt kan informatie over leerstijlen en leerstrategieën van VO-leerlingen bijdragen aan optimalisering van de onderwijskundige inrichting van het studiehuis, de Tweede Fase van het voortgezet onderwijs zoals die in de nabije toekomst gestalte krijgt. Twee belangrijke uitgangspunten van deze vernieuwing zijn 'actief en zelfstandig leren' en 'omgaan met verschillen' (Stuurgroep Profiel 2e Fase V.O., 1994). De conceptualisatie van leerstijlen zoals hierboven geschetst laat zien dat leerstijlen voor een belangrijk deel zijn gedefinieerd in termen van leeractiviteiten en zelfstandig leren. Een belangrijk verschil tussen leerlingen waarmee in het studiehuis rekening gehouden dient te worden is dat in leerstijlen. Over beide onderwijskundige pijlers van het studiehuis kan onderzoek naar leerstijlen van VO-leerlingen dus belangrijke ontbrekende informatie verschaffen.

Theoretisch gezien is de vraag interessant hoe de leerstijlen van leerlingen en studenten zich ontwikkelen gedurende hun onderwijsloopbaan. Zo vond Klatter (1995) bij brugklassers nog maar relatief zwakke samenhangen tussen leerstrategieën, mentale leermodellen en leeroriëntaties. Hieruit valt af te leiden dat het leergedrag van deze groep leerlingen nog maar in beperkte mate wordt gestuurd door hoe zij denken over leren en wat ze met leren willen bereiken (Vermunt, 1995). Bovendien bleek in haar onderzoek dat bij brugklassers nog weinig differentiatie aanwezig was in leeroriëntaties en mentale leermodellen. De ongerichte, reproductiegerichte en betekenisgerichte leerstijl leken in aanleg aanwezig, maar nog lang niet zo uitgesproken als bij studenten in het hoger onderwijs. De toepassingsgerichte leerstijl bleek nog niet als afzonderlijke leerstijl voor te komen, maar was nog ondergedompeld in de betekenisgerichte. Wanneer we deze gegevens van Klatter vergelijken met die zoals die bij studenten in het hoger onderwijs zijn verzameld, doet de vraag zich voor hoe die ontwikkeling in leerstijlen van brugklasser tot student in het hoger onderwijs verloopt.

Praktisch gezien blijkt er een grote behoefte te bestaan op de scholen om een beeld te krijgen van de leerstijlen van hun leerlingen, met name van hun bovenbouw-leerlingen. Deze toenemende belangstelling kan wellicht verklaard worden uit de maatschappelijke ontwikkelingen, met name de nadering van het studiehuis-concept. Zoals boven reeds aangeduid is kennis over leerstijlen van leerlingen van belang vanuit de optiek van zowel actief en zelfstandig leren als omgaan met verschillen tussen leerlingen. Een handzaam diagnostisch leerstijlinstrument kan, getuige de vragen die in toenemende mate de universiteit bereiken, tegemoet komen aan een duidelijke behoefte vanuit de onderwijspraktijk.

In deze bijdrage wordt een onderzoek gerapporteerd naar leerstijlen van leerlingen uit de bovenbouw van het voortgezet onderwijs. Doel van het onderzoek is meer inzicht te verwerven in de aard van die leerstijlen, om zo een bijdrage te leveren aan het optimaal verloop van bovenbesproken maatschappelijke ontwikkelingen, de verheldering van theoretische kwesties en de tegemoetkoming aan praktische behoeften.

Vraagstelling. De vraagstelling van het hier gerapporteerde onderzoek is drieledig:

1. Welke leerstijlen kunnen worden onderscheiden bij leerlingen uit de bovenbouw van het voortgezet onderwijs, met name bij leerlingen van de vijfde klas VWO?
2. Welke samenhangen zijn er tussen de verwerkings- en regulatiestrategieën die leerlingen gebruiken, hun mentale leermodellen en hun leeroriëntaties?
3. Hoe hangen de leerstijlen samen met de behaalde rapportcijfers, met name voor de vakken Natuurkunde en Economie?

METHODE

Leerlingen

Het onderzoek is uitgevoerd bij 107 leerlingen uit alle drie vijfde klassen van het VWO van een school in Tilburg, die de vakken Natuurkunde en/of Economie volgden (58 meisjes en 49 jongens). Deze leerlingen werden gekozen, omdat het hier beschreven onderzoek naar leerstijlen deel uitmaakte van een groter experiment bij de vakken Natuurkunde en Economie. Aangezien vrijwel alle leerlingen of Natuurkunde of Economie of beide vakken in hun vakkenpakket hadden, betreft het hier dus vrijwel de gehele groep VWO-5 leerlingen van deze school.

Inventaris Leerstijlen Voortgezet Onderwijs (ILS-VO)

Voor het bepalen van de leerstijl-aspecten van de leerlingen is gebruik gemaakt van een aangepaste versie van de Inventaris Leerstijlen voor het Hoger Onderwijs (ILS-HO, Vermunt, 1992): de ILS-VO. De ILS-HO is een leerstijldiagnostisch instrument dat vier domeinen meet: cognitieve verwerkingsstrategieën, metacognitieve regulatiestrategieën, mentale leermodellen en leeroriëntaties. In totaal bevat de ILS-HO 120 items die worden verwerkt in 16 tot 20 schalen. Het onderdeel 'Studie-activiteiten' bevat 55 uitspraken waarbij studenten wordt gevraagd op een 5-puntschaal aan te geven in hoeverre zij bij het studeren gebruik maken van de beschreven studie-activiteit. De schaal varieert van (1) ik doe dit zelden of nooit, tot (5) ik doe dit (vrijwel) altijd. Het onderdeel 'Studie-motieven' bevat 25 uitspraken over leeroriëntaties en het onderdeel 'Studie-opvattingen' 40 uitspraken over mentale leermodellen. Bij deze laatste twee onderdelen wordt studenten gevraagd bij elke uitspraak op een 5-puntschaal aan te geven in hoeverre het verwoorde motief of de beschreven opvatting overeenkomt met hun eigen motieven of opvattingen. Bij deze onderdelen varieert de schaal van (1) helemaal mee oneens tot (5) helemaal mee eens.

De ILS-HO werd aangepast aan de nieuwe doelgroep: leerlingen van de vijfde klas VWO. De aanpassingen werden zo minimaal mogelijk gehouden en bestonden er voornamelijk uit dat bepaalde begrippen, zoals studiestof, theorie, docent, studeren en dergelijke, werden vervangen door termen als lesstof, onderwerp, leraar en leren. De aangepaste ILS-VO is getoetst op begrijpelijkheid en duidelijkheid bij vijf leerlingen van verschillende VWO-scholen. Op basis van deze toetsing is de uiteindelijke onderzoeksversie vervaardigd. In tabel 1 worden de schalen van de ILS-VO en het aantal items per schaal vermeld.

Procedure

Het hier gerapporteerde onderzoek naar leerstijlen van VO-leerlingen maakte deel uit van een omvangrijker onderzoek. In wezen fungeerde de afname van de ILS-VO als voormeting van een complex en omvangrijk experiment bij de vakken Natuurkunde en Economie, waarop hier niet verder zal worden ingegaan. De ILS-VO is begin oktober op hetzelfde tijdstip afgenomen bij alle leerlingen van de klassen VWO-5 die Natuurkunde en/of Economie volgden. De leerlingen waren verdeeld over drie leslokalen. In twee lokalen waren de docenten belast met de afname van de ILS-VO en in het derde lokaal de onderzoeker. De afname vond plaats op een voor deze gelegenheid extra gecreëerd lesuur. De invultijd varieerde van 25 tot 50 minuten.

Dataverwerking

De onderzoeksgegevens werden statistisch verwerkt met behulp van het software pakket SPSS/PC+. Op de schalen van de ILS werden hoofdcomponenten-analyses met Varimax-rotatie uitgevoerd. Regressie-analyses werden uitgevoerd om de directe bijdrage van regulatiestrategieën, mentale leermodellen en leeroriëntaties aan de variantie in verwerkingsstrategieën te analyseren, alsmede om de bijdrage van mentale leermodellen en leeroriëntaties aan de variantie in regulatiestrategieën vast te kunnen stellen. Samenhangen tussen ILS-VO schalen en rapportcijfers zijn berekend met behulp van Pearson correlatie-coëfficiënten.

RESULTATEN

In tabel 1 staan achtereenvolgens in de kolommen vermeld het aantal items per schaal en subschaal van de ILS-VO, de interne consistenties (Cronbach α) van deze schalen, de gemiddelde item-gemiddelden en de gemiddelde standaarddeviatie van de items. Zoals in de tabel is te zien, kunnen de subschalen 'relateren en structureren' en 'kritisch verwerken' worden gecombineerd tot de hoofdschaal 'diepte-verwerking'. Datzelfde geldt voor de subschalen 'memoriseren en herhalen' en 'analyseren', die samengevoegd kunnen worden tot de hoofdschaal 'stapsgewijze verwerking'. Ook op het gebied van regulatiestrategieën kunnen hoofd- en subschalen worden onderscheiden (zie tabel 1). Voor de ILS-VO hoofdschalen varieert de Cronbach α van .53 tot .82. De beide subschalen van de hoofdschaal externe sturing hebben een lagere α dan .53. Van de 24 hoofd- en subschalen van de ILS-VO hebben er 23 (96%) een α groter dan .50, 18 (75%) een α groter dan .60, 9 (38%) een α groter dan .70 en één schaal (4%) een α groter dan .80. Vergelijking met de α -waarden die Vermunt (1992) voor de ILS-HO heeft gevonden laat zien, dat de ILS-VO coëfficiënties over het algemeen wat lager zijn. Bij de verwerkingsstrategieën wordt de hoogste gemiddelde score gehaald op de schaal 'memoriseren en herhalen', bij de regulatiestrategieën op de schaal 'externe sturing leerresultaten'. Ook is opvallend dat de gemiddelde item-standaarddeviaties bij de VO-populatie beduidend lager zijn dan bij HO-studenten. De variatie in leerstijl-aspecten lijkt dus bij VO-leerlingen kleiner dan bij HO-studenten.

Tabel 2 geeft de relaties tussen de ILS-VO schalen weer in een vier-factoren hoofdcomponenten-analyse met varimax-rotatie. Op de eerste factor hebben de beide diepte-verwerkingschalen, de beide zelfgestuurde regulatie-schalen, het mentale leermodel 'opbouw van kennis' en de persoonlijk geïnteresseerde leeroriëntatie relatief hoge ladingen. Ook de analytische en concrete verwerkingsstrategieën vertonen een redelijke lading op deze factor. Met uitzondering van de lading van de analytische strategie komt dit patroon aan ladingen sterk overeen met het patroon dat in het hoger onderwijs werd gevonden. Daar werd deze factor geïnterpreteerd als een *betekenisgerichte leerstijl*.

De tweede factor representeert wat in het hoger onderwijs de *ongerichte leerstijl* werd genoemd. Kenmerkend zijn hoge ladingen van stuurloos leergedrag, de mentale leermodellen stimulerend onderwijs en samen studeren, en de ambivalente leeroriëntatie. Eveneens kenmerkend is het ontbreken van enige lading van een verwerkingsstrategie. In tegenstelling tot de bevindingen in het hoger onderwijs heeft ook de beroepsgerichte leeroriëntatie een vrij hoge

Tabel 1. Aantal items, interne consistentie (Cronbach α), gemiddeld item-gemiddelde (M items) en gemiddelde item-standaarddeviatie (SD items) van de ILS-VO-schalen.

ILS-schaal	N items	Alpha	M items	SD items
Verwerkingsstrategieën				
Diepteverwerking	11	.82	2.77	0.73
Relateren & structureren	7	.79	2.90	0.80
Kritisch verwerken	4	.64	2.59	0.83
Stapsgewijze verwerking	11	.71	3.08	0.58
Memoriseren & herhalen	5	.64	3.26	0.79
Analyseren	6	.61	3.02	0.69
Concrete verwerking	5	.65	2.54	0.70
Regulatiestrategieën				
Zelfsturing	11	.74	2.42	0.62
Leerproces en -resultaten	7	.60	2.57	0.69
Leerinhoud	4	.58	2.18	0.71
Externe sturing	11	.59	3.30	0.53
Leerproces	6	.52	2.85	0.72
Leerresultaten	5	.39	3.76	0.58
Stuurloos leergedrag	6	.64	2.02	0.65
Mentale leermodellen				
Opbouw van kennis	9	.59	3.21	0.59
Opname van kennis	9	.53	3.48	0.64
Gebruik van kennis	6	.63	3.83	0.52
Stimulerend onderwijs	8	.79	3.07	0.75
Samen studeren	8	.73	2.40	0.75
Leeroriëntaties				
Persoonlijk geïnteresseerd	5	.65	3.05	0.62
Certificaatgericht	5	.60	3.93	0.64
Testgericht	5	.78	3.24	0.89
Beroepsgericht	5	.73	4.02	0.67
Ambivalent	5	.79	1.94	0.77

lading op deze factor. Kennelijk is in het reguliere voortgezet onderwijs het hebben van een op beroep en werk gerichte oriëntatie niet bevorderlijk voor doelgericht studiegedrag.

De derde factor wordt gekenmerkt door hoge ladingen van de beide stapsgewijze verwerkingsschalen, de beide extern gestuurde regulatieschalen, opname van kennis als mentaal leermodel, en certificaat- en beroepsgerichte leeroriëntaties. Deze ordening komt in hoge mate overeen met die welke bij studenten in het hoger onderwijs werd gevonden en daar werd getypeerd als een *reproductiegerichte leerstijl*. Een uitzondering betreft de beroepsgerichte leeroriëntatie; deze maakte bij HO-studenten geen deel uit van de reproductiegerichte leerstijl.

Op de vierde factor hebben de concrete verwerkingsstrategie, het mentale leermodel 'gebruik van kennis', de persoonlijk en testgerichte leeroriëntaties en de externe regulatiestrategie gericht op leerresultaten relatief hoge ladingen. Van alle vier factoren wijkt dit patroon het meest af van de bevindingen gedaan bij HO-studenten. De twee eerstgenoemde aspecten vormen daar samen met de beroepsgerichte leeroriëntatie de *toepassingsgerichte leerstijl*. De verschillen met het hier gerapporteerde onderzoek betreffen vooral de oriëntatie-componenten en de overwegend externe aansturing van deze stijl.

Samenvattend kan worden gesteld dat de eerste drie factoren in grote mate overeenkomen met de ladingenpatronen van de betekenisgerichte, ongerichte en reproductiegerichte leerstijlen zoals die ook in het hoger onderwijs zijn vastgesteld. De vierde factor combineert enkele elementen van de toepassingsgerichte leerstijl met voornamelijk persoonlijke interesse. Dit duidt erop dat in de bovenbouw van het reguliere voortgezet onderwijs toepassingsgerichte leerstrategieën door leerlingen veel meer op basis van persoonlijke interesses worden gehanteerd dan op basis van beroepsgerichte leeroriëntaties. In ontwikkelingspsychologisch perspectief duidt de aparte vierde factor er wel op dat bij deze VWO-5 leerlingen de toepassingsgerichte leerstijl zich reeds heeft losgemaakt uit de betekenisgerichte.

Tabel 2. Factorladingen van ILS-schalen in een 4 factoren Varimax-oplossing.

ILS-schaal	F1 Betekenis- gericht	F2 Ongericht	F3 Reproductie- gericht	F4 Toepassings- gericht
Verwerkingstrategieën				
Diepte-verwerking				
Relateren & structureren	.73			
Kritisch verwerken	.72			
Stapsgewijze verwerking				
Memoriseren & herhalen			.52	
Analyseren	.37		.61	
Concrete verwerking	.36			.60
Regulatiestrategieën				
Zelfsturing				
Leerproces & -resultaten	.86			
Leerinhoud	.65	-.27		
Externe sturing				
Leerproces			.62	
Leerresultaten			.40	.40
Stuurloos leergedrag		.82		
Mentale leermodellen				
Opbouw van kennis	.67			
Opname van kennis	-.26	.30	.66	
Gebruik van kennis				.72
Stimulerend onderwijs		.48		
Samen studeren		.49		
Leeroriëntaties				
Persoonlijk geïnteresseerd	.30	-.30		.65
Certificaatgeicht			.53	
Testgericht				.40
Beroepsgericht		.52	.48	.27
Ambivalent		.74		
Eigenwaarde	4.2	2.4	1.9	1.4
% verklaarde variantie	21.1	12.0	9.5	7.0
Cumulatief %	21.1	33.2	42.7	49.7

Hoofddcomponentenanalyse, N= 107, ladingen >-.25 en < .25 zijn weggelaten.

Om de samenhangen tussen de verwerkingsstrategieën, regulatiestrategieën, mentale leermodellen en leeroriëntaties op een andere manier te analyseren zijn in totaal tien multiple regressie-analyses uitgevoerd. Vijf multiple regressie-analyses zijn uitgevoerd met de verwerkingsstrategieën relateren & structureren, kritisch verwerken, memoriseren & herhalen, analytisch verwerken en concrete verwerking als afhankelijke variabelen en als onafhankelijke variabelen de regulatiestrategieën, de mentale leermodellen en de leeroriëntaties. De resultaten van deze analyses staan weergegeven in tabel 3.

De mate waarin leerlingen gebruik maken van de verwerkingsstrategieën 'relateren en structureren' en 'kritisch verwerken', blijkt voornamelijk samen te hangen met de mate waarin zij hun leerprocessen en -resultaten zelf sturen. Naarmate leerlingen meer gebruik maken van de kritische verwerkingsstrategie, zijn ze vaker persoonlijk geïnteresseerd in de studiestof en zijn ze minder van opvatting dat 'leren' overeenkomt met het opnemen van bestaande kennis.

Het gebruik van de memoriserende en herhalende verwerkingsstrategie hangt eveneens samen met de mate waarin de leerlingen hun leerproces en -resultaten zelf sturen. Daarnaast wordt het gebruik van deze strategie bepaald door een mentale leermodel waarin leren wordt gezien als het opnemen van kennis, en door een leeroriëntatie gericht op het uittesten van de eigen capaciteiten.

Tabel 3. Betagewichten van leeroriëntaties, mentale leermodellen en regulatiestrategieën als predictoren van verwerkingsstrategieën, op basis van het totale regressiemodel (DF=15,89), en significantieniveaus van de bijbehorende F-waarden.

Verwerkingsstrategieën					
Predictoren	Relateren & structureren	Kritisch verwerken	Memoriseren & herhalen	Analyseren	Concreet verwerken
Regulatiestrategieën					
Zelfsturing					
Leerproces & -resultaten	.41**	.33***	.31*	.17	.31*
Leerinhoud	.07	-.09		.18	
Externe sturing					
Leerproces	.10	.07		.36***	
Leerresultaten	-.14	.07	.12		
Stuurloos leergedrag	-.11			-.15	.34**
Mentale leermodellen					
Opbouw van kennis	.31	.16	-.15	-.08	-.17
Opname van kennis	-.15	-.26*	.36*	.06	-.20
Gebruik van kennis	.26	-.21	.14	.13	.40**
Stimulerend onderwijs	.09	.07			-.07
Samen studeren		-.11	.07	-.08	
Leeroriëntaties					
Persoonlijk geïnteresseerd		.24*	-.13		.31**
Certificaatgericht	-.09			.06	
Testgericht			.21*		
Beroepsgericht	.15	.11		.29**	.07
Ambivalent			-.13	.25*	
R ²	.39***	.43***	.25*	.39***	.40***
F	3.7	4.5	2.0	3.7	4.0

Noot. N=107; gewichten >-.05 en <.05 weggelaten

Significantieniveau: *: $p < .05$; **: $p < .01$; ***: $p < .001$.

Aan de mate waarin leerlingen gebruik maken van de analytische verwerkingsstrategie levert vooral de externe gestuurde regulatiestrategie gericht op leerprocessen een bijdrage. Daarnaast wordt deze strategie ook aangestuurd door beroepsgerichte en ambivalente leeroriëntaties.

Het gebruik van de concrete verwerkingsstrategie hangt in bijna gelijke mate samen met de mate waarin leerlingen hun leerprocessen en leerresultaten zelf sturen, en het ervaren van stuurloos leergedrag. Daarbij wordt de concrete verwerking van de leerstof door de leerlingen in hoge mate bepaald door de opvatting dat kennis ervoor dient om gebruikt te worden en in mindere mate door een oriëntatie die voortkomt uit persoonlijke interesse voor een bepaald vakgebied.

Opvallend is dat de zelfgestuurde regulatiestrategie gericht op leerprocessen en -resultaten op één na alle verwerkingsstrategieën aanstuurt, met uitzondering van de analytische strategie. De externe regulatiestrategie hangt alleen samen met het gebruik van de analyserende strategie, en niet met de andere verwerkingsstrategieën. Stuurloos leergedrag hangt samen met het gebruik van de concrete verwerkingsstrategie.

Een vergelijking met de resultaten zoals gevonden in het hoger onderwijs laat zien, dat er een behoorlijke mate van overeenkomst bestaat tussen de resultaten van beide onderzoeken. Wel blijken bij VWO-5 leerlingen leeroriëntaties een grotere rol spelen ter verklaring van het gebruik van leerstrategieën dan bij studenten uit het hoger onderwijs. Bij HO-studenten bleken leeroriëntaties namelijk nauwelijks bij te dragen tot variantie in het gebruik van leerstrategieën (Vermunt, 1992, pp. 120-121).

In tabel 4 zijn de resultaten opgenomen van vijf multiple regressie-analyses waarbij de regulatiestrategieën zelfsturing leerproces en -resultaten, zelfsturing leerinhouden, externe sturing leerprocessen, externe sturing leerresultaten en stuurloos leergedrag zijn opgenomen als afhankelijke variabelen en de mentale leermodellen en leeroriëntaties als onafhankelijke variabelen.

Tabel 4. Betagewichten van mentale leermodellen en leeroriëntaties als predictoren van regulatiestrategieën op basis van het totale regressiemodel (DF=10,94), en significantieniveaus van de bijbehorende F-waarden.

Regulatiestrategieën	Zelfsturing leerproces & -resultaten	Zelfsturing leerinhoud	Externe sturing leerproces	Externe sturing leerresultaten	Stuurloos leergedrag
Predictoren					
Mentale leermodellen					
Ophouw van kennis	.69***	.34***		-.13	-.08
Opname van kennis	-.05		.16	.21	.15
Gebruik van kennis	-.11		-.07	.25*	.08
Stimulerend onderwijs	-.16		.09		.16*
Samen studeren	.10	-.10	.15	-.07	.13
Leeroriëntaties					
Persoonlijk geïnteresseerd	.25	.29**			-.15
Certificaatgericht	.07				-.14
Testgericht	.07		.07	.08	
Beroepsgericht		.07	.09	.06	-.09
Ambivalent			-.21		.26**
R ²	.34***	.33***	.13	.13	.39***
F	4.9	4.6	1.4	1.4	6.0

Noot. N=107; gewichten >-.05 en <.05 weggelaten.

Significantieniveau: *: $p < .05$; **: $p < .01$; ***: $p < .001$.

Tabel 5. Correlaties tussen (1) ILS-VO-schalen en (2) rapportcijfer voor Natuurkunde ($N=72$) en Economie ($N=71$).

ILS-VO schaal	Natuurkunde r 1-2	Economie r 1-2
Verwerkingsstrategieën		
Diepteverwerking		
Relateren & structureren	.33**	.15
Kritisch verwerken	.22*	.11
Stapsgewijze verwerking		
Memoriseren & herhalen	-.10	-.04
Analyseren	.14	.07
Concrete verwerking	.11	.05
Regulatiestrategieën		
Zelfsturing		
Leerproces en resultaten	.29*	.17
Leerinhoud	.14	.15
Externe sturing		
Leerproces	-.04	-.06
Leerresultaten	-.07	.06
Stuurloos leergedrag	-.14	-.07
Mentale leermodellen		
Opbouw van kennis	.08	.14
Opname van kennis	-.02	.17
Gebruik van kennis	.11	.29**
Stimulerend onderwijs	-.09	-.02
Samen studeren	-.06	.15
Leeroriëntaties		
Persoonlijk geïnteresseerd	-.04	.22*
Certificaatgericht	.11	-.11
Testgericht	-.08	.03
Beroepsgericht	.07	.14
Ambivalent	-.17	-.07
R ²	.29	.33
DF	20,50	20,49
F	1.0	1.2

Significantieniveau: *: $p < .05$, **: $p < .01$ (tweezijdige toetsing).

De mate waarin leerlingen zelf hun leerprocessen en -resultaten sturen blijkt vooral samen te hangen met het mentale leermodel waarbij leren wordt gezien als het opbouwen van kennis. Het door de leerlingen medebepalen van leerinhouden hangt samen met zowel de opvatting dat leren overeenkomt met het zelf opbouwen van eigen kennis en inzichten, als met de oriëntatie waarbij leren uit persoonlijke interesse in het vakgebied centraal staat.

Het door de leerlingen uit handen geven van het bewaken en toetsen van hun leerresultaten, hangt samen met een mentaal leermodel waarbij het gebruik van kennis wordt benadrukt. De mate van stuurloos leergedrag wordt vooral verklaard uit waarde hechten aan stimulerend onderwijs en een onzekere, twijfelachtige houding ten aanzien van leren. Opvallend is dat de mate

waarin leerlingen gebruik maken van een extern gestuurde regulatiestrategie veel minder kan worden verklaard vanuit mentale leermodellen en leeroriëntaties dan het gebruik van zelfgestuurde regulatiestrategieën of stuurloos leergedrag.

In tabel 5 staan de (Pearson) correlatie-coëfficiënten weergegeven tussen de laatst behaalde rapportcijfers (overgangsrapport 4-5 VWO) voor de vakken Natuurkunde en Economie en de scores op de ILS-VO schalen. Uit de tabel blijkt dat bij deze leerlingen hoge rapportcijfers voor Natuurkunde samenhangen met het gebruik van drie aspecten van de betekenisgerichte leerstijl: de relaterende/structurerende en kritische verwerkingsstrategieën en de zelfgestuurde regulatiestrategie gericht op leerprocessen en -resultaten. Het rapportcijfer voor Economie hangt significant positief samen met het mentale leermodel waarin het gebruik van kennis wordt benadrukt en de persoonlijk geïnteresseerde leeroriëntatie. Ook kan worden geconstateerd dat de samenhangen tussen rapportcijfer en het gebruik van de beide diepte-verwerkingsstrategieën en de beide zelfgestuurde regulatiestrategieën in dezelfde richting wijzen als bij het vak Natuurkunde.

CONCLUSIES EN DISCUSSIE

Leerstijlen van leerlingen uit de bovenbouw van het voortgezet onderwijs

Uit de hier gevonden resultaten blijkt dat de samenhangen tussen leerstrategieën, leeroriëntaties en mentale leermodellen bij deze bovenbouw-leerlingen sterker zijn dan bij brugklassers. Het leergedrag staat bij hen dus sterker onder invloed van studie-opvattingen en leermotieven dan bij de brugklassers uit het onderzoek van Klatter (1995), maar nog niet zo sterk als bij studenten in het hoger onderwijs (Vermunt, 1992). Bovendien blijken alle vier stijlen zoals die in het hoger onderwijs werden gevonden al goed herkenbaar bij deze leerlingen, en zijn ze al meer gedifferentieerd dan bij brugklassers. Ook in leeroriëntaties en mentale leermodellen is er sprake van een grotere differentiatie dan bij brugklassers. De toepassingsgerichte leerstijl heeft zich nu losgemaakt van de betekenisgerichte en is duidelijk zichtbaar als een eigen stijl. Wel valt op dat deze toepassingsgerichte leerstijl duidelijk 'nog in ontwikkeling' is. Merkwaardig is bijvoorbeeld de associatie tussen stuurloos leergedrag, concreet verwerken en verschillende andere aspecten van deze leerstijl. Naarmate deze VWO-leerlingen beroepsgericht zijn in hun leeroriëntaties, gaan ze eveneens meer stuurloos te werk bij het leren en staan ze meer ambivalent tegenover hun studie. 'Leren' zien in relatie met zich bekwamen in een beroep kan bij deze leerlingen dus beter worden vermeden. Wellicht dat het studiehuis-concept, met zijn grotere nadruk op het leren van vaardigheden, deze associatie teniet doet.

Sturen van leerprocessen van leerlingen

Evenals bij studenten in het hoger onderwijs blijkt zelfsturing van leerprocessen vrijwel alle verwerkingsstrategieën aan te sturen, terwijl externe sturing alleen samenhangt met het gebruik van een analytische verwerkingsstrategie. Wellicht kan dit erop wijzen dat een hoge mate van externe sturing, waarbij leerlingen langdurig door middel van vragen, opdrachten, aanwijzingen, en dergelijke stap voor stap wordt aangegeven wat ze moeten doen, niet de beste manier is om het gebruik van diepe verwerkingsstrategieën door leerlingen te bevorderen. Misschien is systematisch werken aan het vergroten van het vermogen tot zelfsturing van leerprocessen bij leerlingen een betere weg om een diepteverwerking van de studiestof te bevorderen. In het studiehuis zouden didactische principes als 'leren leren', 'overhevelen van leerfuncties', en 'procesgericht onderwijs' dan ook een centrale plaats in moeten nemen in de onderwijskundige inrichting. Het abrupt wegnemen van externe sturing, zonder serieus te werken aan het ontwikkelen van de vaardigheid tot zelfsturing, lijkt een van de gevaarlijkste mogelijkheden om de pijlers onder het studiehuis weg te nemen.

Samenhangen tussen leerstrategieën en toetsprestaties

In dit onderzoek bleken diepte-verwerkingsstrategieën en een zelfgestuurde regulatiestrategie, allebei aspecten van de betekenisgerichte leerstijl, bij één van de twee onderzochte vakken

significant met behaalde rapportcijfers samen te hangen. Bij het andere vak wezen de correlaties in dezelfde richting, maar bereikten geen significantieniveau. Stapsgewijze en concrete verwerkingsstrategieën en een externe gestuurde regulatiestrategie hingen bij beide vakken niet significant samen met behaalde rapportcijfers. Uiteraard kunnen deze betrekkelijk lage correlaties te maken hebben met de manier waarop de docenten het vak geven en de toetsing vormgeven. Leerlingen stemmen hun leergedrag immers vaak af op de taakeisen die ze percipiëren (Schellings, 1995). Het ontwikkelen van toetsen die het gebruik van hoogwaardige leerstrategieën belonen is dan ook een belangrijke taak voor de inrichters van een studiehuis.

Praktische implicaties

De onderzoeksresultaten suggereren dat, om diepgaand en constructief leergedrag tot stand te brengen, het belangrijk is leerlingen gelegenheid te geven tot zelfgestuurd leren. In deze zin vormen de resultaten een argument vóór invoering van actief en zelfstandig leren als leidend didactisch concept voor het studiehuis. De resultaten wijzen er echter ook op, dat reproductiegerichte leerstijlen bij deze leerlingen meer voorkomen dan betekenis- en toepassingsgerichte leerstijlen. Dit onderstreept de noodzaak tot het ontwerpen en invoeren van onderwijsvormen en toetsen die de geleidelijke ontwikkeling van betekenis- en toepassingsgerichte leerstijlen stimuleren en het gebruik van reproductiegerichte en ongerichte leerstijlen ontmoedigen. Procesgerichte onderwijsvormen, met hun nadruk op een geleidelijke en systematische overheveling van leerfuncties van de docent naar de leerling, vormen hiervoor een goed uitgangspunt (zie bijv. Roosendaal, Thijssen en Kunst-Bruyns, 1994; De Jong & Van Hout-Wolters, 1994; Volet, 1995; Kaldewey, Haenen, Wils & Westhoff, 1996). De resultaten vormen ook een waarschuwing tegen het plotsklaps wegnemen van externe sturing, zoals 'doen aan zelfstandig leren' soms ten onrechte wordt geïnterpreteerd. Het wegnemen van externe sturing zonder dat bij leerlingen de vaardigheid tot zelfstandig leren is opgebouwd, kan niet anders dan leiden tot fricties tussen leer- en doceerstrategieën en tot mislukking van een nastrevenswaardige innovatie.

Een belangrijke praktische opbrengst van het hier gerapporteerde onderzoek is ook de eerste stap die is gezet in de ontwikkeling van een diagnostisch leerstijlinstrument voor het voortgezet onderwijs. Verder normerings- en validatie-onderzoek van dit instrument is echter noodzakelijk.

Suggesties voor verder onderzoek

De ideeën voor verder onderzoek zijn talrijk. Nader onderzoek is bijvoorbeeld gewenst naar de ontwikkeling van leerstijlen gedurende, en na afloop van, de studieloopbaan. Ook de vraag of deze ontwikkeling geleidelijk dan wel discontinu verloopt is van theoretisch en praktisch belang. Zijn er bijvoorbeeld kritische momenten aan te duiden in de studieloopbaan van leerlingen waarop zij het meest open staan voor interventies gericht op verbetering van hun leerstijl? Van praktisch belang is ook kennis over verschillen in leerstijlen tussen havo- en vwo-leerlingen, onder meer in verband met de vormen van zelfstandig leren die in beide schooltypen mogelijk zijn. Verder onderzoek is ook gewenst naar de wijze waarop de overheveling van leerfuncties concreet kan worden vormgegeven in verschillende leeromgevingen. In welk tempo kunnen bepaalde leerfuncties gedurende de studiehuis-jaren worden overgeheveld, kan dat bij sommige leerfuncties sneller dan bij andere, verschillen vakken in de snelheid waarmee en de wijze waarop leerfuncties kunnen worden overgeheveld, en dergelijke zijn in dit verband belangrijke vragen. Hiermee verband houdt onderzoek naar de wijze waarop betekenis- en toepassingsgerichte leerstijlen het best kunnen worden ontwikkeld. Noodzakelijk tenslotte ook is onderzoek naar de constructie van toetsvormen die het gebruik van diepe, analytische en concrete verwerkingsstrategieën belonen en bekrachtigen, en het gebruik van dergelijke toetsvormen in het studiehuis.

DANKBETUIGING

Het onderzoek is uitgevoerd in het kader van het VF-programma 'Leren en instructie van cognitieve strategieën' van het Instituut voor de Lerarenopleiding van de Universiteit van Amsterdam.

LITERATUUR

- Boekaerts, M. & Simons, P.R.J. (1993). *Leren en instructie: Psychologie van de leerling en het leerproces*. Assen: Dekker & Van de Vegt.
- Busato, V.V., Prins, F.J., Hamaker, C., & Visser, K.H. (1995). Leerstijlenonderzoek gerepliceerd; de samenhang tussen leerstijlen en intelligentie. *Tijdschrift voor Onderwijsresearch*, 20, 332-340.
- De Jong, F.P.C.M. & Van Hout-Wolters, B.H.A.M. (1994). Knowledge construction, self-regulation and process-oriented instruction. In F.P.C.M. de Jong & B.H.A.M. Van Hout-Wolters (eds.). *Process-oriented instruction and learning from text* (pp. 3-12). Amsterdam: VU University Press.
- Kaldeway, J., Haenen, J., Wils, S., & Westhoff, G. (red.). (1996). *Leren leren in didactisch perspectief*. Groningen: Wolters-Noordhoff.
- Klatte, E. (1995). Leerstijlen in de brugklas – onderzoek naar een vakspecifieke leerstijl. In H.C. Schouwenburg & J.T. Groenewoud (red.). *Studievaardigheid en leerstijlen* (pp. 169-191). Groningen: Wolters-Noordhoff.
- Lonka, K. & Lindblom-Ylänne, S. (1996). Epistemologies, conceptions of learning, and study practices in medicine and psychology. *Higher Education*, 31, 5-24.
- Roosendaal, L.A., Thijssen, A.P.J., & Kunst-Bruyns, B.L.M. (1994). *Werken met zelfstandig leren volgens de procesgerichte instructie*. Tilburg: Theresialyceum & Katholieke Universiteit Brabant.
- Schellings, G.L.M. (1995). *An educational approach to selecting main points in texts. Motives, methods, tasks, and individual differences*. Academisch proefschrift, Universiteit van Amsterdam.
- Schouwenburg, H.C. (1996). Een onderzoek naar leerstijlen. *Tijdschrift voor Onderwijsresearch*, 21, 151-161.
- Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs (1994). *Tweede fase: scharnier tussen basisvorming en hoger onderwijs*. Den Haag: Stuurgroep Profiel Tweede Fase V.O.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs – Naar procesgerichte instructie in zelfstandig denken*. Amsterdam/Lisse: Swets & Zeitlinger.
- Vermunt, J.D.H.M. (1995). Leerstijlen: een overzicht en recente onderzoeksgegevens. In H.C. Schouwenburg & J.T. Groenewoud (red.). *Studievaardigheid en leerstijlen* (pp. 51-72). Groningen: Wolters-Noordhoff.
- Vermunt, J.D. (1996). Metacognitive, cognitive and affective aspects of learning styles and strategies: A phenomenographic analysis. *Higher Education*, 31, 25-50.
- Volet, S. (1995). Process-oriented instruction: A discussion. *European Journal of Psychology of Education*, 10, 449-459.

Manuscript ontvangen 1-5-1996

Manuscript aanvaard 21-6-1996

Gedecentraliseerde besluitvorming binnen mbo-instellingen: een meetmodel voor domeinen van besluitvorming

Pieter L.J. Boerman en Timo M. Bechger

Vakgroep Onderwijskunde, SCO-Kohnstamm Instituut, Universiteit van Amsterdam

ABSTRACT

This article deals with decentralization of decision making within colleges for vocational education. While the literature on decision making within school organizations shows a diversity in conceptualizations, little is known about the measurement and empirical evidence of these models. The present study concerns decentralization as a function of decision making issues at different organizational levels. We have compared three models of decision making based on work by Bacharach et al. (1990), Hanson (1991), and Slegers (1991) to a simple unidimensional model for decentralization of the decision making system. The various models are tested within Dutch colleges for vocational education (MBO) by means of *confirmative factor analysis* (LISREL8). Data from 298 middle-managers of 87% of all MBO schools in the Netherlands have been used to (a) show the utility of a three domain model, proposed by Hanson, to examine decision making and (b) empirically identify three decision making domains.

1. INLEIDING

Binnen de context van het dereguleringsbeleid van de centrale overheid hebben onderwijsinstellingen in het middelbaar beroepsonderwijs (MBO) te maken met autonomieverschuivingen ten opzichte van die overheid. Deze autonomieverschuivingen houden vergroting van beleidsruimte en zeggenschap in voor instellingen op tal van terreinen, en roepen tevens vragen op van organisatorische aard (Van Wieringen, 1989; Boerman, 1995; Van Amelsvoort & Witziers, 1995). Vraagstukken rond de inrichting van de organisatie(structuur), versterking van het management, beheersbaarheid van de organisatie, enzovoort, zijn hierdoor de afgelopen jaren sterker onder de aandacht gekomen van zowel bestuur en management van de MBO-instellingen als van wetenschappers die onderzoek doen op het terrein van (school)organisatiekunde (Pelkmans & de Vries, 1992).

De vormgeving van het besluitvormingssysteem is een van de voornaamste vragen met betrekking tot het ontwerp van de organisatiestructuur (Butler, 1991). Het ontwerp van het besluitvormingssysteem betreft de volgende vragen: Op welk organisatieniveau ligt welke verantwoordelijkheid? Wie krijgt waar de beslissingsbevoegdheid over? Wie heeft invloed op welke type beslissingen? Het begrip *decentralisatie* is dus gekoppeld aan de structuur van de besluitvorming. De mate van decentralisatie is een functie van het onderwerp van besluitvorming (inhoud) en het organisatieniveau waarop de besluitvorming plaatsvindt (Kochen & Deutsch, 1980).

Het vertrekpunt van deze studie is dat er in de literatuur diverse modellen van besluitvormingsstructuren binnen schoolorganisaties bekend zijn, maar dat er nog weinig empirische evidentie voor de validiteit van deze modellen is gevonden. Dit wordt uiteengezet in de paragraaf over de theoretische verkenningen. Vervolgens gaan wij in op de onderzoeksopzet, modelspecificatie en gaan wij in op de analyses van de diverse modellen. Wij eindigen het artikel met de conclusies en discussie.

Adres: Vakgroep Onderwijskunde, SCO-Kohnstamm Instituut, Universiteit van Amsterdam, Wibautstraat 4, 1091 GM, Amsterdam. Tel. 020-5251259. E-mail: pieter@educ.uva.nl, timo@sco.educ.uva.nl

2. VRAAGSTELLING

Gegeven de toename in autonomie is het te verwachten dat instellingen voor MBO zullen verschillen wat betreft hun decentralisatiegraad. De centrale vraag in dit artikel is wat wij onder de decentralisatiegraad van een schoolorganisatie verstaan. Kent de besluitvorming bepaalde domeinen, ofwel clusters van onderwerpen, met een eigen decentralisatiegraad? Of is de mate van decentralisatie juist onafhankelijk van het besluitvormingsonderwerp? Dit laatste geval rechtvaardigt uitspraken als: "Deze schoolorganisatie is in sterke mate gecentraliseerd". Bij een onderscheid in domeinen kan het gaan om verschillen in decentralisatie per besluitvormingsdomein, bijvoorbeeld een gedecentraliseerd onderwijskundig domein of een sterk gecentraliseerd beheersmatig domein. Dergelijke voorstellingen van modellen (zie ook de beschrijvingen verderop) worden in de literatuur weliswaar beschreven, maar een stevige empirische ondersteuning ontbreekt vooralsnog. De onderzoeksvragen bij dit artikel formuleren wij als volgt: a. Is vast te stellen welke domeinen van besluitvorming kunnen worden onderscheiden? en b. Welke eventuele verschillen in decentralisatiegraad zijn per domein te bepalen?

3. DECENTRALISATIE EN BESLUITVORMING: THEORETISCHE VERKENNINGEN

Decentralisatie als structuurkenmerk

De decentralisatiegraad van een organisatie wordt beschouwd als een kenmerk van de structuur van die organisatie (Pugh e.a., 1968; Mintzberg, 1979; Kochen & Deutsch, 1980; de Leeuw, 1988; Kastelein, 1990). Onder de structuur van een organisatie verstaan wij het geheel aan verdeling van taken en de coördinatie daarvan (Mintzberg, 1979; de Leeuw, 1988; 1994).

De organisatiestructuur betreft zowel het resultaat van gegroeide verhoudingen (de structuur als representant van de feitelijke gang van zaken) als de bewust gecreëerde regulering van de verhoudingen (Ax, 1993). Samengevat heeft de structuur van een organisatie betrekking op:

- a. de ordening van taken en mensen: het onderbrengen van taken bij formele groepen en hun onderlinge hiërarchie (organisatiestructuur in enge zin);
- b. de structuur van de leiding (bevel- en gezagslijnen);
- c. de beleidsstructuur (de wijze waarop beleid tot stand komt en strategische beslissingen worden genomen (o.c.)).

De structuur van een organisatie kan meer of minder gecentraliseerd zijn. Als de formele macht om beslissingen te nemen berust bij een klein aantal posities in de organisatie, dan is die organisatie sterk gecentraliseerd (Price, 1972). Macht verwijst naar de mate waarin een persoon het vermogen heeft om andere personen taken te laten verrichten (o.c.). In een school, als voorbeeld van een organisatie, zou een maximum aan centralisatie bestaan indien alle macht bij een enkel individu lag (bijv. centrale directie); een minimum aan centralisatie indien alle macht gelijk verdeeld is over alle medewerkers.

Mintzberg (1979) en de Leeuw (1994) onderscheiden *verticale* en *horizontale* decentralisatie. Verticale decentralisatie is het naar beneden afstaan van beslissingsbevoegdheden. Dit begrip wordt geïntroduceerd ter onderscheid van de horizontale afhankelijkheid van uitvoerende groepen: in hoeverre hebben non-managers formele beslissingsbevoegdheden. In het kader van dit artikel zijn wij geïnteresseerd in de feitelijke invloedsverdeling over de verticale lijn binnen de organisatie. Het organisatiekundige begrip decentralisatie betreft hier dus slechts de verticale invloedsverdeling binnen de schoolorganisaties.

Decentralisatie: Enkelvoudige of meervoudige domeinen?

In onderzoek naar de mate van decentralisatie van een organisatie komen wij doorgaans twee invalshoeken tegen. Enerzijds onderzoek naar een enkelvoudige maat voor de decentralisatiegraad van een organisatie (bijv. Pugh e.a., 1968; Miller & Dröge, 1986; Hoeksema, 1995). Vanuit deze hoek wordt het begrip decentralisatie één-dimensioneel uitgelegd op het continuüm centraal-decentraal ofwel besluitvorming in de top van de organisatie of bij de professionals als

uitvoerders van de kerntaken. Een voorbeeld van dergelijk onderzoek naar verschillen in decentralisatiegraad komen wij tegen als het gaat om verschillen tussen organisatiestructuren van het politie-apparaat en ziekenhuizen (Kastelein, 1990).

Anderzijds komen wij onderzoek naar decentralisatiegraad tegen dat is toegespitst op diverse domeinen van besluitvorming. Het gaat bij dit type onderzoek meestal om de vraag hoe het is gesteld met de mate van decentralisatie binnen een bepaald domein van besluitvorming. Vindt bijvoorbeeld besluitvorming over bepaalde onderwerpen altijd plaats op het niveau van het topmanagement of juist altijd op het professionele niveau?

In de volgende secties werken wij beide invalshoeken nader uit met het oog op hypothesevorming.

Enkelvoudige invalshoek

Voorname onderzoeken vanuit een enkelvoudige invalshoek zijn de zgn. Astonstudies van Pugh e.a. (1968). Pugh e.a. onderscheiden in hun conceptuele model een aparte dimensie voor de mate van centralisatie. Gegevens over organisatiestructuren zijn verzameld bij 46 instellingen. Met behulp van exploratieve factoranalyse op de items over de organisatiestructuur is ondersteuning gevonden voor het onderscheiden van een op zichzelf staande centralisatiedimensie.

Miller & Dröge (1986) hebben onderzoek verricht onder 93 bedrijven naar dimensies van organisatiestructuur. Gebruikmakend van een exploratieve factoranalyse identificeren zij drie structuurkenmerken: formalisatie, centralisatie en integratie. Centralisatie is een onafhankelijke dimensie en wordt gedefinieerd als de verticale verdeling van besluitvormingsmacht in een organisatie. De dimensie beantwoordt de vraag: op welk niveau in de organisatie vindt de besluitvorming plaats? Hoe meer besluiten er worden genomen bij de top, des te sterker heet de organisatie gecentraliseerd. Hoeksema (1995) repliceert het onderzoek van Miller & Dröge onder instellingen in Nederland en bevestigt de resultaten.

De betekenis voor de hypothesevorming bij deze enkelvoudige insteek is duidelijk: in deze opvatting kennen ook schoolorganisaties één dimensie van besluitvorming. Scholen kunnen volgens dit model wat hun besluitvormingsstructuur betreft worden gekenmerkt op het continuüm centraal-decentraal. Dit noemen wij het *bi-polaire model* (vgl. Kerlinger, 1986).

Meervoudige invalshoek

In onderzoek naar de decentralisatiegraad van een organisatie wordt het begrip decentralisatie doorgaans gekoppeld aan domeinen van besluitvorming. Vragen die dan aan de orde komen, zijn: welke domeinen van besluitvorming kunnen worden onderscheiden? Hoort een bepaald domein 'typisch' thuis bij een bepaald organisatieniveau of met andere woorden, is de aard van de beslissing bepalend voor het organisatieniveau? Zijn er verschillen tussen organisaties in mate van decentralisatie binnen bepaalde besluitvormingsdomeinen?

Drie theoretische benaderingen ofwel modellen zijn in de literatuur bekend, te weten de klassieke dichotome benadering (professioneel versus beheersmatig), het 'Interacting Sphere Model' van Hanson (1991) en de multidomein benadering van Bacharach e.a. (1990). Deze benaderingen zullen wij achtereenvolgens bespreken. Per sectie gaan wij in op de toepassing binnen instellingen voor MBO.

De dichotome benadering: twee domeinen

In schoolorganisatorische modellen en onderzoek komen wij vaak een onderscheid tussen twee domeinen tegen, namelijk het domein van de professionals (de docenten) en het domein van de bureaucraten of administratoren (bestuurders, schoolleiders, administratie) (vgl. van Vilsteren, 1984; Sleegers, 1991; Witziers, 1992). Schoolorganisaties worden vaak getypeerd als zogenaamde 'professionele bureaucratieën' (Mintzberg, 1979; Hoy & Miskel, 1987; Sleegers, 1991). Kenmerkend voor deze organisaties is dat ze zowel professionele als bureaucratische aspecten bezitten, hetgeen door o.a. Sleegers de 'dichotome benadering' is genoemd. Professionele kenmerken komen deels naar voren door de professionele opleiding van docenten en het beschikken over hoogwaardige specialistische kennis. Deze categorie heeft betrekking op het primaire

proces in de school: het lesgeven en het werken met leerlingen/studenten. Bureaucratische kenmerken hebben betrekking op verdeling van taken en functies binnen scholen, het instellen van schoolwerkplannen, formele criteria en regels voor leerlingen en hun prestaties (Sleegers, 1991).

Hypothese-vormend kan gesteld worden dat volgens de dichotome benadering te verwachten is dat een model met de scheiding tussen de twee domeinen de beste beschrijving van de besluitvorming op MBO-scholen oplevert. Docenten en hun verantwoordelijkheid in de klas zijn relatief ver verwijderd van de verantwoordelijkheden van de schoolleiding of centrale directie.

Hansons' Interacting Spheres Model: drie domeinen

In Hansons benadering (1991) worden bovengenoemde twee domeinen verbonden door een onderhandelingsgebied (contested zone), dat ontstaat doordat de beide domeinen niet volledig los van elkaar staan maar elkaar gedeeltelijk overlappen als gevolg van gezamenlijke belangen. Hanson noemt dit drie domeinen model het 'Interacting Spheres Model'. Besluitvorming in dit onderhandelingsgebied is enerzijds schoolgericht maar anderzijds verbonden met het werk van de docenten. Een voorbeeld van een dergelijk besluitvormingsonderwerp in MBO-instellingen is de organisatie van stages, die voornamelijk op opleidingsniveau worden geregeld, maar waarvan de centrale directie zegt dat er bepaalde standaarden voor de hele school gelden. Om een en ander te verduidelijken, gaan wij nader in op het onderzoek van Sleegers (1991) en de bespreking daarvan door Imants (1992).

Sleegers (1991) vindt in zijn onderzoek onder scholen in het voortgezet onderwijs geen bevestiging voor het bestaan van drie domeinen. Zijn resultaten wijzen naar de dichotome benadering. Imants (1990) merkt echter in zijn boekbespreking over het onderzoek van Sleegers op dat Sleegers in feite geen onderscheid maakt tussen besluiten uit het onderwijskundig domein en beheersmatig domein. In zijn optiek betreffen besluiten in het leerkrachtendomein van Hanson die besluiten die heel direct op het werk van de leerkracht met de leerling betrekking hebben. Volgens Imants betreffen veel van de besluiten in het onderzoek van Sleegers besluiten die uit hun eigen aard wel een zekere mate van overleg/afstemming inhouden, zoals *op schoolniveau* afspraken maken over allerlei didactische aangelegenheden. Sleegers rekent deze besluiten tot het leerkrachtendomein van Hanson, maar doet hiermee volgens Imants onrecht aan het Interacting Sphere Model (ISM). Volgens het ISM horen deze taken thuis in het onderhandelingsdomein. Sleegers heeft daarom in de optiek van Imants slechts het onderscheid aangetoond tussen het administratieve domein en de contested zone; in feite heeft het leerkrachtendomein geen rol gespeeld in zijn onderzoek.

Rekening houdend met de kritiek van Imants (o.c.), onderscheiden wij volgens Hansons model voor de besluitvormingsstructuur in het MBO drie domeinen overeenkomstig hun geaardheid in onderwijskundig, beheersmatig en een onderhandelingsdomein. De hypothese bij dit model luidt dat het drie domeinen model (ISM) van Hanson de meest adequate beschrijving van de besluitvormingsstructuur in het MBO geeft.

Bacharachs multi-domein benadering: vier domeinen

Als laatste bespreken wij het model van Bacharach e.a. (1990). Bacharach e.a. hebben in de Verenigde Staten onderzoek gedaan naar domeinen van besluitvormingsparticipatie van docenten in het basis en voortgezet onderwijs. Zij volgden daarbij een multidomein-benadering, waarbij vier domeinen van besluitvorming worden onderscheiden. De domeinen worden gekenmerkt door twee dimensies: 1. operationele vs. strategische besluitvorming en 2. individuge-richte vs. organisatiegerichte besluitvorming. Tegen elkaar afgezet ontstaan vier cellen ofwel domeinen van besluitvorming (zie tabel 1).

Bij operationeel-organisatiegerichte besluitvorming moeten we denken aan taken ten aanzien van organisatie-ontwikkeling en -evaluatie in het algemeen en relaties tussen docenten en leerlingen/ouders in het bijzonder. Bij operationeel-individu gerichte besluitvorming gaat het om individuele taakimplementatie: zaken als gebruik en inzet van materialen en technieken bij

Tabel 1. Domeinen van besluitvormingsparticipatie volgens Bacharach e.a. (1990).

	Operationeel	Strategisch
Organisatiegericht	Issues t.a.v. ontwikkeling en evaluatie	Issues t.a.v. verdeling van middelen
Individugericht	Issues t.a.v. gebruik technieken en materialen	Issues t.a.v. loopbaan

het primaire proces. Strategisch-organisatie gerichte besluitvorming heeft betrekking op organisatiedoelen, verdeling van middelen op lange termijn van zowel materiaal als werk, budgettaire beslissingen, binnenhalen van financiën en aanspreken van hulpbronnen. Tot slot, strategisch-individu gerichte besluitvorming heeft te maken met individuele carrièremogelijkheden, lange termijn taaktoewijzing en salariering.

In hun analyse wijzen vier factoren uit consistent te zijn met de theoretische domeinen. Met behulp van exploratieve factoranalyse op 19 items wordt door Bacharach e.a. een ondersteuning gevonden voor de voorgestelde structuur. De vier factoren verklaren 62% van de variantie van de items. De vier factoren zijn benoembaar zoals voorgesteld in hun model. De factoranalyse geeft volgens Bacharach e.a. een beeld van de schoolorganisatie als een netwerk van kritieke issues, ieder reflecterend naar verschillende aspecten van de 'micropolitics' van de school.

Hypothese-vormend wordt gesteld dat volgens de multidomein benadering van Bacharach e.a., het onderscheid in vier besluitvormingsdomeinen de beste beschrijving geeft van de besluitvormingsstructuur in het MBO.

Uitgeleide

In deze paragraaf hebben wij vier modellen voor besluitvorming in schoolorganisaties besproken. De modellen zijn afzonderlijk en elkaar uitsluitend gepresenteerd. De bedoeling van de nu volgende analyse is na te gaan welk model de situatie ten aanzien van besluitvorming binnen MBO-instellingen het best weergeeft. In de volgende paragraaf starten wij met de beschrijving van de onderzoeksopzet, de constructie van de variabelen en de modelspecificaties.

4. ONDERZOEKSOPZET, INSTRUMENTONTWIKKELING EN MODELSPECIFICATIE

Meting

In totaal zijn 60 items ontwikkeld om de mate van decentralisatie te meten. Besluitvormings-items zijn overgenomen uit schalen van Bacharach e.a. (1990), Sleegers (1991), Witziers (1992), Pelkmans & de Vries (1992) en Ax e.a. (1993). De lijsten van besluitvormingsonderwerpen die deze auteurs gebruiken vertonen grote overlap, hetgeen aangeeft dat hiermee het brede scala aan onderwerpen van besluitvorming wordt gedekt. In bijlage 1 staan de items opgesomd. De items zijn aan de situatie in het MBO aangepast.

Voor de scoring gebruikten wij een continue lijn van 0 tot 10. Hierbij betekent een '0' dat de besluitvorming centraal (directieniveau) plaatsvindt, een '10' betekent dat de besluitvorming decentraal (opleidingsniveau: docenten en coördinatoren) plaatsvindt en een '5' dat de invloed van zowel decentraal als centraal niveau op de besluitvorming even groot is. Alle tussenliggende waarden kunnen eveneens gescoord worden. Een voorbeeld van de scoringslijn wordt gegeven in figuur 1.

De respondenten is gevraagd aan te geven waar de feitelijke besluitvorming plaatsvindt. Het antwoord op elk van de 60 vragen werd gegeven door middel van een verticaal streepje.

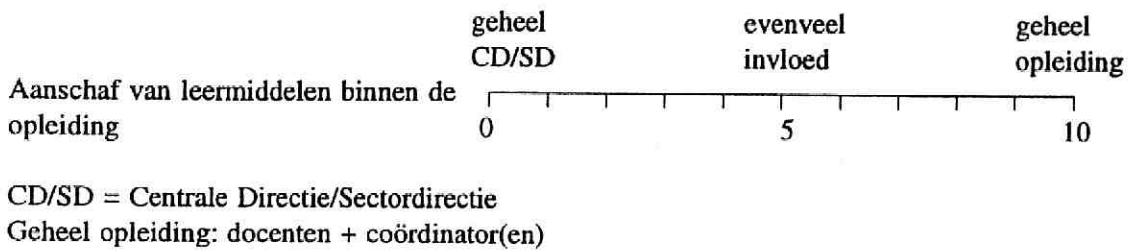


Fig. 1. Voorbeeld van scoringslijn besluitvormingsitems.

Populatie en respons

Een vragenlijst over besluitvorming en invloedsverdeling is voorgelegd aan zgn. coördinatoren van colleges voor middelbaar beroepsonderwijs in Nederland. Coördinatoren zijn in dit geval opleidings-, afdelings- en/of vakgroepscoördinatoren: die mensen in de schoolorganisatie die een coördinerende taak ten aanzien van het onderwijs hebben. Om de vergelijkbaarheid tussen organisaties te optimaliseren is gekozen om organisatieleden, die op hetzelfde organisatieniveau werkzaam zijn, te betrekken in het onderzoek. Coördinatoren zijn leden van de organisatie die in de 'lijn' van de organisatie werkzaam zijn. In de literatuur worden dergelijke middenmanagers aangeduid als zogenaamde '*key-persons*' (Seashore e.a., 1983; Moch e.a., 1983).

Van alle scholen met een sector techniek en/of economie zijn de coördinatoren aangeschreven. Dit leverde een bestand op van 481 coördinatoren verdeeld over (precies) 100 scholen. Deze scholen vormen samen de *hele populatie*. De respons bedroeg 62% (298 ingevulde en bruikbare vragenlijsten) die verdeeld zijn over 87% van alle MBO-scholen in Nederland met een sector techniek en/of economie.

Variabelenconstructie

De 15 achterliggende schalen zijn met behulp van klassieke betrouwbaarheidsanalyse (Cronbachs alfa's) en factoranalyse (PCA) geanalyseerd. In bijlage 2 wordt de procedure nader toegelicht (zie bijlage 2). Op grond van de betrouwbaarheid van de schalen is besloten samengestelde variabelen te construeren. In totaal gaat het om 15 samengestelde variabelen (zie tabel 2):

Per schaal zijn gemiddelde somscores per respondent berekend. Het voordeel van het midden van somscores is dat de resultaten volgens de oorspronkelijke antwoordschaal kunnen worden geïnterpreteerd. Voor de verdere analyses zoals beschreven in dit artikel is gebruikt gemaakt van deze samengestelde variabelen.

Modelspecificatie

De modelspecificatie is in tabel 3 weergegeven. Latente variabelen staan vetgedrukt; de geobserveerde variabele standaard gedrukt. De modellen worden achtereenvolgens besproken waarbij de gebruikte afkortingen van de latente variabelen worden toegelicht.

Tabel 2. Overzicht samengestelde variabelen.

s1. didactiek-micro	s9. strategisch personeelsbeleid
s2. leerroute- en keuzemogelijkh. leerlingen	s10. infomanagement leerlingen
s3. toetsorganisatie	s11. infomanagement docenten
s4. roostering vakken	s12. externe contacten opleidingen
s5. modulering	s13. pr/werving leerlingen
s6. stage-organisatie	s14. strategisch beleid
s7. werkorganisatie	s15. operationeel beheer
s8. operationeel personeelsbeleid	

Tabel 3. Modelspecificatie.

Model 1.0	Model 2.0	Model 3.0	Model 4.0
DecGr s1 to s15	ProfDom s1 to s7; s10 AdminDom s8, s9; s11 to s15	ProfDom s1 s2 s3 s4 s5 OnderhDom s6 s7 s8 s10 s12 s15 AdminDom s9 s11 s13 s14	OperOrg s5 s6 s7 s10 s15 OperInd s1 s2 s3 s4 s8 StratOrg s12 s13 s14 StratInd s9 s11

Naamgeving van de variabelen:

- s1. didactiek-micro

s2. leerroute- en keuzemogelijkh. leerlingen

s3. toetsorganisatie

s4. roostering vakken

s5. modulering

s6. stage-organisatie

s7. werkorganisatie

s8. operationeel personeelsbeleid
- s9. strategisch personeelsbeleid

s10. infomanagement leerlingen

s11. infomanagement docenten

s12. externe contacten opleidingen

s13. pr/werving leerlingen

s14. strategisch beleid & beheer

s15. operationeel beleid & beheer

Model 1.0. Bipolaire model

Dit model kent een bipolaire besluitvormingsstructuur, waarbij één-dimensionale index voor de decentralisatiegraad voldoet. De decentralisatiegraad wordt opgevat als één latente achterliggende variabele (DecGr) waarop alle geobserveerde variabelen s1 tot en met s15 laden. Er is geen onderscheid in besluitvormingsdomeinen. Er is sprake van een bi-polair model: scholen variëren op de dimensie centraal-decentraal.

Model 2.0. Dichotoom model

Binnen het dichotome model valt de besluitvorming in schoolorganisaties uiteen in twee onafhankelijke domeinen: een domein voor professionals en een domein voor de administratoren. Beide domeinen zijn op te vatten als latente variabelen (ProfDom en AdminDom) die respectievelijk de samenhang in de geobserveerde variabelen kunnen verklaren.

Model 3.0. Interacting Spheres Model van Hanson

Binnen het Hanson-model valt besluitvorming in schoolorganisaties uiteen in drie domeinen: het zogenoemde Interacting Spheres Model. Eén domein van onderwijskundige beslissingen (latente variabele: ProfDom), één domein van beheersmatige beslissingen (AdminDom) en één domein waarbinnen zowel het topmanagement als het docentenniveau invloed heeft op de besluitvorming (het onderhandelingsdomein) (OnderhDom).

Model 4.0. Multi domeinen model van Bacharach e.a

Het Bacharach-model bestaat uit vier besluitvormingsdomeinen. Besluitvorming valt uiteen in twee dimensies, die vier domeinen representeren: strategisch vs. operationeel en organisatiegericht vs. individu-gericht (Bacharach, 1990). De variabelen zijn in Bacharachs schema ingedeeld volgens bovenstaande tabel 3. De vier domeinen vormen de latente variabelen: OperOrg, OperInd, StratOrg en StratInd.

5. RESULTATEN

De bedoeling van de analyse is om de genoemde theoretische modellen uit paragraaf 2 te toetsen op grond van de verzamelde gegevens. Alvorens over te gaan op de modeltoetsing werpen wij een blik op de verdeling van de geobserveerde variabelen (zie tabel 4).

Uit tabel 4 blijkt dat de verdelingen van de samengestelde variabelen geen perfecte normaal verdelingen weergeven. Bijv. variabele s1, didactiek-micro, kent weinig spreiding en gaat gepaard met een scheve verdeling naar rechts. Dit laatste betekent dat de besluitvorming over dit item op bijna alle MBO-scholen in Nederland op een laag organisatieniveau plaats vindt. De overige variabelen kennen een minder scheve of platte verdeling.

Plaatsen wij de gemiddelden in afnemende volgorde op een rijtje, dan ontstaat het plaatje zoals in figuur 2 is weergegeven. De verticale as geeft de mate van decentralisatie op een bepaald besluitvormingsonderwerp weer.

Op het eerst oog valt op dat de typische 'onderwijskundige' of 'professionele' zaken in deze rangordening enigszins clusteren. Ook voor items die qua aard als 'beheersmatig' of als 'des schoolleidings' te typeren zijn, geldt dat ze – op het oog – clusteren. De vraag is wat deze clustering voorstelt in het licht van de theoretische modellen van besluitvormingsdomeinen in schoolorganisaties die in paragraaf 2 zijn besproken. In de hiernavolgende sectie zullen wij de diverse modellen toetsen.

Confirmatieve factoranalyse: toetsing van de modellen

In deze sectie gaan wij nader in op het meetmodel voor de variabelen: de structuur tussen latente variabelen (besluitvormingsdomeinen) en geobserveerde variabelen (besluitvormingsitems). In LISREL-termen is dit de uitwerking van het factoranalytisch gedeelte van een LISREL-model (Jöreskog & Sörbom, 1993; Bechger & Hox, 1994).

De analysestrategie is als volgt. Uit het theoretisch gedeelte volgt dat wij vier basismodellen voor besluitvormingsdomeinen onderscheiden: het bi-polaire model met één decentralisatiedimensie; het dichotome model met twee domeinen: een professioneel en een beheersmatig domein; een Hanson model met drie domeinen: onderwijskundig, beheersmatig en een onderhandelingsdomein; tot slot het vier-domeinen-model van Bacharach e.a.

Tabel 4. Gemiddelden en standaardafwijkingen van de samengestelde variabelen.

Samengestelde variabele	gem.	st.dev.	kurt.	skew.
1. didactiek-micro	9.07	.82	2.15	-1.28
2. leerroute/keuzemogelijkheden lln	5.67	2.01	-.59	-.02
3. toetsorganisatie	7.68	2.02	.60	-.92
4. roostering vakken/lln	5.50	2.68	-.97	-.18
5. modulering	7.09	2.23	-.17	-.62
6. stage-organisatie	7.07	2.41	.23	-.87
7. werkorganisatie	4.98	2.40	-.93	-.01
8. oper.pers.beleid	4.28	2.17	-.24	.09
9. strat.pers.beleid	2.90	2.15	-.41	.68
10. info leerlingen	7.07	2.31	-.02	-.72
11. info docenten	3.11	2.48	-.63	.63
12. ext.contacten opleidingen	4.57	2.65	-.86	.12
13. pr/werving	3.56	2.17	-.45	.16
14. strat.beleid&beheer	2.16	1.38	-.12	.70
15. oper.beleid&beheer	3.79	2.09	-.48	.12

Noot. n=298.

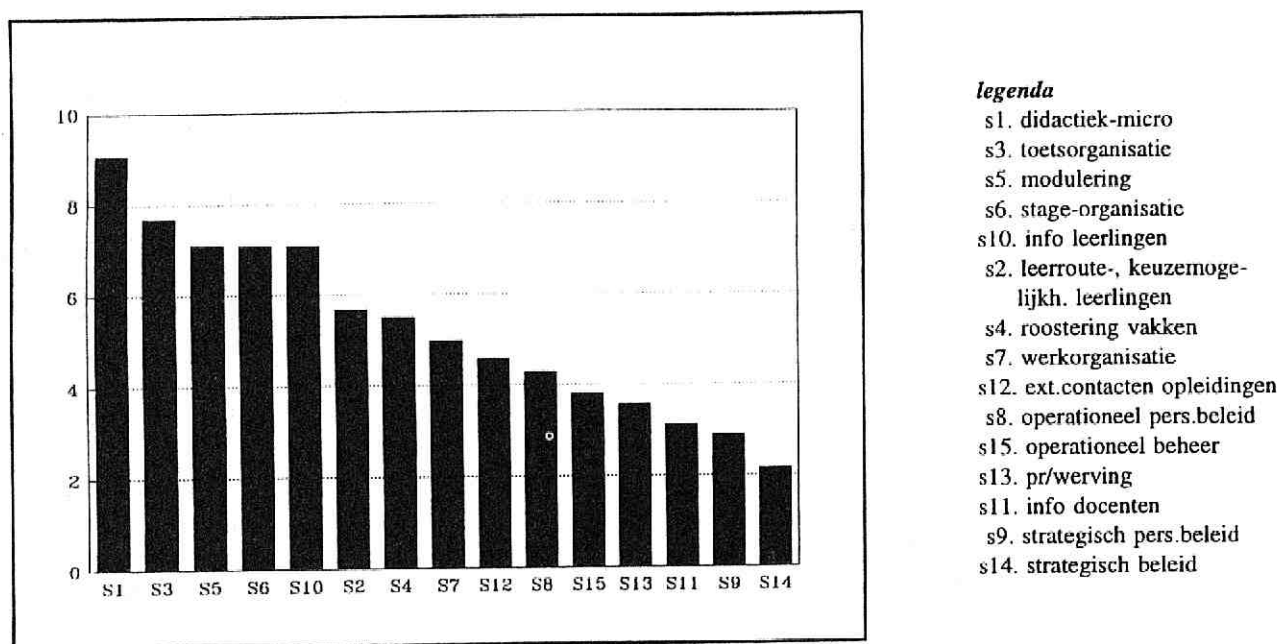


Fig. 2. Rangorde van de gemiddelde scores van de samengestelde variabelen ($n = 298$).

De vier basismodellen van besluitvormingsdomeinen worden geconfronteerd met de data. Het doel is vast te stellen welk model de beste passing ('fit') vertoont met de data. De passing van de basismodellen wordt met elkaar vergeleken. Daarna worden eventuele modelmodificaties uitgevoerd, voor een verbetering van de passing.

De variabelen zijn niet noodzakelijkerwijs normaal verdeeld over het centralisatie-decentralisatie continuüm. De data geven hiertoe ook geen aanleiding (zie tabel 4). Om deze reden maken wij gebruik van de ongewogen kleinste kwadraten methode om parameters te schatten van ieder van de basismodellen (ULS: Jöreskog & Sörbom, 1993).

Resultaten van modelschattingen

In deze sectie gaan wij in de uitkomsten van de modeltoetsingen. In tabel 5 staan de resultaten van de modelschattingen.

Wanneer de ULS-schattingmethode wordt gebruikt kan de chi-kwadraat test alleen als passingsmaat worden gebruikt (Bollen, 1989). Startend vanuit het bi-polaire model 1 kan modelverbetering optreden door meer parameters toe te laten, ten koste van de spaarzaamheid van het model, uitgedrukt in zogenaamde vrijheidsgraden.

Van model 1 naar model 2 is een duidelijke modelverbetering zichtbaar: de chi-kwadraat neemt af met ruim 800 punten ten koste van 1 vrijheidsgraad. De betekenis hiervan is dat besluitvorming in MBO-instellingen geen bi-polair gegeven is. Wij vinden in onze data minimaal twee onafhankelijke domeinen: bijv. een professioneel en een bureaucratisch domein.

Model 3 is met overtuiging het beste model om de data te beschrijven: Op alle criteria uit tabel 5 staat model 3 bovenaan. Modelverbeteringen vinden plaats door error-covarianties toe te staan tussen s8 en s15 (model 3.1) en tussen s3 en s5 (model 3.2). Een error-covariantie tussen twee variabelen wil zeggen dat deze variabelen een latente error gemeenschappelijk hebben. In alle gevallen is dit te wijten aan zekere overeenkomst bij de inhoud van de variabelen. Inhoudelijk gezien betekenen deze twee modelmodificaties dat besluitvorming t.a.v. het operationele personeelsbeleid in sterke mate overeenkomt met besluitvorming t.a.v. het operationele beheer. Operationeel personeelsbeleid heeft betrekking op dat deel van het personeelsbeleid dat direct slaat op het primaire proces van de school (de kernactiviteiten). Het operationele beheer heeft betrekking op beslissingen over aanschaf van apparatuur en software ten behoeve van het

Tabel 5. Resultaten modelanalyse.

Model	CHI-SQ	d.f.	GFI	AGFI	RMR	AIC
1.0	4104.16	90	0.98	0.97	0.34	4164.16
2.0	3292.09	89	0.98	0.98	0.30	3354.09
2.1	3101.07	88	0.98	0.98	0.29	3165.07
2.2	2940.45	87	0.99	0.98	0.29	3006.45
3.0	2905.95	87	0.99	0.98	0.29	2971.92
3.1	2643.76	86	0.99	0.98	0.27	2711.76
3.2	2418.26	85	0.99	0.98	0.26	2488.26
4.0	3502.00	84	0.98	0.98	0.31	3574.00

Noot. $n=298$.

GFI: goodness of fit index
 AGFI: adjusted goodness of fit index
 RMR: Root Mean Square Residual
 AIC: Akaike's Information Criterion

criterium:

zo dicht mogelijk bij 1.00
 zo dicht mogelijk bij 1.00
 zo klein mogelijk
 zo klein mogelijk

primaire proces. Uit de analyse blijkt dat er zodanige overlap in deze besluitvormingsonderwerpen zit, dat wij genooddaakt worden een deel als gemeenschappelijk op te vatten.

Uit tabel 5 blijkt dat de stap van model 3.0 of model 3.2. naar 4.0 een achteruitgang betekent. Bij de stap van model 3.2 naar model 4.0 neemt de chi-kwadraat flink toe (ongeveer 1100 punten, bij een afname van 1 vrijheidsgraad). Dit betekent dat model 4.0 geen betere beschrijving van de data oplevert dan model 3.2. Dezelfde redenering geldt voor de stap van model 3.0 naar model 4.0. De conclusie is dat model 3 de voorkeur heeft boven model 4.

In figuur 3 wordt het volledige en definitieve model gepresenteerd.

Nadere gegevens van de factoren

De drie factoren van model 3.0 kennen een onderlinge samenhang die er als volgt uit ziet (zie tabel 6):

Hetgeen opvalt is dat de correlaties tussen de factoren hoog zijn. De betekenis daarvan is dat indien een school een hoge decentralisatiescore heeft in het professionele domein, de school ook een hoge score heeft in het onderhandelingsdomein. Hetzelfde geldt voor de beide andere domeinen.

De schaal van ieder van de factoren is gelijkgesteld aan die van de meest kenmerkende activiteit. Hierdoor kan de variantie worden geïnterpreteerd op de schaal van 0 tot 10. De meeste scholen variëren binnen een interval van twee maal de standaarddeviatie rond het gemiddelde. Opvalt dat scholen weinig verschillen in de mate waarin besluitvorming rondom activiteiten die in het operationele domein vallen zijn georganiseerd. De variantie binnen het strategische - en het onderhandelingsdomein is aanzienlijk groter.

6. CONCLUSIES EN DISCUSSIE

In dit artikel staat de vraag centraal in hoeverre de mate van decentralisatie binnen schoolorganisaties als een één-dimensioneel of als een meer-dimensioneel begrip opgevat dient te worden. Om deze vraag te beantwoorden zijn diverse modellen over besluitvorming in schoolorganisa-

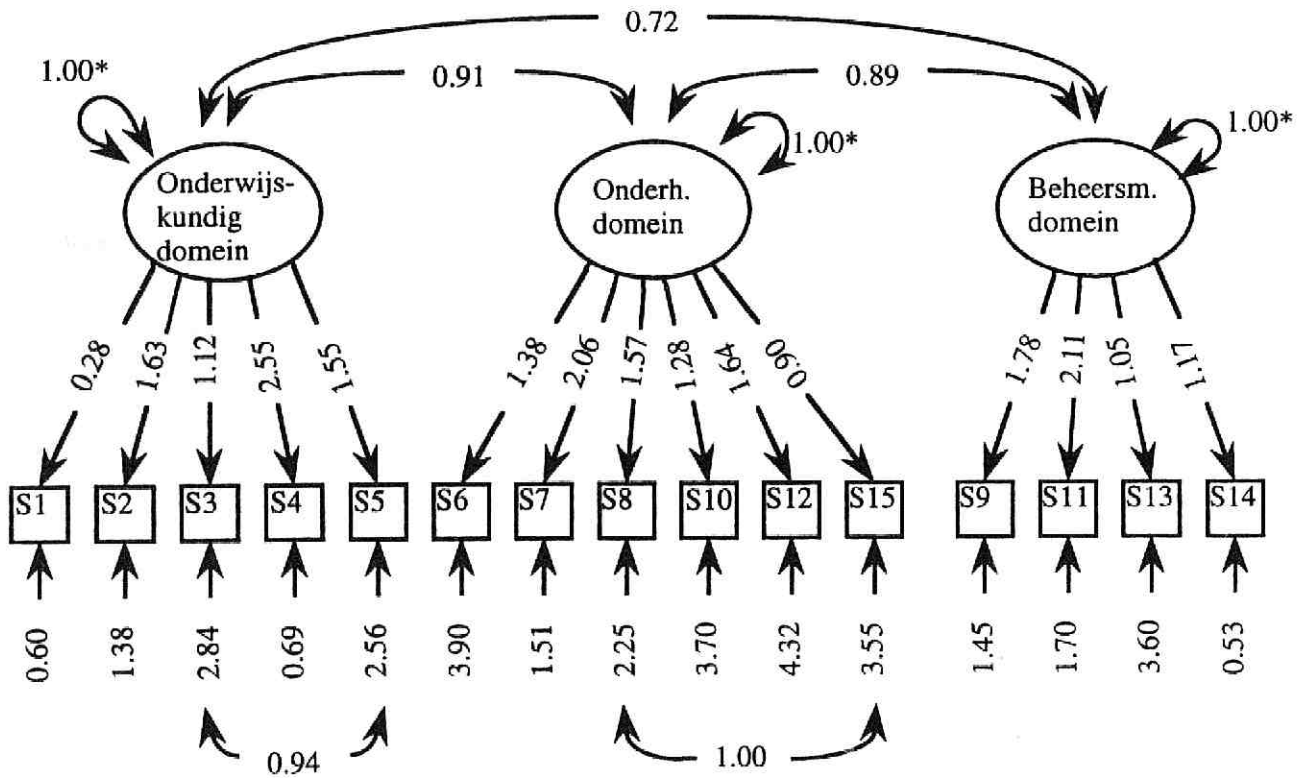


Fig. 3. Het drie-domeinenmodel (factormodel).

Tabel 6. Correlaties van de factoren uit model 3.0.

	Onderwijskundig	Onderhandeling	Beheersmatig
Onderwijskundig	1.00 (.08)		
Onderhandeling	.91	1.00 (2.46)	
Beheersmatig	.72	.89	1.00 (1.37)

Noot. n=298.

Tussen haakjes staan de varianties van de afzonderlijke factoren.

ties die bekend zijn uit de literatuur beschreven. Vervolgens is nagegaan of deze modellen, indien geconfronteerd met de verzamelde gegevens van de MBO-scholen, overeenind blijven. Of in andere termen, welk conceptueel model voor besluitvorming in schoolorganisaties in het MBO blijkt het best de data te beschrijven?

Uit het onderzoek kan geconcludeerd worden dat het drie domeinen model de beste weergave voor de besluitvormingsstructuren in het MBO biedt. De voornaamste opbrengst van het model is dat duidelijk is geworden dat een scala aan besluitvormingsonderwerpen feitelijk een drietal achterliggende factoren bevat. Deze achterliggende factoren betreffen de theoretische specificaties van het drie domeinen model: een onderwijskundig domein, een beheersmatig domein en een onderhandelingsdomein.

Tabel 7 geeft schematisch weer hoe de indeling eruit ziet

Tabel 7. Overzicht variabelen per besluitvormingsdomein

Onderwijskundig domein	Onderhandelingsdomein	Beheersmatig domein
1. micro-didactische aspecten	6. organisatie van stages	9. strategisch personeelsbeleid
2. leerroutes en keuzemogelijkheden leerlingen	7. werkorganisatie	11. informatievoorziening m.b.t. docenten
3. organisatie van proefwerken	8. operationele personeelsbeleid	13. pr/werving van leerlingen
4. roostering van vakken	10. informatievoorziening m.b.t. leerlingen	14. strategisch beleid
5. modulering van de opleiding	12. externe contacten van opleidingen	
	15. operationele beheer	

Plaatsen wij deze bevindingen in het licht van de kritiek van Imants (1992) op het onderzoek van Slegers (1991) dan merken wij het volgende op. Imants interpreteert Hanson wat betreft het onderwijskundig domein voor scholen in het voortgezet onderwijs heel strikt: precies dat deel van de werkzaamheden van docenten waarover geen overleg of afstemming op schoolniveau plaats vindt. Het is de vraag in welke mate deze situatie ook voor het MBO geldt. De organisatiekundige situatie in het MBO is meer gedifferentieerd dan in het voortgezet onderwijs in die zin dat opleidingen en afdelingen als organisatiekundige eenheden functioneren. In onze optiek is het te verdedigen om besluitvorming die afstemming op opleidingsniveau vereist, toch tot het onderwijskundige domein te rekenen. Afspraken binnen het onderwijskundige domein betreffende besluiten over het primaire proces/kernactiviteiten vereisen immers geen afstemming op schoolniveau, maar blijven dicht tegen het docentenniveau aanzitten. Hierbij valt te denken aan besluiten over micro-didactische aspecten, de roostering van vakken en de organisatie van proefwerken. Besluitvorming overstijgt niet de eenheid van de opleiding of afdeling en betreft onderwerpen die dicht tegen de taakuitvoering van docenten aan ligt.

In het onderzoek is gebruik gemaakt van coördinatoren als 'informanten'. De theoretische aanzet gaat vooral over de invloed van managers versus die van professionals op de besluitvorming over verschillende onderwerpen. Dit duidt op het vraagstuk van tot welke groep de coördinator eigenlijk behoort. In de literatuur worden ze aangeduid als middenmanagers: ze komen uit de groep van professionals (docenten) en hebben binnen de groep een managementfunctie. Dit betekent dat ze dit deel van de organisatie goed kennen. Wel is relevant dat ze doorgaans geen leidinggevende macht hebben over de groep professionals. Door het topmanagement worden ze min of meer gebruikt als aansturingspunt of loopjongen. In die zin zijn ze op de hoogte van de managementzijde van de schoolorganisatie. Bij elkaar genomen vormen zij dus een scharnierpunt in de organisatie en zijn ze goed op de hoogte van beide partijen. De keuze voor coördinatoren als informatiebron is gemaakt om de vergelijkbaarheid tussen organisaties te optimaliseren. Zij zitten op een vergelijkbare sleutelposities binnen de onderwijsinstellingen. De informatiewaarde van het onderzoek zou zijn verhoogd indien ook directieleden waren betrokken bij het onderzoek. De keuze om ook informanten uit de directiestructuur te bevragen is negatief uitgevallen vanwege de tijd en kosten. Bovendien is het de vraag of een optimale overeenstemming tussen coördinatoren en directieleden van dezelfde school haalbaar is via schriftelijk vragenlijstonderzoek gezien de geringe bereidwilligheid van directies mee te werken.

Eerder merkten wij al op dat de correlaties tussen de achterliggende factoren hoog zijn, met name tussen de scores op het onderwijskundig en onderhandelingsdomein. Dit betekent niet dat de domeinen identiek zijn. Wel zegt het dat de mate van decentralisatie binnen beide domeinen in hoge mate gelijk opgaat. In de modeltoetsingen is duidelijk geworden dat een driedeling moet worden gehandhaafd: de huidige factoroplossing geniet verreweg de voorkeur boven alle andere modellen.

De eindconclusie luidt dat het conceptuele model van Hanson (1991) met deze resultaten een empirische onderbouwing heeft gekregen voor wat betreft de besluitvormingsstructuren in het middelbaar beroepsonderwijs. Naast de toetsing van theoretische modellen denken wij dat de vragenlijst (zie bijlage), de indeling in schalen en de gepresenteerde gegevens hun nut kunnen bewijzen in toekomstig onderzoek. Ten slotte hebben de gegevens betrekking op 87 van de 100 MBO-instellingen met sectoren techniek en economie.

DANKBETUIGING

Met dank aan dr. J. Imants voor zijn commentaar op een eerdere versie.

REFERENTIES

- Amelsvoort, H.W.C.H., van & Witziers, B. (1995). Deregulering en de marktpositie van scholen. *Tijdschrift voor Onderwijsresearch*, 20, 115-132.
- Ax, J. (1993). 'Organisatiestructuur'. In P.J.J. Stijnen, J. Scheerens, A.M.L. van Wieringen & H.L.W. Münstermann (red.). *Transformatie van schoolorganisaties*. Alphen a/d Rijn: Samsom H.D. Tjeenk Willink.
- Ax, J., Bosma, Y. & Wieringen, A.M.L., van, (1993). *Funding and policy-making in colleges for vocational education*. Amsterdam: SCO-Kohnstamm Instituut. (rapport nr. 346).
- Bacharach, S.B., Bamberger, P., Conley, S.C., & Bauer, S. (1990). The dimensionality of decision participation in educational organization: The value of a multi domain evaluative approach. *Educational Administration Quarterly*, 26, 126-167.
- Bechger, T.M. & Hox, J.J. (1994). *Covariance structure analysis with LISREL8*. Universiteit van Amsterdam: POW (MLS-publ.nr.53).
- Boerman, P.L.J. (1995). Autonomieverschuivingen in het beroepsonderwijs. *Nederlands Tijdschrift voor Onderwijsrecht en Onderwijsbeleid*, Juni, 57-68.
- Bollen, K.A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.
- Butler, R. (1991). *Designing organizations. A decision making perspective*. London: Routledge & Kegan Paul.
- Hanson, E. (1991). *Educational administration and organizational behavior*. (3rd edn; 1st edn: 1979). Boston: Allyn & Bacon.
- Hoeksema, L.H. (1995). *Learning strategy as a guide to career succes in Organizations*. (3rd ed.). Leiden: DSWO Press, Leiden University.
- Hover, C. (1988). *Consistent configureren. Modellen van onderwijsorganisatie*. Den Haag: HBO-raad.
- Imants, J. (1992). Boekbespreking Slegers (1991). *Tijdschrift voor Onderwijsresearch*, 17, 260-266.
- Jöreskog, K. & Sörbom, D. (1993). *LISREL8. Structural equation modeling with the SIMPLIS command language*. Chicago: SSI.
- Kastelein, (1990). *Modulair organiseren. Tussen autonomie en centrale beheersing*. Groningen: Wolters-Noordhoff.
- Kerlinger, F.N. (1986). *Foundations of behavioral research*. London: Holt, Rinehart & Winston.
- Kochen, M. & Deutsch, K.W. (1980). *Decentralization. Sketches toward a rational theory*. Cambridge, MA: Oelgeschlager, Gunn & Hain.
- Krogt, F.J., van der & Oosting, J. (1988). *Relaties tussen scholen en hun omgeving. Een onderzoek naar de externe contacten en het arbeidsmarktbeleid van MBO-scholen*. Enschede: Universiteit Twente/OCTO.
- Leeuw, A.C.J., de (1988). *Organisaties: Management, analyse, ontwerp en verandering; een systeemvisie*. (3e druk: 1e druk: 1982). Assen/Maastricht: Van Gorcum.
- Leeuw, A.C.J., de (1994). *Besturen van veranderingsprocessen. Fundamenteel en praktijkgericht management van organisatieveranderingen*. Assen: Van Gorcum.
- Mintzberg, H. (1979). *The structuring of organizations. A synthesis of the research*. Englewood Cliffs: Prentice Hall.
- Moch, M., Cammann, C. & Cooke, R.A. (1983). Organizational structure: Measuring the distribution of influence. In S.E. Seashore, E.E. Lawler, P.H. Mirris, & C. Cammann (eds.). *Assessing organizational change. A guide to methods, measures, and practices* (pp.177-201). New York: John Wiley.

- Pelkmans, A. & Vries, B. de, (1992). *De V van SVM. Inventarisatie van vernieuwingen in onderwijs en organisatie in de SVM-instellingen voor Middelbaar beroepsonderwijs (1990-1991)*. Nijmegen: ITS.
- Pugh, D., Hickson, D., Hinnings, C. & Turner, C. (1968). Dimensions of organization structure. *Administrative Science Quarterly*, 12, 65-105.
- Seashore, S.E., Lawler, E.E., Mirris, P.H., & Cammann, C. (eds.) (1983). *Assessing organizational change. A guide to methods measures, and practices*. New York: John Wiley.
- Sleegers, P. (1991). *School en beleidsvoering*. Nijmegen: Katholieke Universiteit Nijmegen. (dissertatie).
- Vermculen, M. & Visser, H.J. (1992). *Personeelswerk in MBO-colleges*. Amsterdam: SCO.
- Visscher, A.J. (1993). *Design and evaluation of a computer-assisted management information system for secondary schools*. Enschede: Universiteit Twente. (dissertatie).
- Wieringen, A.M.L., van (1989). *Bestuur en management van onderwijsinstellingen*. Groningen: Wolters-Noordhoff.
- Witziers, B. (1992). *Coördinatie binnen scholen voor voortgezet onderwijs*. Enschede: Universiteit Twente. (dissertatie).

Ontvangen 28-11-1995

Definitief aanvaard 20-6-1996

BIJLAGE 1

OVERZICHT BESLUITVORMINGSITEMS UIT DE VRAGENLIJST

1. Aanschaf van leermiddelen binnen de opleiding
2. Keuze van leermiddelen
3. Invoering concrete leerstof (per vak)
4. Verandering van opleidingsroutes of leerwegen
5. Afstemming van de inhoud van het studieprogramma op specifieke wensen bedrijfsleven
6. Werkvormen en didactische aanpak
7. Normering van proefwerken
8. Inhoud van proefwerken
9. Frequentie van proefwerken
10. Totstandkoming van rapportcijfers
11. Individuele begeleiding/ondersteuning van leerlingen
12. De plaats van de stages binnen het opleidingstraject
13. Hoeveelheid huiswerk
14. Evaluatie van leerlingresultaten
15. Roostering van vakken binnen de opleiding
16. Roostering van proefwerken
17. Groepering van leerlingen
18. Modulering: *begin- en eindtijden* van modules in de opleiding
19. Modulering: *omvang* in punten van modules in de opleiding
20. Modulering: *inhoud* van de modules in de opleiding
21. Voorkomen lesuitval
22. Keuzemogelijkheden leerlingen/cursisten
23. Aantal afstudeerrichtingen (uitstroomprofielen)
24. Taakverdeling over docenten m.b.t. stages
25. Coördinatie van stagecontacten
26. Afspraken met instellingen over stagecontacten
27. Vergaderfrequentie opleidingsteam
28. Groepering van het onderwijspersoneel binnen de sector
29. Inzet van het aantal docenten bij de opleiding
30. Mate van zelfstandigheid van de opleidingen
31. Scholing/deskundigheidsbevordering van het onderwijspersoneel
32. Taakspecialisatie
33. Taakverruiming
34. Carrièreperspectief van het onderwijspersoneel
35. Flexibilisering inzet onderwijspersoneel
36. (Extra of tijdelijk extra) Beloning van het personeel
37. Aanneمة/selectie van het personeel
38. Verhouding lesgevende / niet-lesgevende taken
39. Initiatief tot uitvoering van functionerings- en beoordelingsgesprekken
40. Vormgeving aan uitvoering van functionerings- en beoordelingsgesprekken
41. Terugkoppeling en bespreking leerlingresultaten
42. Voortgangsregistratie van leerlingen
43. Registratie verzuim van leerlingen
44. Evaluatie functioneren van docenten
45. Registratie taakbelasting van docenten
46. Ziekteregistratie docenten
47. Onderwijsbeleid op lange termijn ('visie-ontwikkeling')
48. Contacten met directies andere scholen
49. Contacten met landelijke opleidingscommissies

50. Contacten met verwante opleidingen
51. PR- en reclame-activiteiten
52. Werven van leerlingen
53. Gebruik leerlinggegevens (leerprestaties, doorstroom, arb.m.situatie enz.) voor beleidsontwikkeling
54. Gebruik personeelsgegevens (ervaring, leeftijd, ziekteverzuim) voor beleidsontwikkeling
55. Middelentoewijzing aan afdeling (hoogte budget)
56. Huisvesting
57. Toerusting praktijklokalen
58. Bibliotheek/mediatheek
59. Aanschaf apparatuur
60. Aanschaf software

BIJLAGE 2

VARIABELENCONSTRUCTIE

Factoranalyse (varimax-rotatie) op de 60 items levert een 15-factoroplossing op, overeenkomstig de voorgestelde structuur. Nagegaan is in hoeverre de items die op de factoren laden betrouwbare schalen vormen. De betrouwbaarheid van de schalen is bepaald met behulp van Cronbachs alfa, een maat voor de interne consistentie of homogeniteit van een schaal. Een hoge alfa (liefst 0.80 of hoger) geeft de mate aan waarin de geformuleerde items betrekking hebben op hetzelfde fenomeen (Swanborn, 1987). De alfa-waarden van de schalen staan in tabel B1.

Op grond van de betrouwbaarheid van de schalen is besloten samengestelde variabelen te construeren.

Per schaal zijn gemiddelde somscores per respondent berekend. Het voordeel van het middelen van somscores is dat de oorspronkelijke scoringsschaal in tact blijft. Voor de analyses in dit artikel is gebruikt gemaakt van de samengestelde variabelen.

Een beschrijving van de samengestelde variabelen is als volgt (zie tabel B2).

Tabel B1. Betrouwbaarheidscoëfficiënten van de schalen.

Schalen	α	# items
1. didactiek-micro	.72	6
2. leerroute- en keuzemogelijkh. leerlingen	.77	4
3. toetsorganisatie	.75	3
4. roostering vakken	.80	3
5. modulering	.85	3
6. stage-organisatie	.82	3
7. werkorganisatie	.80	4
8. operationeel personeelsbeleid	.84	3
9. strategisch personeelsbeleid	.83	4
10. infomanagement leerlingen	.80	3
11. infomanagement docenten	.84	3
12. externe contacten opleidingen	.85	2
13. pr/werving leerlingen	.75	2
14. strategisch beleid & beheer	.76	6
15. operationeel beleid & beheer	.83	4

Noot. $n=298$.

Tabel B2. Beschrijving samengestelde variabelen.

Samengestelde variabelen	Beschrijving
1. didactiek-micro	Didactiek in de micro-situatie heeft te maken met die activiteiten die zich in de klassituatie afspelen tussen docent en leerling. Het gaat om beslissingen m.b.t. de soorten activiteiten: keuze van leermiddelen, invoering van concrete leerstof, gebruikte werkvormen en inhoud van proefwerken (vgl. Witziers, 1992).
2. leerroute/keuzemogelijkheden leerlingen	De leerroute heeft betrekking op de wijze waarop leerlingen het studieprogramma doorlopen. Beslissingen hierover hebben betrekking op de keuzemogelijkheden die leerlingen worden geboden (aantal zgn. uitstroomprofielen) en beslissingen over aanpassingen aan specifieke wensen (vgl. Pelkmans & de Vries, 1992).
3. toetsorganisatie	Toetsorganisatie heeft te maken met beslissingen over frequentie en roostering van proefwerken (vgl. Witziers, 1992).
4. roostering vakken	Roostering heeft betrekking op beslissingen ten aanzien van het lesrooster: hoe worden vakken geroosterd en hoe worden leerlingen gegroepeerd (vgl. Witziers, 1992).
5. modulering	Beslissingen t.a.v. modulering heeft te maken met de organisatie van de modules: wie beslist er over begin/eindtijden, de omvang en de inhoud van de modules (vgl. Pelkmans & de Vries, 1992).
6. stage-organisatie	De stage-organisatie betreft beslissingen over taakverdeling en coördinatie t.a.v. de stages (vgl. Pelkmans & de Vries, 1992).
7. werkorganisatie	Beslissingen t.a.v. de werkorganisatie heeft betrekking op de vergaderfrequentie van het opleidingsteam, toewijzing van docenten aan klassen en de verhouding lesgevende / niet-lesgevende taken (vgl. Hover, 1988).
8. operationeel personeelsbeleid	Met operationeel personeelsbeleid wordt dat deel van het personeelsbeleid bedoeld dat direct betrekking heeft op het primaire proces (lesgeven als kernactiviteit) van de schoolorganisatie. Het betreft hier beslissingen m.b.t. taakverruiming, taakspecialisatie en scholing/deskundigheidsbevordering (vgl. Vermeulen & Visser, 1992).
9. strategisch personeelsbeleid	Met strategisch personeelsbeleid wordt dat deel van het personeelsbeleid bedoeld dat betrekking heeft op de hoofddoelstellingen van de schoolorganisatie. Het gaat dan om beslissingen m.b.t. carrièreperspectief, beloning en selectie van personeel (vgl. Vermeulen & Visser, 1992).
10. infomanagement leerlingen	Deze variabele verwijst naar beslissingen m.b.t. informatiemanagement t.a.v. leerlinggegevens zoals voortgangsregistratie en absentie (vgl. Visscher, 1993).
11. infomanagement docenten	Informatiemanagement t.a.v. docentgegevens heeft betrekking op beslissingen t.a.v. registratie taakbelasting, functioneren en ziekteregistratie van docenten (vgl. Visscher, 1993).
12. externe contacten op opleidingsniveau	Hierbij gaat het om beslissingen t.a.v. externe contacten op opleidingsniveau (Van de Krogt & Oosting, 1988).
13. pr/werving	Hierbij gaat het om beslissingen t.a.v. pr- en reclame activiteiten en werving van leerlingen (Van de Krogt & Oosting, 1988).
14. strategisch beleid	Bij strategisch beheer gaat het om beslissingen t.a.v. allocatie van middelen aan de opleiding en huisvestingskosten (Pelkmans & de Vries, 1992).
15. operationeel beheer	Bij operationeel beheer gaat het om beslissingen over de aanschaf van apparatuur en software t.a.v. het onderwijs (als kernactiviteit) (Pelkmans & de Vries, 1992).

In de tabel B3 zijn de correlaties van de samengestelde variabelen afgedrukt.

Tabel B3. Correlatie matrix samengestelde variabelen s1 to s15.

S1	—														
S2	.33**	—													
S3	.37**	.40**	—												
S4	.25**	.69**	.56**	—											
S5	.41**	.56**	.56**	.66**	—										
S6	.26**	.43**	.19**	.53**	.35**	—									
S7	.21**	.65**	.46**	.76**	.53**	.50**	—								
S8	.17*	.55**	.28**	.58**	.39**	.41**	.68**	—							
S9	.09	.49**	.28**	.56**	.37**	.36**	.64**	.64**	—						
S10	.34**	.38**	.41**	.53**	.43**	.37**	.45**	.26**	.38**	—					
S11	.10	.53**	.31**	.64**	.40**	.38**	.65**	.56**	.74**	.52**	—				
S12	.16*	.56**	.29**	.53**	.42**	.38**	.50**	.45**	.43**	.24**	.40**	—			
S13	.03	.24**	.12	.34**	.23**	.32**	.28**	.27**	.40**	.35**	.42**	.33**	—		
S14	.07	.58**	.27**	.58**	.43**	.35**	.65**	.60**	.70**	.36**	.63**	.59**	.41**	—	
S15	.16*	.33**	.11	.29**	.19**	.32**	.39**	.51**	.36**	.08	.27**	.37**	.16*	.44**	—
S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	

Noot. n = 298.
N of cases: 298 1-tailed Signif: * - .01 ** - .001

BULLETIN

ORGAAN VAN DE VERENIGING VOOR ONDERWIJSRESEARCH

Redactie-adres

Dr. M. Mulder,
Faculteit der Toegepaste Onderwijskunde,
Universiteit Twente,
Postbus 217, 7500 AE Enschede,
tel. 053-4893652
fax 053-4893759

Jrg. 20, nr. 2

1996

Inhoudsopgave

Jaarverslag VOR 1995-1996	368
Verslag VOR ledenvergadering	370
ORD in perspectief	373
Lezing ORD 96	378

Jaarverslag VOR 1995-1996

Dr. Martin Mulder, secretaris

1. Algemene gegevens

De VOR telt ultimo 1995 ruim 550 leden. Namens de divisie B&BO is Dr. A.M. Versloot toegetreden tot het bestuur, namens de divisie Curriculum Prof.dr. J. van den Akker en namens de divisie Methodologie en Evaluatie heeft Prof.dr. J. Scheerens plaats gemaakt voor Dr. P. Sanders. Dr. J. Stijnen en Dr. F. van Bussel hebben afscheid genomen van het DB en zij zijn opgevolgd door Dr. Tj. Imbos (penningmeester) en Dr. M. Mulder (secretaris).

Wat betreft de financiële situatie van de VOR betreft, kan worden opgemerkt dat de reserves van de vereniging gering zijn en de lasten hoog. Daarom is voorgesteld divisiebudgetten weliswaar te oormerken, maar om tegelijkertijd 50% van de voor divisies bestemde budgetten te reserveren om beleidsruimte voor de VOR als geheel te creëren. Als de voor de divisies bestemde budgetten volledig zouden worden geoormerkt voor divisie-activiteiten, dan zou het verenigingsvermogen volledig zijn besteed. Door deze maatregel door te voeren, krijgen de divisies redelijke budgetten ter beschikking. De ervaring leert dat de beschikbare divisiebudgetten nauwelijks worden aangesproken. De verwachting is derhalve dat dit beleid niet op problemen zal stuiten bij de divisies.

2. VOR-TOR

Tijdens de ALV van de VOR op 20 juni 1995 is het bestuursstandpunt in verband met de positie van TOR vastgesteld. Dit standpunt luidt:

- Het TOR blijft het verenigingsblad van de VOR. De VOR garandeert het voortbestaan van TOR. Met de uitgever is afgesproken dat de VOR een afname van 300 exemplaren per jaar garandeert.
- De redactie van TOR zal er naar streven het karakter van TOR – onder handhaving van de huidige kwaliteit – zo te wijzigen, dat het blad aantrekkelijk wordt voor meer leden van de vereniging. De divisies zijn gevraagd contactpersonen voor te dragen die de band tussen de divisies en het TOR kunnen verstevigen.
- De leden van de VOR krijgen, desgewenst de mogelijkheid TOR niet te ontvangen. De reductie die zij daarvoor krijgen kan desgewenst gebruikt worden voor een abonnement op een ander blad. Deze regeling gaat in per 1 januari 1997.

Er is een subsidie voor het TOR verkregen van het Ministerie van OC en W van f 5000,- per jaar gedurende een periode van drie jaren. Er wordt geconstateerd dat de kosten voor het TOR zwaar drukken op de VOR-begroting.

3. Opheffing SVO

Het bestuur heeft uitgebreid van gedachten gewisseld over de consequenties van de opheffing van SVO voor de vereniging. Het bestuur heeft getracht invloed uit te oefenen op het Ministerie van OC en W om zo gunstig mogelijke overgangscondities te scheppen. De volgende punten waren in dezen van groot belang:

- Er moet een oplossing komen voor eventuele wachtgeldaanspraken.
- NWO moet een overgangsperiode inlassen om de verschillen in tariefstelling tussen NWO en SVO te kunnen overbruggen.
- De zeggenschap over de verdeling van de onderzoeksmiddelen dient in overeenstemming te zijn met het volume aan onderzoeksgeld uit de verschillende disciplines waarover de stichting, waarin het onderzoek wordt ondergebracht, zal beschikken.
- De programmering van het onderzoek dient te worden uitgevoerd door deskundigen.
- De verspreiding van de onderzoeksresultaten moet worden gewaarborgd door NWO.

4. ORD

De organisatiecommissie van de ORD te Tilburg heeft voorgesteld het format van de ORD-sessies te vernieuwen. Het voornemen is te streven naar meer interactie tussen de onderzoekers en hun doelgroepen. Het bestuur van de VOR heeft aan de organisatiecommissie van de ORD te Tilburg voorgesteld ook toe te staan dat individuele paperpresentaties worden georganiseerd. De organisatiecommissie heeft dit verzoek ingewilligd.

Besloten is de ORD 97 te organiseren in Leuven. Voorzitter zal zijn Prof.dr. E. De Corte.

5. EERA/ECER

De ontwikkeling van de EERA verloopt voorspoedig. Een probleem blijft de toezending van het EERA Bulletin. De kosten van deze nieuwsbrief zijn te hoog voor de VOR om alle leden individueel een exemplaar te sturen. De VOR krijgt 50 exemplaren toegestuurd, waarvan er vijf naar iedere divisie gaan. De resterende exemplaren gaan naar bestuursleden.

Prof. dr. N. Lagerweij is benoemd tot programmavoorzitter van de ECER '96 conferentie in Sevilla. Er zijn verschillende themacoördinatoren benoemd. De netwerken die zijn voorgesteld vanuit de divisies Curriculum en B&BO zijn goedgekeurd. Voorzitters zijn respectievelijk Prof.dr. J. van den Akker en Dr. M. Mulder.

6. VOR-congres

Gesproken is over het feit dat er afgezien van de ORD geen activiteiten meer plaatsvinden die voor alle VOR-leden interessant zijn. Dat vindt het bestuur een gemis. Er is derhalve een initiatief genomen om te bezien of de Vakgroep Onderwijskunde van de Universiteit Utrecht geïnteresseerd zou zijn een congres te organiseren dat tegelijkertijd interessant is voor veel VOR-leden en een maatschappelijk relevant thema omvat.

Verslag VOR-ledenvergadering 5 juni 1996: 12.19 – 13.30 uur KUB te Tilburg (tijdens ORD '96)

Aanwezig: Leden: M. de Bekker-Ketelaars, F. van Bussel, M. Geerlings, B. van Hout-Wolters, G. Kanselaar, P. Kirschner, J. van der Linden, P. Meijer, F.S.J. Rie-
mersma, A.M. Scholten, J.H. Slavenburg, P. Stijnen, K.J. Westerhof, A.M.L.
van Wieringen, Th. Wubbels
Bestuursleden: J. van den Akker, J. Ax, J. Dronkers, T. Imbos (penningmeester), N. Lagerweij
(voorzitter), M. Mulder (secretaris), N. Verloop

1. Opening

De voorzitter opent de vergadering om 12.19 uur.

Bericht van verhindering is ontvangen van H. van Hout, Mw. Peters-Sips, P. Span en R. Vietor.

2. Mededelingen

De voorzitter deelt mee dat het bestuur heeft besloten de Stichting Onderwijs Research (SOR) op te heffen. De SOR is ooit in het leven geroepen om financiële risico's van de ORD af te dekken. De organisatie van de ORD is echter al jaren de verantwoordelijkheid van de organiserende instelling. De VOR-prijs wordt er ook uit betaald. Het reservebedrag (ca. f 6.000,-) wordt toegevoegd aan het verenigingsvermogen. De voorzitter constateert dat de vergadering met dit besluit instemt.

3. Verslag van de VOR-vergadering van 20 juni 1995

Het verslag wordt ongewijzigd goedgekeurd.

4. Jaarverslag

De voorzitter memoreert dat dit het eerste jaar is dat de VOR een jaarverslag heeft gemaakt. Het bevat een beschrijving van de hoofdpunten waaraan het bestuur heeft gewerkt. Besloten wordt het jaarverslag na correctie te publiceren in het VOR-bulletin. Stijnen merkt op dat hij het doen verschijnen van een jaarverslag een goed initiatief vindt.

5. Vaststellen balans en winst- en verliesrekening 1995

De stukken worden ter vergadering uitgereikt en door de penningmeester toegelicht. Hij wijst op het nadelig saldo van f 4706,09, maar stelt tevens dat dit tekort was voorzien en dat daarom ook in 1995 is besloten de contributie te verhogen. Tevens wijst hij op het feit dat het TOR loodzwaar op de begroting drukt en dat ook het lidmaatschap van de EERA kosten met zich meebrengt, maar dat dit een nuttige besteding is. Voorts stelt hij dat de kosten van divisies verschillen.

Van Bussel merkt op dat het gebruikelijk was het promovendi-overleg te subsidiëren, maar dat hij aan de winst- en verliesrekening niet kan zien dat de kosten hiervoor ook inderdaad zijn gemaakt. De penningmeester antwoordt hierop dat het promovendi-overleg geen verzoek om subsidie heeft ingediend, maar er wel recht op heeft. Misschien komt er nog een verzoek met terugwerkende kracht.

Wat betreft de balans wordt opgemerkt dat de beschikbare reserves voor de helft naar de divisies zijn gegaan en voor de helft naar de algemene reserves van de VOR. Tenslotte wordt erop gewezen dat divisies de naam van hun divisie goed moeten aangeven op declaraties bij de VOR. De vergadering stelt de balans en winst- en verliesrekening 1995 vast.

6. Rapport kascommissie en decharge kascommissie

Stijnen rapporteert namens de kascommissie en deelt mee dat de commissie van mening is dat de penningmeester uitstekend werk heeft verricht, en dat dit een goede voortzetting is van het werk van de vorige penningmeester. De commissie wil graag drie opmerkingen maken:

- de uitgaven voor TOR zijn onverantwoord hoog;
- declaraties van divisies zijn soms onduidelijk;
- de inkomsten uit renten zouden door beter cash-flow management kunnen worden verhoogd.

Opgemerkt wordt dat er zeker aandacht is voor het cash-flow management. Beide andere punten hebben ook de aandacht van het bestuur.

De voorzitter bedankt de kascommissie voor de verrichte werkzaamheden en stelt voor haar te dechargeren. Aldus wordt besloten.

7. Instellen nieuwe kascommissie

De voorzitter draagt dezelfde personen (mevr. Peters-Sips en de heren Stijnen en Van Wieringen) voor de nieuwe kascommissie. De vergadering neemt dit voorstel over en de voorzitter polst de aanwezigen of zij bereid zijn wederom de kascommissie te vormen. Zij stemmen hiermee in en worden bij acclamatie benoemd.

8. Vaststellen begroting 1997

De begroting voor het jaar 1997 wordt toegelicht en vastgesteld.

9. Bestuursverkiezing

De voorzitter deelt mee dat Verloop de vergadering zal verlaten, hetgeen vervolgens gebeurt. Daarna deelt de voorzitter mee dat het bestuur Verloop voorstelt als kandidaat voor het voorzitterschap van het bestuur van de vereniging. De kandidaat heeft zich bereid verklaard een eventuele benoeming te aanvaarden. De vergadering wordt gevraagd of er bezwaren tegen deze kandidatuur zijn. Van Wieringen vraagt dat van de kandidaten wordt verwacht dat zij een beleidslijn voor de vereniging zullen presenteren. Dat hoeft niet tijdens de vergadering, maar dat kan uiteraard in een later stadium. Ter toelichting wordt gesteld dat het goed zou zijn als het bestuur een plan zou maken waarin is verwoord waar de vereniging naar toe gaat.

Er zijn geen bezwaren tegen de kandidatuur en de vergadering stemt in met de benoeming van Verloop tot voorzitter van de VOR.

Verloop komt vervolgens de vergaderzaal binnen en de voorzitter deelt het besluit van de vergadering mee. De scheidend voorzitter geeft de nieuwe voorzitter een uitnodiging mee een visie te ontvouwen over hoe het verder moet met de vereniging. Verder bedankt hij de aanwezigen, de secretaris en de penningmeester voor de prettige samenwerking en spreekt de wens uit dat de VOR meer aandacht zal hebben voor de maatschappelijke relevantie van de onderwijskunde en dat zij daarmee meer naar buiten zal treden. Vervolgens geeft hij de leiding aan de nieuwe voorzitter. Verloop dankt de vergadering voor het in hem gestelde vertrouwen. Vervolgens spreekt hij de scheidend voorzitter toe en wijst op zijn wens het maatschappelijk belang van de onderwijskunde meer te benadrukken. Verder deelt hij mee dat gezien zal worden of en zo ja hoe er een afscheids-symposium zal worden georganiseerd waarin de relatie tussen theorie en praktijk tot uitdrukking komt. Ook stelt hij dat de scheidend voorzitter de VOR een goede positie heeft bezorgd binnen de EERA. Tenslotte krijgt de scheidend voorzitter een bos bloemen aangeboden.

10. Kosten TOR

De voorzitter (Verloop) memoreert de afspraken die zijn gemaakt tijdens de vorige ALV over het TOR. Vervolgens krijgt Westerhof als redactie-secretaris van het TOR het woord over het beleid dat de TOR-redactie sedert de vorige ALV heeft gevoerd. Westerhof noemt zes punten:

1. per divisie zou een contactpersoon worden aangesteld. Er zijn inmiddels van vier divisies contactpersonen (Jungbluth voor Onderwijs en Samenleving, Terwel voor Curriculum, Hulshof voor Hoger Onderwijs en Mulder voor Beroepsonderwijs en Bedrijfsopleidingen);
2. Er zitten inmiddels vier themanummers in de pijplijn (op de domeinen van de divisies Leren en Instructie, Curriculum, Beroepsonderwijs en Bedrijfsopleidingen en Evaluatie en Methodologie);

3. Er wordt een begin gemaakt met overzichtsartikelen;
4. Er wordt actief geworven voor de rubriek notities en commentaren;
5. Er wordt gewerkt aan een leading article met commentaar daarop; het onderwerp is leerstijlen;
6. Aangezien er een nieuwe redacteur is voor de boekbesprekingen zullen de resultaten daarvan eerst in het volgend jaar zichtbaar worden.

De voorzitter constateert duidelijke vorderingen, maar aangezien de implementatie van het nieuwe beleid een paar jaar in beslag neemt zal dit punt tijdens de volgende ALV terug moeten komen.

Vervolgens wordt ingegaan op de kosten van het TOR. Tot voor kort berekende Swets & Zeitlinger f 90,- voor ieder abonnement. Het lidmaatschapsgeld van de VOR bedraagt f 135,-. Swets en Zeitlinger wil de abonnementsprijs verhogen naar f 115,- bij een minimumafname van 300 exemplaren. Dat vindt het VOR-bestuur onaanvaardbaar hoog. Er is daarom uitgekeken naar een andere uitgever. Het ziet er naar uit dat Swets en Zeitlinger geen problemen heeft met het beëindigen van het uigeeefcontract. Het is derhalve aantrekkelijk om een overgang naar een andere uitgever nader te exploreren. Het bestuur vraagt de vergadering mandaat te beslissen over deze overgang.

Verder stelt het bestuur voor de ontkoppeling van het lidmaatschap van de VOR en het abonnement op TOR een jaar uit te stellen. Dit hangt samen met de tariefstructuur voor het lidmaatschap en de overige tijdschriften waarop de leden zich met korting kunnen abonneren. Gedacht wordt aan een basisbedrag voor het VOR lidmaatschapsgeld van f 110,-. Voor een abonnement op het TOR komt daar dan f 25,- bij. De verwachting is dat er bij een dergelijke tariefstructuur weinig opzeggingen komen. Verder wordt gedacht aan andere tijdschriften waarop VOR-leden zich met korting kunnen abonneren. Divisies zullen worden uitgenodigd voor 1 januari 1997 voorstellen te doen voor een tijdschrift dat kan worden toegevoegd aan de lijst van tijdschriften waarop VOR-leden zich met korting kunnen abonneren. De divisies zullen desgewenst onderhandelingen moeten voeren met de uitgevers. Tijdens de ALV van 1997 kan dan een voorstel worden gedaan voor de tariefstructuur vanaf 1 januari 1998.

Na enige discussie besluit de vergadering in te stemmen met de bestuursvoorstellen. Er wordt wel opgemerkt dat er zou moeten worden omgezien naar mogelijkheden voor het elektronisch uitgeven van het TOR, dat op termijn wellicht het basisbedrag voor het VOR-lidmaatschap omlaag zou kunnen en dat het VOR-bestuur de onderhandelingen met uitgevers over aanvullende tijdschriften wellicht beter zelf zou kunnen doen om te voorkomen dat verschillende divisies eventueel met dezelfde uitgevers gaan onderhandelen.

11. Adressenbestand

Het voorstel van het bestuur is het adressenbestand van de VOR kosteloos ter beschikking te stellen voor VOR- en divisie-activiteiten en om voor alle overige activiteiten een bedrag van f 1,- per adres te berekenen. Uit de vergadering wordt naar voren gebracht dat alleen dan adressen ter beschikking moeten worden gesteld als de activiteit waarvoor het gebruik is bedoeld de doelstellingen van de vereniging ondersteunt. Er dient verder restrictief te worden omgegaan met het ter beschikking stellen van de adressen. Het bestuur zegt toe aldus te handelen en afgesproken wordt tijdens de ALV van volgend jaar te rapporteren over dit punt.

12. Rondvraag

De Jong merkt op dat de contacten tussen EERA en EARLI kunnen worden aangemoedigd. Kanselaar merkt op dat de EERA als doel heeft een vereniging van verenigingen te zijn, maar dat een individueel lidmaatschap nu ook mogelijk is.

Verder vraagt P. Meijer of er nog VOR-folders zijn. De secretaris merkt op dat ze op zijn en dat er wordt gewerkt aan nieuwe folders. De kosten voor een nieuw ontwerp bedragen ca. f 1.000,-. Het bestuur is wel van plan die uitgave te doen.

13. Sluiting

De voorzitter sluit de vergadering om 13.30 uur.

De Onderwijs Research Dagen in het perspectief van het begin

Welkomstwoord bij de opening van de ORD'96, op 4 juni 1996 te Tilburg

Prof. dr Hans G.L.C. Lodewijks

Onderwijs- & Opleidingspsychologie, Katholieke Universiteit Brabant, Tilburg

Het is een buitengewoon genoeg om u na 12 jaar weer welkom te mogen heten bij de opening van de Onderwijs Research Dagen aan de Katholieke Universiteit Brabant. Na de 'Tilburgse' ORD's van 1982 en 1984 is het nu de derde keer dat de KU Brabant het genoeg heeft om onderwijsonderzoekend Nederland en Vlaanderen te mogen ontvangen. We zijn nu in Tilburg toe aan de 23ste aflevering van de ORD en mogen dus vaststellen, dat de jaarlijkse toogdagen van het onderwijsonderzoek inmiddels kunnen bogen op een respectabele traditie, zeker als we bij die kwalificatie de betrekkelijke jonge leeftijd van ons wetenschapsgebied in acht nemen.

Het lijkt voor een gelegenheid als deze een goed idee om nog eens terug te kijken naar het begin, en wel naar de eerste Onderwijs Research Dag, op 29 maart 1974, die, zoals de seniores zich ongetwijfeld zullen herinneren, slechts één dag duurde en in de Amsterdamse RAI plaats vond. Daarom heb ik het verslag van die allereerste ORDag er nog maar eens bijgenomen, dat door Tjeenk Willink te Groningen werd uitgegeven in de serie "Empirische Studies over Onderwijs". De toenmalige Werkgroep Onderwijsresearch (WOR), waaruit een jaar later de Vereniging voor Onderwijs Research (VOR) zou voortkomen, was met medewerking en financiële steun van de toenmalige Stichting voor Onderzoek van het Onderwijs SVO, waarvan vele jaren later, zoals we inmiddels weten, en sommigen ook zeer betreuren, heel veel minder zou overblijven, verantwoordelijk voor deze uitgave.

Omdat het een blik achter de toenmalige coulissen van het schouwtoneel mogelijk maakt, is het heel curieus dat in dit conferentie-verslag op aandringen van SVO de passus werd opgenomen, dat het gepresenteerde geenszins een volledige representatie biedt van hetgeen er toen in Nederland aan onderwijsonderzoek werd verricht. Dit tekent de verhoudingen in het toenmalige landschap van de onderwijsresearch, en het laat ook zien, dat er – althans naar de mening van de SVO – te weinig aandacht op die eerste ORD was voor het onderzoek, dat via de SVO in uitvoering was.

Dit laatste had wellicht te maken met een heel ander probleem, namelijk met de vraag welk onderzoek de kwalificatie 'wetenschappelijk' verdiende en (dus?) geschikt was om te worden gepresenteerd op de ORD. Een probleem, dat toen – naar ik me herinner – ook bij de onderzoekers zelf leefde. Want wat er nou precies onder 'empirisch' c.q. 'wetenschappelijk' moest worden verstaan, was in die dagen lang niet voor iedereen zonneklaar. Daar werden toen in de nieuwe Vereniging voor OnderwijsResearch vaak ferme -en naar mijn oordeel, lang niet altijd even heldere- discussies over gevoerd.

Het was wat dit betreft maar goed, dat de toenmalige voorzitter van de VOR, Egbert Warries, dit probleem voor eens en voor altijd oploste, door de even eenvoudige als liberale vuistregel te hanteren, dat er sprake was van empirisch onderzoek, als de onderzoeker tenminste iets had gemeten. Meten hield dan in dat de onderzoeker in elk geval de moeite had genomen om iets te tellen, en – als het even meezat – daarna ook nog wat gemiddelden had uitgerekend.

Hoe dan ook, het eerste ORD-boek bevatte nog de volledige weergave van alle bijdragen, die tijdens het congres werden gepresenteerd, en zelfs waren de samenvattingen opgenomen van de papers die *niet* werden voorgedragen tijdens het congres.

Maar ook als we met een hedendaagse bril naar de bijdragen kijken en de maatstaven van vandaag aanleggen bij het beoordelen van het gebodene, kunnen we tot de conclusie komen dat

de meeste ORD'74-papers in ruime mate voldeden aan de eis dat het om empirisch onderzoek gaat. Dit kon ook nauwelijks anders, als we letten op degenen die toen in commissie-verband het wetenschappelijk gehalte van de voorstellen beoordeelden. Die commissie had als voorzitter A.D. De Groot, en bestond verder uit de collega's Brus, Hofstee, Mellenbergh en Vaags.

De voorstellen voor de eerste ORD moesten volgens deze commissie aan de volgende criteria voldoen:

- relevantie voor de onderwijsresearchdag;
- mate waarin er *gemeten* onderzoeksresultaten gepresenteerd konden worden;
- mate waarin het voorstel direct betrekking had op het onderwijs, en tenslotte
- minimale wetenschappelijke normen qua design en verwerking van het onderzoek.

Overigens haastte de beoordelings-commissie zich wel om hieraan onmiddellijk toe te voegen, dat "de voorstellen niet in even grote mate aan alle criteria behoefden te voldoen om voor selectie in aanmerking te komen", hetgeen te denken geeft over het vertrouwen, dat de WOR en de beoordelingscommissie toen hadden in de haalbaarheid van een 'full fledged' wetenschappelijk congres over de onderwijsresearch in Nederland.

De proceedings zelf bieden aardige lectuur voor iemand die geïnteresseerd is in de ontwikkelingen binnen de onderwijsresearch, de thema's die daarin centraal staan en de methoden die daarbij gebruikt worden. Wat dit betreft, zijn ze nog immer lezenswaardig. Bij de congres-thema's ging het in 1974 om:

- doelstellingen-analyse;
- beoordelingsmethoden in het onderwijs;
- onderzoeks-methodologie;
- individuele studie- en instructiesystemen;
- studietoetsen en toetsontwikkeling;
- evaluatie- & innovatie-onderzoek;
- onderzoek naar cognitieve functies;
- survey-onderzoek naar leerlingkarakteristieken;
- survey-onderzoek naar kenmerken van leerkrachten, en tenslotte om
- diverse aspecten van het leren.

Bij *doelstellingen-analyse* richtte men zich met name op onderzoek en toetsing van de bruikbaarheid van classificatie-systemen voor onderwijsuitkomsten. Dit soort onderzoek wierp zijn schaduw vooruit naar latere – veel minder empirisch geïnspireerde – discussies binnen het onderwijs en het onderwijsbeleid over zaken als kerndoelen, eindtermen, streefdoelen en dergelijke, maar het onderwerp zelf is toch als afzonderlijke benadering binnen de onderwijsresearch nu goeddeels verdwenen. Men zal in het ORD-programma van dit jaar danook vergeefs zoeken naar een voorbeeld van doelstellingen-onderzoek.

Het onderwerp *beoordelingsmethoden* ging over het meten van de studievoortgang in met name de praktijk van het medisch onderwijs. Het spreekt voor zich, dat dit thema veel verwantschap toonde met het doelstellingenthema. Ook dit thema is als afzonderlijk aandachtsgebied voor de onderwijsresearch sinds 1974 goeddeels verdwenen, althans in de programmering van de ORD's.

Bij het thema over *onderzoeks-methodologische kwesties* was de aandacht in 1974 onder meer gericht op de betrouwbaarheid, de validiteit en de bruikbaarheid van selectie-instrumenten in de propaedeuse van het wetenschappelijke onderwijs, een thema dat toen erg actueel was, omdat allereerst discussies rond het zogenaamde 'selectie-rapport' van A.D. De Groot in alle hevigheid woedden. Wat dit betreft is het overigens opvallend dat, nu dit thema van selectie aan de poort tot het Hoger Onderwijs recent weer op de onderwijspolitieke agenda is teruggekeerd, daarvan nauwelijks een neerslag is terug te vinden in het ORD-programma van dit jaar.

Verder ging het in 1974 bij het thema onderzoeks-methodologie over uiteenlopende inhoudelijke kwesties rond, of om methodische en statistische aspecten van Aptitude-Treatment-Interaction (ATI) onderzoek. De ATI-onderzoeksbenadering trok toen duidelijk veel belangstelling,

zeker ook van in oorsprong relatief fundamenteel-wetenschappelijk gerichte onderzoekers, die de relevantie van hun onderzoeksresultaten voor de praktijk van het onderwijs poogden te verbeteren. Een streven dat naar mijn mening samenhang met de snelle opkomst van de Onderwijswetenschappen, die hun praktische nut nog moesten bewijzen ten overstaan van nieuwe opdrachtgevers en subsidiënten, maar ook volledig in overeenstemming was met de optimistische geest van de tijd, waarin de samenleving – en dus ook het onderwijs – als maakbaar werden beschouwd.

In de jaren tachtig en negentig verdwijnt het ATI-onderzoek langzamerhand volledig naar de achtergrond, onder meer omdat de oogst aan eenduidige en bruikbare AT-interacties door velen als te mager werd beoordeeld en omdat 'procesgerichte en kwalitatieve analyses' als onderzoeksbenadering de overhand kregen, boven de puur statistische werkwijzen binnen het ATI-onderzoek. Ik meen echter intussen toch ook weer enkele voortekenen waar te nemen, die erop duiden dat dit type onderzoek aan een voorzichtige come-back bezig is. Het is inmiddels gelukkig wel ontdaan van een teveel aan pretenties en beter ingebed binnen moderne onderzoeksbenaderingen als 'proces-gerichte' instructie. Dat dit zo is moge blijken uit het feit, er op het ORD-programma van dit jaar op beperkte schaal weer een aantal 'interactie'-studies voorkomt, waarin relaties tussen persoonskenmerken en cognitieve vaardigheden bij het leren en studeren in verschillende leeromgevingen aan bod komen.

Een belangrijk deel van de bijdragen ging in 1974 over de ontwikkeling van *individuele studie- en instructiesystemen*, waarbij de computer in het onderwijsleerproces werd ingeschakeld. Dit onderwerp zal een paar jaar later geheel uit het vizier verdwijnen, waarschijnlijk als gevolg van gebrek aan succes, het ontbreken van technologische faciliteiten en van mogelijkheden om de computer meer te laten zijn dat een administratief hulpmiddel of als een 'drill-and-practice'-voorziening in het onderwijsleerproces, en tenslotte waarschijnlijk ook als gevolg van het verdwijnen van de RWO-bureaus aan de universiteiten, waar dit onderzoek veelal plaats vond.

Maar intussen is het onderzoek naar en met technologische voorzieningen – zeker mede als gevolg van de snelle technologische ontwikkelingen van de laatste jaren – weer helemaal terug en prijkt het thans weer opvallend op de onderzoeksagenda's van veel van onze collega's. Wat dit betreft, verwijs ik op deze plaats slechts naar het onderwerp 'krachtige leeromgevingen' op de ORD'96, waarin onder meer onderzoek naar de mogelijkheden om de computer te integreren in de onderwijsleersituatie veel aandacht krijgt.

Een ander onderdeel van het ORD'74-programma bestond uit presentaties over *survey-onderzoek naar leerlingkarakteristieken*, en naar kenmerken van leerkrachten (zo werden leraren aan de basisschool, toen nog *lagere school*, in die jaren in het algemeen aangeduid). In het eerste type onderzoek was de aandacht vooral gericht op schoolloopbaan-onderzoek en op kwesties van aansluiting tussen bijvoorbeeld de schoolloopbaan en het beroep of de relaties tussen dialect-gebruik en schoolsucces.

Deze onderzoekslijn heeft, als welhaast geen andere, alle stormen en ontwikkelingen binnen de onderwijsresearch-wereld van na 1974 glansrijk doorstaan. Nog immer is de aandacht voor schoolloopbanen groot binnen de onderwijsresearch, al moet gezegd, dat de onderwijssociologische representatie binnen het veld van onderzoekers er zeker niet groter op is geworden.

In de andere onderzoekslijn, *survey-onderzoek naar kenmerken van leerkrachten*, ging het in 1974 om een tamelijk gevarieerd aanbod. Er was aan de ene kant onderzoek naar loopbanen en naar functies en taken van leraren, dat naar opzet veel overeenkomst vertoonde met het schoolloopbaan-onderzoek bij leerlingen. Aan de andere kant werd er bijvoorbeeld ook onderzoek gepresenteerd over de relatie tussen 'denken en doen' van leraren. Je zou kunnen zeggen, dat het latere onderzoek naar 'teacher thinking' en 'proces-gerichte' instructie mede hierin zijn vertrekpunt heeft gevonden.

Het onderwerp '*diverse aspecten van het leren*' tenslotte, bestond in 1974 slechts uit een schameel aantal van vier presentaties: zij gingen over de relatie tussen creativiteit en intelligentie, de invloed van extra lessen gymnastiek, en over het lezen van blinde kinderen. Zo bezien moest het

onderzoek op het gebied van leren en instructie, wellicht met uitzondering van het onderzoek naar het lezen, en ook afgezien van het ATI-onderzoek dat, als vermeld, tijdens de ORD'74 afzonderlijk was vertegenwoordigd, in 1974 nog bijna geheel tot ontwikkeling komen. De bloei die dit onderzoek vandaag de dag doormaakt kon toen nog nauwelijks worden voorzien. Hoewel die bloeiperiode zich ook internationaal manifesteert, denk ik, dat de latere ontwikkeling van bijvoorbeeld een SVO-onderzoeksthmagroep op het gebied van onderwijsleerprocessen in ons land een stevige impuls heeft gegeven aan het leren- en instructie-onderzoek. In ieder geval staat het thema 'Leren & Instructie' al sinds jaar en dag prominent op de agenda van de ORD, en is het aantal bijdragen op dit terrein het veelvoudige van de vier die de eerste ORD te bieden had.

Wat ik me zelf herinner van die eerste ORD, is dat het programma bestond uit vele en altijd (te) korte zogenaamde paperlezingsen, met aan het slot een nog kortere periode van discussie met de zaal, als er althans tijd overbleef. In mijn beleving, maar ik was toen nog maar een net niet afgestudeerde student te Amsterdam, was dat het beeld van een echt wetenschappelijk congres. Een beeld, dat – eenmaal afgestudeerd – sindsdien vele malen werd bevestigd: een echt wetenschappelijk congres kenmerkt zich door (a) een teveel aan aanbod; (b) te vaak onduidelijke of onaffe presentaties; en vooral door (c) te weinig tijd voor interactie en discussie.

Hoewel opeenvolgende organisaties van ORD's er sinds 1974, bijvoorbeeld via alternatieve programmering, variaties in 'werkvormen' en sessies, er al van alles aan hebben gedaan om het karakter van de ORD's, de aard van de presentaties en de samenstelling van de groep deelnemers te beïnvloeden, is onze conferentie door de jaren heen toch in essentie dezelfde gebleven: nog steeds vallen de Onderwijs Research Dagen namelijk vooral op door de rijke en gevarieerde verzameling van presentaties en discussies over bevindingen die uit het onderwijs-onderzoek voortkomen, losjes georganiseerd onder brede thema's, die heel verschillende onderwijskundige en/of wetenschappelijke perspectieven toelaten. En uiteindelijk ben ik nu maar tot de slotsom gekomen, dat het zo wat mij betreft wel goed is, en dat het ook maar zo moet blijven.

Er kleven echter zeker ook aperte nadelen aan door jaren ingeslepen tradities zoals bij de ORD. Tradities zijn door gewoonte gevormde vaste gedragspatronen en verwachtingen, die – omdat zij in inmiddels onbewust geworden behoeften voorzien – welhaast resistent zijn tegen elke poging tot verandering. Hoewel het onze gewoonte hier bij Onderwijs- & Opleidingspsychologie in Tilburg is, om nu juist van dergelijke vaste gedragspatronen studie te maken en te onderzoeken, of er hierbij bijvoorbeeld sprake is van zoiets als een conferentie-stijl, en of er redenen zijn om te proberen iets aan die stijl te veranderen, hebben we dit streven, na overleg met de VOR, voor de ORD deze keer maar nagelaten. Vooral omdat de bezoeker van de ORD nu eenmaal vaste verwachtingen koestert, die moeten worden vervuld om het risico op 'destructieve fricties' bij de informatieverwerking – zoals wij dit in ons jargon uitdrukken – te vermijden.

Kort gezegd, wij hebben dus de traditionele paperlezing gehandhaafd in het programma, omdat het waarschijnlijk niet zozeer de vorm, maar de inhoud en de kwaliteit van het gebodene zijn, die bepalen of we leerzame ervaringen zullen opdoen en of er aanleiding tot discussie en interactie zal worden gezocht. Wij blijven het natuurlijk jammer vinden, dat – door tegemoet te komen aan deze eisen – de interactiviteit tijdens het congres door ons niet verder kon worden verhoogd. Dit zullen alle deelnemers er dus zelf moeten zien in te brengen. En zo hoort het eigenlijk ook bij een groep onderzoekers, van wie velen 'zelfverantwoordelijk leren' tot het hoogste goed rekenen.

Ook deze aflevering van de ORD zal gekenmerkt worden door een breed spectrum aan benaderingen en onderwerpen. Ook dit jaar zijn de bijdragen geordend onder de acht bekende thema's, die overeenkomen met de divisie-structuur van de Vereniging voor OnderwijsResearch (VOR). Naast deze acht, loopt dit jaar het door de KU Brabant geprogrammeerde thema 'krachtige leeromgevingen', als een rode draad door het programma. Voor dit centrale thema is gekozen, omdat in brede kring de overtuiging groeit, dat de wijze waarop we al jaren gewend zijn scholen en scholing in te richten aan een fundamentele heroverweging toe is.

De 'school zoals we die kennen' slaagt er – ondanks alle goede bedoelingen en inzet – steeds moeilijker in om leerlingen tot hoogwaardige vormen van leren te brengen, d.w.z. tot leren waarvan de resultaten duurzaam, flexibel, (multi)-functioneel en toepasbaar zijn ook buiten de school of in de vervolgopleiding. In dit thema komt recent onderzoek aan de orde, dat mogelijke aanknopingspunten biedt voor zo'n heroverweging.

De thematiek van de 'krachtige school' komt verder tot uitdrukking in de twee plenaire inleidingen van respectievelijk Prof. Dr. M. Boekaerts (RU Leiden) en Prof. Dr. P.R.J. Simons (KU Nijmegen) die aan het begin van de congresdagen zijn geprogrammeerd en die allebei gaan over het onderwerp: *De school van de toekomst, of de toekomst van de school*.

Moge het, in de geest van alle vorige afleveringen van de ORD, wederom een goede conferentie worden, waar u een heleboel interessante en leerzame ervaringen zult opdoen, die na de ORD hun waarde nog zullen bewijzen in het werk dat u in het kader van onderwijs en onderwijsresearch verricht!

Lezing ORD 96: De kwaliteit van onderwijsonderzoek

Nijs Lagerweij

Ik wil graag met het goede nieuws beginnen

Vanochtend stond het onderwijsonderzoek op de voorpagina. Paul Jungbluth liet in de Volkskrant (van 5 juni 1996) zijn licht schijnen over de problematiek van ongelijkheid in onderwijs en ongelijkheid in scholen. Hij is er wederom in geslaagd zijn favoriete thematiek onder de aandacht van de pers te brengen. Een goede start van de ORD 96. Wat Jungbluth doet, kunnen vele andere onderzoekers ook. Alleen lijken zij er minder attent op. Nog te weinig timmert de onderwijsresearch aan de weg. Dat vind ik niet verstandig. Een gevolg van dit bescheiden gedrag is dat de maatschappelijke waardering voor onderwijsonderzoek gering is. De politiek staat niet aan onze kant. Onze PR verdient aanzienlijk meer aandacht.

Vandaag ter inleiding van de Tilburgse ORD wil ik een paar woorden besteden aan deze ORD96 en aan een paar belangwekkende rapporten voor de onderwijskunde. Is dit wellicht het slechte nieuws?

1. ORD 96

Namens de VOR wil ik de organisatie van de ORD 96 graag complimenteren met de organisatie van het programma dat de komende twee dagen gepresenteerd wordt. Het congresboek toont zoals gebruikelijk een groot scala aan thema's. Lodewijks typeert dit in zijn inleiding met een treffend beeld: een marktplaats.

Dat heeft naast iets aantrekkelijks ook een schaduwzijde: het leidt ook gauw tot versnippering en gebrek aan coherentie. Zo lang er sprake is van onderwijsonderzoek is dat als een probleem naar voren gebracht. Tot op de dag van vandaag zijn we er echter niet goed in geslaagd tot samenhangende programmering te komen. Ik begin me af te vragen of we dit gewoon maar als een onmogelijke opgave moeten aanvaarden. Tenslotte is op de veel grotere Amerikaanse researchconferentie (de AERA) eveneens een enorme verscheidenheid aan te treffen. Daarover maakt men zich daar ook geen zorgen.

We ordenen de onderzoeksverslagen in divisies en jaarlijks komt er een zelfgekozen thematiek bij. Dit jaar koos de sectie Onderwijs & Opleidingspsychologie voor het eigen thema: "Krachtige leeromgevingen." Daarnaast zijn er afzonderlijke lezingen over de toekomst van het onderwijs. Een verstandige en noodzakelijke keuze. In het kader van ons huidig onderwijsbeleid staan basisvorming en studiehuis centraal. Daarin gaat het om nieuwe impulsen op het gebied van leren en onderwijzen. Activering van leerlingen staat in de belangstelling. Het verwerven van vaardigheden krijgt steeds meer aandacht en waardering. Ook in het Hoger Onderwijs beginnen dergelijke didactische opvattingen door te dringen en vorm te krijgen. In het congresboek zag ik hier en daar thema's die voor al deze actuele ontwikkelingen in het onderwijs zeker van betekenis kunnen zijn.

Onderwijskundige analyses en onderzoek zijn hard nodig om de impulsen tot onderwijsvernieuwing in de praktijk te ondersteunen. Ook onderwijskundige verbeeldingskracht is gewenst, zeker als het gaat om het scheppen van denkbeelden over de toekomst van het onderwijs. Onderwijskundigen kunnen hierbij waardevolle perspectieven bieden. Nog steeds huldig ik de opvatting dat onderwijskunde een wetenschap is die gericht is op het bijdragen aan de oplossing van problemen in het object van studie. Daarin ligt ook het maatschappelijk belang van onderwijskunde. Ik ben van mening dat wetenschapsbeoefening op het gebied van onderwijs in nauwe relatie staat en moet staan met maatschappelijke taken.

2. Hoe staat het met de onderwijsresearch in Nederland?

Naar mijn stellige indruk staat het er op korte termijn allerbelabberdst voor met de voortgang van het onderwijsonderzoek. De Tweede Kamer is akkoord gegaan met de voornemens van het kabinet ten aanzien van de toekomstige regeling van de onderwijsverzorging. Het voornemen is op het gebied van onderwijsonderzoek het fundamenteel alsmede het strategisch, evaluatief en toegepast onderzoek onder te brengen bij de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Een overweging hierbij is dat het onderbrengen van deze typen onderzoek bij één organisatie de onderlinge relatie zou moeten versterken. Voorwaarde is dat de NWO een apart programma voor onderwijsonderzoek opbouwt. Het opstellen van een dergelijk programma gebeurt met behulp van een programmaraad van deskundigen.

Deze ontwikkelingen zijn bedreigend voor de opgebouwde expertise in onderzoeksinstituten. Het VOR-bestuur vond het verstandig dit jaar nader overleg te voeren met de directie van de NWO en het ministerie. Tot nu toe hebben we twee keer van gedachten gewisseld met NWO en het Ministerie. Veel zorgen hebben we naar voren gebracht: verlies aan expertise, financiële problemen als gevolg van ander tariefstelsel en onduidelijkheid in de programmering. Veel zorgen zijn tot op heden niet weggenomen.

Verdient de onderwijsresearch deze onzekere tijden? Ik kan u voor beantwoording van die vraag verwijzen naar twee belangwekkende rapporten die dit afgelopen jaar verschenen zijn:

A. De onderzoekbeoordeling: *Quality Assessment of Research* door de VSNU en onder voorzitterschap van het VOR-erelid Fons Knoers.

B. *Nederlands Onderwijsonderzoek verkend 1996*. Rapport van de Overlegcommissie Verkenningen voor wetenschap en technologie. (Voorzitter De Moor).

Beide gaan dus over hetzelfde object: de Nederlandse onderwijsresearch. Beide rapporten hebben een ding gemeenschappelijk: er zit een kritische bijlage bij. In het eerste geval van de faculteit pedagogische en onderwijskunde van Amsterdam. In het tweede geval een minderheidsstandpunt van een lid van de commissie.

Eind 1995 verscheen het visitatierapport over het onderzoek op het gebied van de pedagogiek en de onderwijswetenschappen. Je moet wel even wennen aan de vermenging van pedagogiek met onderwijskunde. Het interdisciplinaire karakter van de onderwijskunde komt zo niet echt goed uit de verf.

De beoordeling van de commissie Knoers valt goed uit voor de onderwijsonderzoekers. Voor de pedagogiek is dat minder het geval. We werden zelfs in een volgorde van kwaliteit geplaatst. Dat is wel een nieuwe ervaring. Lang niet iedereen was gelukkig met deze gang van zaken. Ik hoop dat de volgende visitatie kan profiteren van deze eerste keer. We werden beoordeeld op criteria als: wetenschappelijke kwaliteit, productiviteit, en gelukkig ook maatschappelijk relevantie en tenslotte levensvatbaarheid en dat laatste op een vijfpuntsschaal. Het is maar te hopen dat we de volgende keer weten waarop beoordeeld wordt. Dat was nu pas achteraf duidelijk.

Het rapport van de Verkenningscommissie bevat een brief van de staatssecretaris over de organisatie en de programmering van het onderwijsonderzoek na het verdwijnen van de WOV. Zij formuleert daarin een aantal uitgangspunten die van belang zijn voor de toekomst van het onderwijsonderzoek.

1. De opgebouwde kwaliteit dient gehandhaafd te blijven.
2. Er dient synergie te ontstaan tussen geldstromen.
3. Het toegepast onderzoek moet in nauw overleg met de onderwijspraktijk geprogrammeerd worden.
4. De programmering dient plaats te vinden binnen de kaders van de NWO-organisatie en met zo min mogelijk bureaucratie omgeven.

Met grote interesse heb ik uitgekeken naar de bevindingen van de verkenningscommissie onderwijsonderzoek. Het rapport verscheen in mei 1996. Men heeft interessante vragen proberen te

beantwoorden, zoals bijvoorbeeld: Wat is de wetenschappelijke kwaliteit van het Nederlandse onderzoek van onderwijs? Hoe is de verhouding tussen fundamenteel strategisch en praktijkgericht onderzoek?

Hoe wordt de kennisbehoefte van de vraagzijde gearticuleerd? Wordt het hele veld bestreken? Hoe staat het met de kennisoverdracht? Hoe kan onderzoek helpen nieuwe informatietechnologie te integreren? Op welke maatschappelijke vraagstukken is onderzoek gericht?

Er verschenen tussentijds een paar interessante inventariserende deelstudies van het ITS en andere instituten. Volgens de verkenningscommissie verdienen twee thema's in de toekomst aandacht:

- leren en instructie;
- management en organisatie van de school en het onderwijs.

Naar mijn mening zijn we met zo'n conclusie niet veel opgeschoten. Hier kan heel onderwijsonderzoekland zich immers wel in vinden. Voor iedereen zit er wel wat van belang bij. Ik zou nog willen toevoegen ook aandacht te besteden aan de relatie tussen de twee thema's. Juist het verband tussen micro-niveau en meso-niveau in scholen is in het kader van toenemende autonomie voor scholen een belangwekkende thematiek.

De commissie heeft ook nog een derde thema aangemerkt, dat zeker in de nabije toekomst nog aan belang zal winnen:

- communicatie- en informatietechnologie in het onderwijs.

Op deze plaats is het nauwelijks mogelijk een gedegen analyse van het rapport te maken. Het ligt meer op de weg van de VOR en haar leden om bijvoorbeeld in het vakblad TOR grondig aandacht te besteden aan de voorstellen. Mijn voorlopige indruk is dat het hier om allesbehalve indrukwekkende vondsten of perspectieven gaat die de commissie ons biedt. Gedurende de werkzaamheden heeft de commissie diverse mensen verzocht te reageren op toekomstscenario's. De uiteindelijke relatie met deze scenario's is mij in het eindrapport niet duidelijk geworden.

Over de programmering, een van ouds netelig probleem, komt de opmerking dat er afstemming nodig is op de programmering van NWO. Lang leve de programmaraad, die er binnenkort wel zal komen. We moeten daarvan kennelijk veel heil verwachten.

Ik ben benieuwd wat dit rapport nog kan betekenen voor de verdere voortgang van het onderwijsonderzoek onder de vleugels van de NWO. Ik maak me er grote zorgen over. In mijn column in *Het Schoolblad* van 20 april 1996 schreef ik er de volgende woorden over:

“Met ingang van 1997 komt er een eind aan het bestaan van een eigen stelsel van regulering van het onderwijsonderzoek. Het geld hiervoor gaat in het vervolg voor het grootste deel naar NWO. Slechts iets meer dan twee miljoen gulden is gereserveerd voor onderzoek vanuit het veld. Dat moet dan het praktijkrelevante onderzoek worden. Dit geld wordt in de toekomst verdeeld door de drie Landelijke Pedagogische Centra. Daar moeten natuurlijk ook de kosten voor de organisatie nog van af. Tel uit je winst.

NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) is een instituut voor zuiver wetenschappelijk onderzoek. Hier maken wetenschappers de dienst uit. Theoriegericht onderzoek luidt er de boodschap. Het geld voor onderwijsonderzoek wordt ondergebracht in een stichting Maatschappij- en Gedragswetenschappen. Een bestuur is verantwoordelijk voor de middelenverdeling. Het einde van de praktijkrelevantie van onderwijsonderzoek is daarmee aangebroken. Wat heeft Didaktief dan nog te melden?

Op het oog lijkt de opheffing van SVO een te verdedigen besluit. Niet alles was rozegeur en maneschijn rondom deze organisatie. Je kunt je zeker afvragen of er voor onderwijsonderzoek wel een afzonderlijk instituut nodig is. In het buitenland is men er niettemin jaloers op. Ik heb daar in diverse bewoordingen door wetenschappers van naam de loftrompet horen blazen, als zij zicht kregen op onze situatie.

Die eer is ons nu te veel geworden. Genadeloos schaft men alle procedures en inspraak af die het veld had bij de SVO. Zonder enig mededogen verdwijnen nu noodgedwongen veel experts

die op een stuk of vijf gespecialiseerde onderzoeksinstituten werkzaam zijn. De NWO hanteert namelijk volstrekt andere procedures en betaalt veel lagere tarieven dan bij SVO het geval was. Weg investeringen. Einde van instituten van naam en faam.

We moeten weer met de kop onder het maaiveld. Dat geheel in lijn met de in ons land zo populaire calvinistische traditie: doe maar gewoon dan doe je al gek genoeg. Welke gek zit achter dit masterplan?"

3. Slot

Over de VOR kan ik kort zijn: het gaat uitstekend met onze vereniging.

Het is denkbaar dat de VOR in de toekomst meer op de maatschappij gerichte activiteiten gaat ontwikkelen. In de jaren vijftig en zestig was de WVO (Werkgemeenschap voor Onderwijsvernieuwing) erg actief in het organiseren van nationale conferenties over vraagstukken inzake onderwijsvernieuwing. Inmiddels zorgen overheid, verzorgingsinstellingen en nascholingsinstituten voor een levendig conferentie-aanbod. Zou hier niet een taak voor de VOR liggen? Er is immers nogal wat te verbeteren aan de communicatie tussen onderwijskunde en de samenleving. Door het naar buiten brengen van belangwekkende informatie over onderwijspolitieke thema's kan aan verbetering van de interactie worden gewerkt.

Inmiddels zijn de voorbereidingen voor een Europese vereniging voor onderwijsonderzoek (de EERA) voorspoedig afgesloten. Onze vereniging neemt daaraan op actieve wijze deel. In oktober 1993 is in Aarau in Zwitserland door zo'n 14 landen het principebesluit genomen tot stichting van een Europese vereniging. Het doel van deze vereniging is het bevorderen van onderzoek en ontwikkeling (research and development) van onderwijs in Europa en zodoende het bevorderen van onderwijskunde (educational science) en het verbeteren van onderwijs.

In september is de eerste zelfstandige conferentie een feit. In het Spaanse Sevilla komen vele onderzoekers uit Europa bijeen. Er hebben zich ongeveer 500 deelnemers aangemeld, die allen iets willen presenteren. Europa trekt ook personen uit de VS, Israël, Brazilië, Rusland en Hongkong.

Volgend jaar vindt de ORD plaats in Leuven. In Vlaanderen is inmiddels ook een vereniging voor onderwijsonderzoekers opgericht. Zij organiseren de Vlaams-Nederlandse ORD97. Ik spreek graag de wens uit dat deze researchdagen veel mogen bijdragen aan het onderhouden en verbeteren van de kwaliteit van onderwijsonderzoek. Het inmiddels bereikte niveau laat zien dat we best iets boven het maaiveld mogen uitkomen.

Artikelen

Gedrag van aanstaande docenten in aandachtseisende situaties in de klas <i>Wilfried Admiraal, Theo Wubbels en Fred Korthagen</i>	162
Schoolloopbaansucces en de breedte van het opleidingsaanbod in het MBO <i>Th.A. van Batenburg en P.R. den Boer</i>	205
Effecten van de aanwezigheid en inrichting van de praktijkcomponent bij scholing van laag opgeleide volwassenen <i>P.R. den Boer</i>	222
Gedecentraliseerde besluitvorming binnen mbo-instellingen: een meetmodel voor domeinen van besluitvorming <i>Pieter L.J. Boerman en Timo M. Bechger</i>	348
Schooladviezen van allochtone leerlingen <i>Hans Bosma en Peter Cremers</i>	262
Het intern rendement van het leerlingwezen <i>G.J. Harms en P.R. den Boer</i>	212
Inleiding: Onderwijs aan jonge kinderen nog nauwelijks onderzocht <i>Egbert Harskamp en Gerry Reezigt</i>	1
Het duale karakter van beroepsopleidingen en hun waardering door werkgevers <i>W.Th.J.G. Hoebe</i>	233
Rendement van het beroepsonderwijs: Beleid en onderzoek <i>W.Th.G.J. Hoebe</i>	197
De factorstructuur van schrijfvaardigheid: een toepassing van multilevel analyse <i>Hans Kuhlemeier en Huub van den Bergh</i>	133
Stimulering van geletterdheid bij kleuters <i>Jef van Kuyk en Ludo Verhoeven</i>	33
Onderwijs aan jonge kinderen: historische ontwikkelingen <i>G.W. Meijnen</i>	95
Onderzoek naar vormgeving en effecten van het onderwijs aan jonge kinderen in Nederland <i>Gerry Reezigt</i>	7
Leerstijlen en zelfstandig leren in het voorportaal van het studiehuis <i>Alex Roosendaal en Jan Vermunt</i>	336
Een onderzoek naar leerstijlen <i>H.C. Schouwenburg</i>	151
Verschillen in leerstijlen: Sekse en sekse-identiteit <i>Sabine Severiens en Geert ten Dam</i>	294

De constructie en evaluatie van een videotoets voor de beoordeling van bekwaamheid in professionele gespreksvoering <i>G.N. Smit en H.T. van der Molen</i>	182
Actuele opvattingen over onderwijs aan jonge kinderen. Een onderzoek onder leer- krachten van groep 2 in het basisonderwijs <i>L.M. Stevens en E.G. Harskamp</i>	81
Leerstijlen en het opzoeken van informatie in tekst <i>Evelien T. Stoutjesdijk en Jos J. Beishuizen</i>	323
Ontwikkeling van tweetaligheid bij allochtone kleuters <i>Ludo Verhoeven en Goretti Narain</i>	54
Inleiding op het thema leerstijlen <i>Jan Vermunt en Johan van der Sanden</i>	291
Impliciet versus expliciet onderwijs in bedrijfseconomische modellen <i>A.T.J. Vernooij</i>	246
Leeromgevingspercepties en leerstrategieën van eerstejaars studenten sociale wetenschappen <i>Ronny F.A. Wierstra en Eva P.M. Beerends</i>	306
Onderwijsexpansie of meritocratisering? Een verklaring voor de afgenomen samenhang tussen het ouderlijk sociaal milieu en het hoogst bereikte opleidings- niveau van kinderen <i>M.H.J. Wolbers en P.M. de Graaf</i>	117
Notities en Commentaren	
Constructieve en destructieve discussies over de interdisciplinaire onderwijskunde. Een reactie op De Klerk <i>B. Creemers, N. Lagerweij, en N. Verloop</i>	284
Scoringsregels en toetsinstructie <i>A. Dirkzwager</i>	272
Het toekomstperspectief van de onderwijskunde <i>Len F.W. de Klerk</i>	278
Treffers schiet mis <i>W.J. van der Linden en M.A. Zwarts</i>	106
Maar raakt wel zijn eigen voet <i>W.J. van der Linden en M.A. Zwarts</i>	111
Rekeninhoudelijke kritiek op ceb-repliek <i>A. Treffers</i>	103