



Tijdschrift voor onderwijsresearch 1999-2000 : een tweemaandelijks tijdschrift voor onderzoekers van onderwijs, gewijd aan theoretische, methodologische en professionele problemen in de onderwijsresearch

<https://hdl.handle.net/1874/214866>

Inhoud

S W

Inleiding op het thema leerconcepties / *pag. 3**G. ten Dam, S. Severiens en J. Vermunt*Een structureel model over de relaties tussen leeropvattingen, regulatie en cognitieve verwerking / *pag. 8**Y.J. Vermetten, H.G. Lodewijks en J.D. Vermunt*

UNIVERSITEITSBIBLIOTHEEK UTRECHT



3174 214 5

Leerconcepties en leerstrategieën cross-sectioneel vergeleken / *pag. 21**S. Severiens en G. ten Dam*De samenhang tussen leerconcepties en verwerkingsstrategieën bij leerlingen in het middelbaar beroepsonderwijs / *pag. 35**A. Slaats, J. v.d. Sanden en H. Lodewijks*Epistemologische opvattingen of mentale leermodellen in het Hoger Onderwijs: een verkenning van het terrein / *pag. 50**G. Schellings, H. Lodewijks en J. v.d. Sanden***Notities en Commentaren***J. Dronkers* - Veranderden leerlingenaantallen in het voortgezet onderwijs in het schooljaar 1998-1999? / *pag. 63**F. Goffree* - Dimensies, dilemma's en didactiek / *pag. 66***Boekbesprekingen***J.J.G. van Merriënboer* - Training Complex Cognitive Skills.A four-component designmodel for technical training / *pag. 73**J. te Nijenhuis* - Comparability of test scores for immigrants and majority group members in the Netherlands / *pag. 75**T.M. Janssen* - Literatuuronderwijs bij benadering / *pag. 77*Jaarverslag Vereniging voor Onderwijs Research 1998-1999 / *pag. 85*

Inhoud

Predictoren van de keuze van natuurkunde in de tweede fase VO / **pag. 91**

K.M. Stokking

De score op een ja/nee-woordenschattoets:

correctie voor raden en persoonlijke antwoordstijl / **pag. 110**

I. Huibregtse en W. Admiraal

Is het voor studiesucces belangrijk om uit te kijken
naar een goede opleidingsplaats? / **pag. 125**

A. Need en U. de Jong

Waargenomen Eigen Competentie, Doeloriëntaties en Studieresultaten
bij Beginners en Gevorderden / **pag. 144**

A. Vrugt, F.J. Oort en C. Zeeberg

Boekbesprekingen

J. Buitink - In-functie opleiden en in-functie leren van aanstaande
docenten / **pag. 156**

M. Gemmeke - Politieke betrokkenheid van kinderen op
de basisschool / **pag. 158**

M.H.J. Wolbers - Diploma-inflatie en verdringing op de arbeidsmarkt.
Een studie naar ontwikkelingen in de opbrengsten van diploma's
in Nederland / **pag. 161**

P. Meijer - Teachers' practical knowledge. Teaching reading comprehension
in secondary education / **pag. 164**

P. Bimmel - Training en transfer van leesstrategieën: training in de moedertaal
en transfer naar een vreemde taal, een effectstudie bij leerlingen uit het
voortgezet onderwijs / **pag. 166**

Verslag Algemene Ledenvergadering VOR d.d. 21 mei 1999 / **pag. 170**

Aankondiging ORD 2000 / **pag. 173**

Inhoud

X Ouderlijk gezin en schoolsucces / **pag. 179**

G. Kraaykamp

Samenwerkend leren in computerondersteund rollenspel / **pag. 195**

G.A. Holsbrink-Engels

X Cultureel kapitaal en onderwijsprestaties: differentiëren naar
etnische groep? / **pag. 215**

G.W.J.M. Driessen

Effecten van curriculumkenmerken op studievoortgang en uitval
van studenten techniek / **pag. 237**

M. van der Hulst en E. Jansen

Het ILS-instrument voor leerstijlen:
een methodologische kanttekening / **pag. 248**

P. van Petegem en S. De Maeyer

De studentenrechtbank: een evaluatieonderzoek naar beoordelingslijsten
voor prestatievaardigheden / **pag. 265**

F.M. Edens, F. Rink en M.J. Smilde

De relatie tussen leerstijl en persoonlijkheid / **pag. 275**

J. Adema, K.I. van der Zee, H.T. van der Molen

Studievoortgang in het wetenschappelijk onderwijs / **pag. 288**

M. van den Berg en A. Hofman

De stabiliteit en consistentie van differentiële schooleffecten / **pag. 308**

R.J. Bosker en H. Luyten

Notities en Commentaren / **pag. 322**

Boekbesprekingen / **pag. 328**

Redactionele medewerkers jaargang 23 en 24 / **pag. 339**

Inleiding op het thema leerconcepties

Geert ten Dam, Sabine Severiens (*Instituut voor de Lerarenopleiding, Universiteit van Amsterdam*), en Jan Vermunt (*Capaciteitsgroep Onderwijsontwikkeling en Onderwijsresearch, Universiteit Maastricht*)

De opvattingen over leren en kennis van leerlingen en studenten zijn de laatste jaren steeds meer in de belangstelling komen te staan (zie bijv. Hofer & Pintrich, 1997). Uit verschillende onderzoeken blijkt dat deze leer- en kennisconcepties voor een belangrijk deel variaties in de leeractiviteiten van leerlingen en studenten kunnen verklaren. Deze samenhangen tussen leeractiviteiten en leerconcepties, vaak aangevuld met motivationele componenten, vormen de basis voor overkoepelende begrippen als 'learning orientations' (Entwistle, 1997), 'study orchestrations' (Meyer, 1991) en 'leerstijlen' (Vermunt, 1992; 1996; Busato, 1998).

Een van de meest bekende onderzoeken naar leerconcepties is verricht door Säljö (1979). Hij vroeg volwassen studenten wat ze nu eigenlijk onder leren verstonden. Op basis van fenomenografische analyse van de antwoorden identificeerde hij vijf kwalitatief verschillende leerconcepties. Leren wordt daarbij opgevat als: (1) een kwantitatieve toename in kennis; (2) memoriseren; (3) het verwerven van feiten, procedures, enz., die onthouden kunnen worden en toegepast in de praktijk; (4) het verlenen van betekenis; en (5) een interpretatief proces gericht op het begrijpen van de werkelijkheid. De belangrijkste scheidingslijn tussen deze vijf leerconcepties lijkt te lopen tussen leerconceptie 1, 2 en 3 aan de ene kant en 4 en 5 aan de andere kant. De eerste drie worden gekenmerkt door een reproductieve visie op leren. Ze zijn gebaseerd op de gedachte dat kennis al bestaat, in boeken, docenten, en dergelijke en dat leren in essentie neerkomt op het overbrengen van deze kennis naar het geheugen van de lerende, alwaar de kennis wordt opgeslagen. Leerconceptie 4 en 5 daarentegen worden gekenmerkt door een meer constructieve visie op leren. Leren houdt dan in proberen kennis op te bouwen door middel van een actief (re)constructieproces door de lerende. Marton, Dall'Alba en Beaty (1993) voegden daar nog een zesde categorie aan toe: leren als een verandering in de persoon zelf.

De visie op kennis die deze leerconcepties mede van elkaar onderscheidt, doet denken aan de stadia die Perry (1970) in een longitudinaal onderzoek waarnam in de ontwikkeling van studenten in hun epistemologische opvattingen gedurende de universitaire studie. Beginnende studenten hadden over het algemeen een absolutistische visie op kennis, waarbij informatie wordt beoordeeld als 'waar' of 'onwaar' (dualisme). Het volgende stadium werd gekenmerkt door een meer relativistische visie: er bestaan meerdere meningen en ideeën en deze zijn alle even legitiem aangezien er niet zoiets als de enige waarheid bestaat. Naarmate de ontwikkeling vorderde traden studenten een fase binnen van betrokkenheid. Men erkent dat er verschillende legitieme opvattingen over en visies op een bepaald verschijnsel mogelijk zijn, maar men kiest voor een bepaalde eigen visie of opvatting die men het meest nuttig of waardevol vindt.

Leerconcepties vormen één van de componenten van leerstijlen zoals Vermunt (1992) die heeft geoperationaliseerd. Hij hanteert een brede invulling van het begrip leerstijl en onderscheidt aan een leerstijl vier componenten: cognitieve verwerkingsstrategieën, metacognitieve regulatiestrategieën, mentale leermodellen (ofwel leerconcepties) en leeroriëntaties (de motivationele component). Leerconcepties vormen dus één van de componenten van leerstijlen. Omdat deze component centraal staat in het huidige themanummer, beschrijven we deze in enig detail. Leerconcepties kunnen worden opgevat als een aspect van metacognitie. Het gaat om de kennis, opvattingen, concepties en denkbeelden die mensen hebben over leren en onderwijs. Op basis van empirisch onderzoek worden vijf mentale leermodellen onder-

scheiden die van elkaar verschillen 'in wat onder leren en studeren wordt verstaan en in de visie op de taakverdeling tussen de student, het onderwijs en medestudenten': (1) leren gezien als een reproductief proces waarbij extern aanwezige kennis moet worden opgenomen in het geheugen van de student ('opname van kennis'); (2) een visie op leren als een constructief proces waarbij door een actieve inzet van leeractiviteiten eigen kennis en inzichten worden opgebouwd ('opbouw van kennis'); (3) een visie op leren waarin veel waarde wordt gehecht aan het gebruiken van de kennis die men verwerft ('gebruik van kennis'); (4) een opvatting waarin het als belangrijke taak van het onderwijs wordt gezien studenten voortdurend te stimuleren tot het uitvoeren van leeractiviteiten ('stimulerend onderwijs'); en (5) een visie waarin veel waarde wordt gehecht aan het verdelen van de taken bij het studeren met medestudenten ('samen leren') (Vermunt, 1992). De vijf modellen zijn min of meer onafhankelijk van elkaar, wat wil zeggen dat ze niet worden beschouwd als fasen in een ontwikkelingsproces.

Kenmerkend voor een leerstijl zijn de karakteristieke combinaties van de genoemde componenten (verwerking, regulatie, conceptie en oriëntatie). Onderzoek onder universitaire studenten leidde tot het identificeren van vier leerstijlen: betekenisgericht, reproductiegericht, toepassingsgericht en ongericht leren. In de betekenisgerichte leerstijl wordt een leerconceptie waarin 'opbouw van kennis' centraal staat, verbonden met de diepe verwerkingsstrategieën (relateren en kritisch verwerken). Deze leerstijl kenmerkt zich tevens door zelfregulatie (controleren van het eigen leerproces etc.) en een leeroriëntatie waarin persoonlijke interesse centraal staat. Lerenden met een reproductiegerichte leerstijl daarentegen zien leren als 'opname van kennis' en gebruiken de stapsgewijze verwerkingsstrategieën (memoriseren en analyseren) om te leren. De regulatie behorende bij deze leerstijl is externe regulatie (gebruiken van externe bronnen voor de regulatie van het eigen leerproces) en de leeroriëntatie bestaat uit een gerichtheid op het halen van tentamens of proefwerken. De derde leerstijl is de toepassingsgerichte leerstijl. Lerenden met deze leerstijl zien leren als het 'gebruiken van kennis', maken veel gebruik van concrete verwerkingsstrategieën tijdens het leren en zijn beroepsgericht in hun leeroriëntatie. Wat betreft regulatie is deze leerstijl tamelijk neutraal: zowel zelfgestuurde als extern gestuurde varianten komen voor. De vierde leerstijl is de zogenaamde ongerichte leerstijl. Deze leerstijl kenmerkt zich doordat lerenden een gebrek aan regulatie ervaren ('stuurloos leergedrag'), een ambivalente leeroriëntatie vertonen, in hun leerconceptie vooral waarde hechten aan de sturing die anderen, zoals medestudenten en docenten, aan hun leren kunnen geven ('samen leren' en 'stimulerend onderwijs') en nauwelijks toekomen aan het gebruik van strategieën om de leerstof te verwerken.

De invulling van het begrip leerstijl als een samenhangend geheel van zowel leerstrategieën (verwerking en regulatie), als leerconcepties en leeroriëntaties, doet recht aan de complexiteit van leerprocessen (Boekaerts, Otten & Simons, 1997). Maar niet alleen vanuit het perspectief van een leerstijl als geheel heeft een brede invulling van leerstijlen meerwaarde, ook vanuit het perspectief van de onderscheiden componenten is er winst. Zo kan onderzoek naar de cognitieve verwerkingsstrategieën van studenten gebruik maken van het inzicht dat de manier waarop lerenden informatie verwerken niet op zichzelf staat maar beïnvloed wordt door de regulatie-activiteiten van de betrokkene en vice versa. Hetzelfde geldt voor de meermaalen geconstateerde verbanden tussen verwerkingsstrategieën en opvattingen over leren en onderwijs. Op basis van interviewonderzoek ontwikkelden Vermunt en zijn collega's de *Inventaris Leerstijlen (ILS)*. Deze ILS is in verschillende onderwijssettings op zijn merites beoordeeld. En alhoewel het instrument met name is toegesneden op de situatie in het hoger onderwijs, laten de betrouwbaarheid en validiteit ervan ook gebruik in de bovenbouw van het voortgezet onderwijs ruimschoots toe. In twee van de vier bijdragen in het huidige themanummer wordt gebruik gemaakt van deze ILS. Het doel van het huidige themanummer is echter niet gelegen in een kritische reflectie op de psychometrische kwaliteiten van de ILS, daarvoor verwijzen we naar onder meer Busato (1998) en Vermunt (1999).

Dit themanummer 'leerconcepties' bestaat uit vier bijdragen die gebruik maken van het theoretisch kader zoals dat ontwikkeld is door Vermunt (1992), al dan niet gezet naast ande-

re uitwerkingen (bijv. Schommer). Het themanummer is de weerslag van een symposium dat plaatsvond op de Onderwijs Research Dagen in 1998. Dit gegeven betekent dat het huidige themanummer niet een uitputtende beschrijving is van het 'leerconcepties' onderzoek zoals dat in Nederland wordt verricht. We verwijzen naar onderzoek van bijvoorbeeld Busato (1998) en Oosterheert (1998) om het beeld verder te complementeren. De vier bijdragen beschrijven onderzoek naar leerconcepties al dan niet in verhouding tot leerstrategieën in verschillende onderwijscontexten. Dat de bijdragen vertrekken vanuit eenzelfde theoretisch kader biedt als voordeel de mogelijkheid tot het vergelijken van de resultaten. In deze inleiding wordt enkele lijnen uitgezet om deze vergelijking te kunnen maken.

In een uitgebreide review-studie pleiten Hofer en Pintrich (1997) voor een helder onderscheid tussen opvattingen over de aard van *leren* aan de ene kant, en opvattingen over de aard van *kennis* aan de andere kant. Terwijl het onderzoek van Vermunt en dat van zijn Zweedse leermeesters Säljö en Marton betrekking hebben op *leer*opvattingen, is het werk van Perry illustratief voor onderzoek naar opvattingen over *kennis*. Alhoewel op conceptueel niveau te onderscheiden, gaan Hofer en Pintrich er vanuit dat leerconcepties en kennisconcepties waarschijnlijk wel met elkaar samenhangen. Deze stelling zal echter op basis van nader onderzoek verder handen en voeten moeten worden gegeven. Verschillende bijdragen aan dit themanummer gaan in op dit onderscheid tussen leer- en kennisconcepties.

Zoals gezegd worden leerconcepties in dit themanummer niet als een op zichzelf staand fenomeen beschouwd. In de verschillende bijdragen worden leerconcepties gerelateerd aan leerstrategieën. De vraag naar het verband tussen beide leerstijl-componenten is zowel vanuit een theoretische als praktische invalshoek van belang. Allereerst levert het onderzoeken van het verband tussen leerconcepties en leerstrategieën een waardevolle bijdrage aan de verheldering van het concept leerstijl: in hoeverre vormen de verschillende onderdelen van het begrip leerstijlen een samenhangend geheel? Indien er sprake is van een onderling verband, wat is de precieze aard (richting) van die samenhang en onder welke condities treedt deze op? Ook vanuit een praktische, verbeteringsgerichte optiek is onderzoek naar het verband tussen leerconcepties en leerstrategieën van belang. We weten uit reeds verricht onderzoek dat leerlingen en studenten met een 'diepe verwerkingsstrategie' (relateren en structureren van de leerstof, kritisch benaderen van de leerstof) in het algemeen tot betere leerprestaties komen dan lerenden met een 'oppervlakkige verwerkingsstrategie'. Anders geformuleerd: de kwaliteit van kennis verworven door actieve kennisconstructie is beter dan die verworven voor passieve kennisopname. Om leerstrategieën te sturen in de gewenste richting kan uiteraard getracht worden het feitelijke leergedrag van leerlingen en studenten te beïnvloeden. Training in geïsoleerde strategieën leverde echter nogal eens teleurstellende resultaten op (Brown, Armbruster & Baker, 1986). Bij een onderlinge relatie tussen leerconcepties en leerstrategieën opent zich echter een tweede weg: ook via het 'sturen' van de opvattingen die lerenden hebben over leren en onderwijs kan worden getracht de manier waarop wordt geleerd te veranderen.

Het themanummer start met een artikel van Vermetten, Lodewijks en Vermunt waarin de onderlinge verbanden tussen cognitieve verwerking, regulatie en leerconcepties zijn vertaald in een structureel model. In het model wordt verondersteld dat leerconcepties onderliggend zijn aan regulatie en cognitieve verwerking. In hun grootschalige onderzoek onder propedeutestudenten van verschillende studierichtingen aan de KUB, wordt deze veronderstelling bevestigd. Bovendien komen er ook andere verbanden in het model aan het licht waardoor meer duidelijkheid ontstaat over de verschillende componenten van een leerstijl. Vermetten *et al.* concluderen dat interventieprogramma's waarin het gaat om het leergedrag van studenten waarschijnlijk succesvoller zijn als overtuigingen van studenten omtrent onderwijs en leren, hun leerconcepties, een belangrijke rol spelen.

Severiens en Ten Dam vervolgen het themanummer met een cross-sectioneel onderzoek naar stabiliteit van leerstrategieën, leerconcepties en de verbanden daartussen. Gedurende de bovenbouw van het voortgezet onderwijs en het hoger onderwijs lijkt het repertoire aan strategieën toe te nemen en tevens veranderen de opvattingen over leren en onderwijs. De onder-

linge verbanden tussen beide componenten maken ook veranderingen door. Voorts blijkt uit de resultaten dat zowel leeftijd als onderwijsniveau een deel van de ontwikkelingen in leerstrategieën, leerconcepties en hun onderlinge verbanden kunnen verklaren. De meeste veranderingen hebben overigens betrekking op die leerconcepties en strategieën die met elkaar verbonden zijn in de betekenisgerichte leerstijl. Severiens en Ten Dam concluderen hieruit dat de componenten in deze stijl zich later ontwikkelen dan de componenten in de reproductiegerichte leerstijl.

In de derde bijdrage van Slaats, Van der Sanden en Lodewijks wordt een onderzoek in het middelbaar beroepsonderwijs beschreven. Met name vanwege de praktijkcomponent in dit type onderwijs, waardoor deze onderwijssetting nogal afwijkt van de oorspronkelijke context en doelgroep waarvoor de ILS werd ontwikkeld, besloten Slaats *et al.* tot een kwalitatief onderzoek. Zij onderzochten de samenhang tussen cognitieve verwerking en leerconcepties met behulp van interviews, die geanalyseerd zijn middels de fenomenografische methode. De gevonden leerstrategieën en leerconcepties lijken sterk op de strategieën en concepties zoals beschreven door Vermunt. De verbanden tussen beide concepten zijn echter alleen in de zogenaamde theorievakken terug te vinden, in de praktijkvakken lijken concepties van leren en strategieën niet met elkaar verbonden. Slaats *et al.* stellen voor in vervolgonderzoek meer aandacht te besteden aan specifieke praktijkleerconcepties: het woord 'leren' roept in een praktijk situatie wellicht andere dan de bedoelde associaties op.

De laatste bijdrage is geleverd door Schellings, Lodewijks en Van der Sanden. Zij leggen een verband tussen leerconcepties en de theorie van Schommer omtrent epistemologische opvattingen. Ze richten zich in eerste instantie op verbanden op het conceptuele vlak. Daarnaast worden twee verkennende studies beschreven waarin epistemologische opvattingen en leerconcepties empirisch worden onderzocht. De epistemologische opvattingen zijn onderzocht met het instrument van Schommer. Dit instrument is echter niet ideaal gebleken. In de discussie van Schellings *et al.* wordt ingegaan op mogelijke problemen, en wordt voorgesteld meer instrumentatie-onderzoek te verrichten.

De bijdragen in dit themanummer laten ten eerste zien dat leerconcepties op uiteenlopende manieren samenhangen met de cognitieve verwerking van studenten. In het hoger onderwijs lijkt het erop, volgens Vermetten *et al.*, dat leerconcepties sturend zijn voor de manier waarop studenten leren. In andere typen onderwijs kan deze conclusie niet zonder meer getrokken worden. Slaats *et al.* laten zien dat in MBO-praktijkleersituaties, leerconcepties niet op dezelfde manier tot uiting komen, en zich ook niet op dezelfde manier laten vertalen in concreet leergedrag. In de bovenbouw van het voortgezet onderwijs lijkt te gelden dat de verbanden tussen concepties en strategieën in de reproductiegerichte leerstijl al wel zijn ontwikkeld, maar de componenten in de betekenisgerichte leerstijl nog volop in ontwikkeling zijn. Het lijkt er echter wel op dat de bevindingen consequenties moeten hebben voor de plaats van leerconcepties in het ontwerpen van strategie-onderwijs. In de conclusies van Vermetten *et al.* wordt hier uitgebreid op ingegaan.

Daarnaast worden in dit themanummer op verschillende plaatsen conceptuele verbanden gesignaleerd tussen leerconcepties en de zogenaamde epistemologische opvattingen. Alleen in de bijdrage van Schellings *et al.* wordt hier, in empirische zin, verder op ingegaan. Hier ligt een duidelijk hiaat aan onderzoek. Het verder afbakenen van concepties van kennis en concepties van leren, het onderzoeken van de verbanden ervan, maar vooral het ontwikkelen van een goed 'epistemologische opvattingen' instrument, is van groot belang voor dit onderzoeksterrein.

Tot slot dringt zich nog een vraag op die de reikwijdte van dit themanummer overstijgt, maar die we desondanks niet onvermeld willen laten. Onder invloed waarvan ontwikkelen de opvattingen die lerenden hebben van leren en onderwijs zich? Leerpsychologisch onderzoek beperkt zich doorgaans tot het onderzoeken en op termijn bevorderen van de gewenste leerprocessen bij individuele leerlingen en studenten. In toenemende mate benaderen onderzoekers (bijv. Cobb & Yackel, 1996; Brown & Campione, 1990) leren echter als sociaal proces.

Hiermee wordt bedoeld dat individuele activiteit en 'classroom discourse' alleen in wisselwerking met elkaar kunnen worden begrepen. Lerenden ontwikkelen hun opvattingen over leren en onderwijs doordat ze participeren in de klas, in de interactie met de docent en met medeleerlingen. Door te participeren leveren ze op hun beurt ook een bijdrage aan het leer-klimaat. Alhoewel het in deze onderzoekslijn uiteindelijk ook gaat om de kwaliteit van het leerproces van individuen, is niet zozeer de individuele leerling het belangrijkste aandachtspunt als wel de kwaliteit van de leergemeenschap ('community of learners'). Wie wil dat lerenden zich ontwikkelen tot actieve lerenden met een opvatting over leren als een constructief proces, zal ruim baan moeten maken voor het gezamenlijke proces van betekenisgeving waarin leerlingen en studenten na kunnen gaan wat leren voor hen inhoudt, wat ze er mee willen, wat hen belemmert, stimuleert, waarin ze van elkaar verschillen, enzovoorts.

REFERENTIES

- Boekaerts, M., Otten, R., & Simons, R.J. (1997). Een onderzoek naar de bruikbaarheid van de ILS. *Leerlijnen in de onderbouw van het voortgezet onderwijs. Tijdschrift voor Onderwijsresearch*, 22, 15-36.
- Brown, A.L., Armbruster, B.B., & Baker, L. (1986). The role of metacognition in reading and studying. In J. Orasanu (Ed.), *Reading comprehension: from research to practice* (pp. 49-75). Hillsdale, NJ: Erlbaum.
- Brown, A. L., & Campione, J. C. (1990). Communities of learning and thinking. *Human Development*, 21, 108-125.
- Busato, V. (1998). *Leerlijnen nader geanalyseerd*. Academisch proefschrift, Universiteit van Amsterdam.
- Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*.
- Entwistle, N. (1997). Contrasting perspectives on learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds), *The experience of learning* (2nd ed., pp. 3-22). Edinburgh: Scottish Academic Press.
- Hofer, B.K. & Pintrich, P.R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88-140.
- Marton, F., Dall'Alba, G., & Beaty, E. (1993). Conceptions of learning. *International Journal of Educational Research*, 19, 277-300.
- Meyer, J.H.F. (1991). Study orchestration: The manifestation, interpretation and consequences of contextualised approaches to studying. *Higher Education*, 22, 297-316.
- Oosterheert, I.E. (1998). *Learning styles in learning to teach*. Paper presented at the joint meeting of the EARLI-SIGs 'Higher Education' and 'Teaching and Teacher Education', Leiden, the Netherlands.
- Perry, W.G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart & Winston.
- Säljö, R. (1979). *Learning in the learner's perspective. I. Some common sense conceptions*, Reports from the Department of Education, University of Göteborg, No. 76.
- Vermunt, J. (1992). *Leerlijnen en sturen van leerprocessen in het hoger onderwijs. Naar procesgerichte instructie in zelfstandig denken*. Amsterdam/Lisse: Swets & Zeitlinger.
- Vermunt, J.D. (1996). Metacognitive, cognitive and affective aspects of learning styles and strategies: A phenomenographic analysis. *Higher Education*, 31, 25-50.
- Vermunt, J. (1999). Recente ontwikkelingen in het onderzoek naar leerstrategieën en leerlijnen. In J.G.L.C. Lodewijks & J.M.M. van der Sanden (Red.), *Op de student gericht - Een bundel opstellen over leren en studeren, opgedragen aan Prof. Dr. Len F.W. de Klerk* (pp. 85-92). Tilburg: Tilburg University Press.

Een structureel model over de relaties tussen leeropvattingen, regulatie en cognitieve verwerking

Yvonne J. Vermetten (*Open Universiteit Nederland, OTEC*), Hans G. Lodewijks (*Katholieke Universiteit Brabant, Sectie Onderwijs- en Opleidingspsychologie*) en Jan D. Vermunt (*Universiteit Maastricht, Capaciteitsgroep Onderwijsontwikkeling en Onderwijsresearch*)

ABSTRACT

This study explored explanations of individual differences in learning resulting from different underlying learning conceptions. A theoretical model, based on a systems model of personality, was proposed in which three levels of learning behaviour are discerned. In this model, learning conceptions underlie regulation strategies, which in turn underlie cognitive processing strategies. The model was evaluated by means of structural equation modelling. Significant and meaningful relations were confirmed between the constructs situated on the different levels. As learning conceptions explained a substantial part of individual differences in learning, the implication for educational practice is that these underlying factors need concurrent attention in the teaching of learning skills. Trying to change learning behaviour without modifying the underlying conceptions and convictions about learning, will probably founder.

1. INLEIDING

In deze studie is nagegaan op welke wijze individuele opvattingen van studenten over leren (zogenaamde leerconcepties, mentale leermodellen), van invloed zijn op de manier waarop zij sturing geven aan het leerproces en op de aard van de verwerkingsactiviteiten die zij ontplooiën. Inzicht in deze kwestie is van belang, onder meer omdat uit (b.v. leerstijl-) onderzoek steeds naar voren komt, dat studenten bepaalde manieren van leren prefereren, ongeacht of deze in een concrete leersituatie de meest effectieve zijn.

De aandacht voor individuele opvattingen en overwegingen als grondslag voor actueel gedrag is sterk beïnvloed door de opkomst van het constructivisme dat een belangrijke rol toekent aan de lerende zelf bij de vormgeving en inhoud van het leerproces (e.g., Shuell, 1988; Simons, 1993). Leren wordt hierbij gezien als een proces dat zich binnen het individu voltrekt en dat is gekenmerkt door het (re)construeren van schema's of conceptuele structuren (mental models).

Het belang van onderzoek dat opvattingen over leren en kennis relateert aan het studeergedrag en de prestaties van studenten wordt ook onderstreept in een recente review over epistemologische opvattingen van Hofer en Pintrich (1997). Zij stellen dat er op dit terrein nog erg weinig bekend is. Kennis over dergelijke relaties kan een belangrijke bijdrage leveren aan het onderzoek over leren en instructie. Perry (1981) suggereerde al dat als studenten andere opvattingen over kennis kregen (zie verderop in de inleiding), dit mogelijk ook zou leiden tot verandering in hun studiestrategieën.

Hofer en Pintrich trekken de vergelijking met het onderzoek naar zelfregulerend leren van Zimmerman (1994), waaruit naar voren komt dat studenten een doel of een criterium moeten

hebben, op grond waarvan ze hun gedrag reguleren (initieën, besluiten te stoppen en dergelijke). Op soortgelijke wijze kan een opvatting over leren (of kennis) fungeren als een doel of als een norm, waarmee het eigen gedrag kan worden vergeleken en getoetst. Zo kan het gedrag van een student dat bestaat uit het alsmat herhalen en repeteren van de leerstof gerechtvaardigd worden vanuit het feit dat dit volgens zijn of haar eigen opvatting er zo hoort uit te zien.

1.1 Leerstijlen-onderzoek

Het vaakst beschreven onderscheid in de wijze waarop studenten in het hoger onderwijs leren, is waarschijnlijk dat tussen oppervlakkig leren (waarbij leerstof op een reproductieve en betekenisloze manier wordt geleerd) en diepgaand leren (waarbij er wordt gezocht naar de betekenis achter de leerstof). Het proefschrift van Vermunt (1992), heeft het hieraan gerelateerde concept 'leerstijlen' volop in de aandacht gebracht. In het huidige onderzoek wordt dieper ingegaan op het model dat ten grondslag ligt aan deze leerstijltheorie, waarin relaties verondersteld worden tussen verschillende componenten van leren. Voor een beschrijving van leerstijlen volgens de theorie van Vermunt wordt verwezen naar de inleiding op dit themanummer (Ten Dam, Severiens & Vermunt, 1999). Hieronder worden de verschillende componenten van een leerstijl nog wel even kort toegelicht.

Cognitieve verwerking bestaat bijvoorbeeld uit het leggen van relaties, het structureren van leerstof of het herhalen van rijtjes. Deze activiteiten zijn direct gericht op de verwerking van de leerstof. Het sturen of reguleren van leerprocessen verwijst naar activiteiten waarmee controle over de cognitieve verwerking wordt uitgeoefend, zoals het bepalen van een strategie en het in de gaten houden van het verloop van het leerproces. Een mentaal leermodel is een geheel van concepties over het verschijnsel 'leren', zoals bijvoorbeeld de opvatting dat leren in wezen bestaat uit het opbouwen van je eigen kennis. Studiemotieven tenslotte zijn iemands persoonlijke doelstellingen, intenties, motieven, verwachtingen, enzovoort bij het leren.

1.2 Gelaagdheid van (leer)gedrag

Tussen de bovengenoemde componenten veronderstelt Vermunt onderlinge verbanden. De manier waarop ze volgens het onderliggende model (Vermunt, 1998) onderling gerelateerd zijn, is overeenkomstig met een systeem-model uit de persoonlijkheidsleer over adaptief gedrag (Hettema, 1979). Hierin worden drie niveaus van gedrag onderscheiden, te weten een cognitief-symbolisch niveau (ideeën en overtuigingen), een controle-niveau en een operationeel niveau. Het controle-niveau heeft hierin een afstemmingsfunctie, tussen het cognitief-symbolische niveau, het operationele gedragniveau en de (steeds veranderende) omgeving (zie ook Vermetten, 1999).

Als het gaat om 'leer'gedrag, verwijzen zowel het cognitief-symbolische niveau als het controle-niveau naar het domein van de metacognitie, terwijl het operationele gedragniveau verwijst naar het domein van de cognitie. Onderzoek naar metacognitie heeft zich voornamelijk gericht op (zelf)regulatievevaardigheden (zie Elen & Lowyck, 1998). Terwijl dit kan worden gezien als de procedurele kant van metacognitie, en geplaatst kan worden op het controle-niveau, vormen leerconcepties de declaratieve kant van metacognitie, gerelateerd aan het cognitief-symbolische of conceptuele niveau.

Toegepast op de componenten van leerstijlen, worden de cognitieve verwerkingsactiviteiten beschouwd als concrete gedragingen, zich aldus afspelend op het operationele gedragniveau. De sturing of regulatie van leerprocessen speelt zich af op het tussenniveau van de controle, waar afstemming plaatsvindt tussen het conceptuele niveau, het operationele niveau en de omgeving. Leeropvattingen horen tot het conceptuele of cognitief-symbolische niveau, van waaruit de onderliggende niveaus worden aangestuurd. In het huidige onderzoek wordt een structureel model getoetst waarin deze drie niveaus zijn vertegenwoordigd, met als doel verschillend leergedrag te verklaren vanuit onderliggende leeropvattingen.

1.3 Onderzoek naar leeropvattingen

Een invloedrijke studie naar de aard en ontwikkeling van opvattingen over kennis (epistemological beliefs) betreft het longitudinale onderzoek van Perry (1970) onder Harvard studenten. Hoewel Perry de opvattingen van studenten over de aard van *kennis* onderzocht, gaat recenter onderzoek ook vaak over de opvattingen van studenten over de aard van *leren* (zie ook de inleiding op dit themanummer). Hofer en Pintrich (1997) duiden op het conceptuele verschil tussen deze twee constructen (zie ook Schellings, Lodewijks en Van der Sanden, dit themanummer). Verschillende onderzoekers maken echter geen onderscheid tussen de twee constructen (b.v. Schommer, 1994; Vermunt, 1996; Marton, Dall'Alba en Beaty, 1993).

Een bekende studie naar vorm en inhoud van leerconcepties is het onderzoek van Säljö (1979, beschreven in Marton, Dall'Alba & Beaty, 1993), die vijf (hiërarchisch geordende) categorieën van leerconcepties onderscheidt. Hierop is reeds uitgebreid ingegaan in de inleiding van dit themanummer. Ook Vermunt betrok leerconcepties in de ontwikkeling van de Inventaris Leerstijlen (ILS), maar gebruikte hiervoor de term mentale leermodellen. In de huidige studie wordt gebruik gemaakt van deze ILS, waarin vijf schalen zijn opgenomen die mentale leermodellen meten (zie Ten Dam, Severiens & Vermunt, dit themanummer). Omdat twee van deze schalen niet echt *leeropvattingen* zijn, is besloten om in dit onderzoek alleen de drie opvattingen over *leren* te betrekken. Deze tonen veel gelijkheden met de leerconcepties die Säljö en Marton et al. onderscheidden.

Bij hoge scores op de ILS-schaal 'opname van kennis' zien studenten *leren* als het reproduceren van feiten en geven zij aan duidelijke en exacte instructies te prefereren. De opvatting 'leren als memoriseren en reproduceren' van de Zweedse onderzoekers heeft ongeveer dezelfde betekenis. De ILS-schaal 'leren als het opbouwen van kennis' impliceert dat men vindt dat het de taak van de lerende is om vragen te bedenken, naar relaties te zoeken, andere bronnen te raadplegen enzovoort. Deze opvatting zou in de indeling van Säljö en Marton et al. passen bij de 'hogere-orde' opvattingen, waarin transformeren en betekenis geven een centrale rol spelen.

De derde opvatting die in de ILS gemeten wordt is 'gebruik van kennis', waarbij men *leren* ziet als het verwerven van kennis met het oogmerk deze te kunnen gebruiken en in te zetten in de praktijk. Het is niet duidelijk of deze categorie zou passen bij de opvatting 'leren als toepassen', die bij Säljö en Marton et al. tot de 'onderste' drie hoorde. In dat geval zou de opvatting geassocieerd zijn met kennis opnemen en opslaan. Het kan ook zijn dat deze opvatting te maken heeft met betekenis verlenen, maar daarnaast het belang onderstreept van het gebruiken van de 'begrepen kennis'. Door de onderlinge relaties met andere constructen te onderzoeken kan het huidige onderzoek hier meer licht op werpen.

1.4 Individuele verschillen in verwerking en regulatie

Uit het onderzoek van Vermunt zijn verschillende vormen van verwerkingsstrategieën en regulatiestrategieën naar voren gekomen waarin individuele studenten kunnen verschillen (zie ook de inleiding op dit themanummer). Binnen het cognitieve verwerkingsdomein zijn dit 'diepe verwerking', 'stapsgewijze verwerking', en 'concrete verwerking'. De eerste twee categorieën kunnen verder worden onderverdeeld in twee subcategorieën. Zo omvat diepe verwerking zowel het relateren en structureren van leerstof, als het kritisch benaderen van de leerstof. Stapsgewijze verwerking verwijst naar het stap voor stap verwerken en bestuderen van details en feiten in de leerstof (in de ILS aangeduid als 'analyseren'), als ook naar het herhalen en uit het hoofd leren van de leerstof. De derde verwerkingsstrategie kan worden omschreven als 'concretiseren en personaliseren' en heeft betrekking op het gebruiken van leerstof buiten de studie en het relateren van het geleerde aan eigen ervaringen. Hoewel deze leerstrategie niet vaak wordt gerapporteerd, onderscheiden ook Geisler-Brenstein, Schmeck, en Hetherington (1996) een soortgelijke elaboratieve verwerkingsstrategie, die verband houdt met concrete, visuele, emotionele en persoonlijk gerichte informatie-verwerking en met zelf-actualisatie (persoonlijke ontwikkeling).

Ook wat betreft regulatiestrategieën zijn er drie verschillende manieren waarop studenten hun leerprocessen reguleren. Zelfregulatie betekent dat de student het eigen leerproces controleert, initieert, plant enzovoorts. Ook verwijst deze schaal naar het uit eigen beweging betrekken van andere bronnen bij de leerstof. Een tweede regulatiestrategie is externe regulatie, die duidt op het gebruiken van externe bronnen voor de regulatie van het eigen leerproces en het zich laten leiden door bijvoorbeeld aanwijzingen van de docent, doelen uit de reader, en vragen in het boek. Daarnaast onderscheidt Vermunt zogenaamde 'stuurloosheid', hetgeen in feite een problematische vorm van regulatie van het eigen leerproces inhoudt. Deze vorm kan men niet echt een strategie noemen, maar duidt eerder op een gebrekkige of ontbrekende strategie. Uit onderzoek (Vermetten, Lodewijks & Vermunt, 1999) is gebleken dat stuurloosheid samenhangt met verschillende soorten gerapporteerde problemen die studenten tijdens het leren tegenkomen, zoals problemen met tijd-management en met het voorbereiden van tentamens.

1.5 Het huidige onderzoek

Terwijl Vermunt (1998) achtereenvolgens verschillende regressiemodellen toetst om de samenhang tussen de bovengenoemde constructen te onderzoeken, maakt de huidige studie gebruik van lineaire structurele vergelijkingsmodellen. Dit heeft als voordeel dat alle veronderstelde relaties tussen de variabelen (ook de indirecte paden) tegelijk in één model kunnen worden opgenomen en dat het model vervolgens in zijn geheel kan worden geëvalueerd. Een van de doelen van dit onderzoek is om een meer omvattend en integratief beeld van de onderlinge relaties tussen verschillende componenten van de ILS te krijgen.

Het hypothetische model dat wordt onderzocht sluit aan bij het systeem-model uit de persoonlijkheidsleer zoals hierboven werd geschetst. Het theoretische uitgangspunt is dat opvattingen over leren onderliggend zijn aan de regulatie van het leren, wat op zijn beurt richting geeft aan de verwerking van de leerstof. Ook kunnen er directe verbanden bestaan tussen leeropvattingen en verwerkingsstrategieën (Vermunt, 1998). De relaties in het hypothetische model zijn gebaseerd op de eerdere resultaten van Vermunt (zie ook methode-sectie). De onderzoeksvragen die in dit artikel centraal staan kunnen als volgt worden samengevat:

- Hoe en in hoeverre kunnen onderliggende leeropvattingen, direct of indirect via regulatiestrategieën, verschillen in verwerkingsstrategieën verklaren?
- Wat dragen de gevonden relaties bij aan een beter begrip van enerzijds de inhoudelijke betekenis en anderzijds de functie van de verschillende constructen?
- Zijn de huidige resultaten vergelijkbaar met de resultaten die Vermunt (1998) rapporteerde en hoe kunnen eventuele verschillen verklaard worden?

2. METHODE

2.1 Deelnemers

Aan het onderzoek namen 750 voltijdstudenten deel, die halverwege het eerste jaar van hun universitaire studie aan de Katholieke Universiteit Brabant waren. De gegevens van deze groep waren verzameld in het kader van een promotie-onderzoek naar consistentie en variabiliteit van leren bij studenten in het hoger onderwijs (Vermetten, 1999). De totale groep bestond voor 55% uit vrouwen, en voor 45% uit mannen. De gemiddelde leeftijd van de studenten was 19,5 jaar, variërend tussen 17 tot 44. De groep studenten was verdeeld over verschillende studie-richtingen: 23% was student rechten; 7,5% studeerde letteren; 29,5% was economie-student en 40% studeerde sociologie of psychologie.

2.2 Instrumenten

De vragenlijst die voor dit onderzoek gebruikt is, is de Inventaris Leerstijlen (ILS). De ILS bestaat uit vier onderdelen waarvan er voor het huidige onderzoek drie gebruikt zijn: (1) men-

tale leermodellen, (2) metacognitieve regulatiestrategieën, (3) cognitieve verwerkingsstrategieën. De leeroriëntaties (het vierde onderdeel van de ILS) worden in deze studie buiten beschouwing gelaten, vooral omdat Vermunt (1998) aantoonde aan dat deze schalen de minste bijdrage hadden aan het verklaren van leerstrategieën. Bovendien is de vraagstelling van het huidige onderzoek hoe verschillend leergedrag kan worden verklaard vanuit verschillende leerconcepties.

De ontwikkeling van de ILS is begonnen met een interview-studie onder studenten over hun ervaringen met leren. Op basis van de interview-resultaten (die fenomenografisch zijn geanalyseerd, zie Vermunt, 1996), is een eerste versie van de ILS ontwikkeld. Door middel van factor- en betrouwbaarheidsanalyses is het aantal items gereduceerd (dit leidde tot 120 items) en zijn schalen ontwikkeld per onderdeel van de ILS (Vermunt, 1998). Voor het huidige onderzoek is een 100-item versie van de ILS gebruikt, waarbij op basis van betrouwbaarheidsanalyses ongeveer per schaal één item werd weggelaten. Omdat leeroriëntaties buiten beschouwing zijn gelaten is het huidige onderzoek gebaseerd op 75 ILS-items. De schalen van de ILS die in de huidige analyses zijn gebruikt zijn weergegeven in Tabel 1. Cronbach alpha-waarden voor de gebruikte schalen varieerden tussen .68 en .85 in de studie van Vermunt (1998).

2.3 Procedure

De vragenlijst werd aan het eind van het eerste semester, tijdens de tentamenperiode, opgestuurd met een begeleidende brief en een portvrije enveloppe. De afname van de vragenlijsten vond plaats in het kader van een breder onderzoek, zoals reeds eerder vermeld. Deelname aan het onderzoek was op vrijwillige basis. De respons voor de verschillende groepen studenten waarop de huidige analyses zijn verricht varieerde tussen 24,8% en 56,8%, met een mediaanwaarde van 42.2. Voor dergelijk survey onderzoek, zijn dit percentages die in de lijn der verwachting liggen (Babbie, 1973; Seltiz, Wrightsman and Cook, 1976).

2.4 Data-analyse

De relaties tussen de variabelen zijn onderzocht door middel van lineaire structurele vergelijkingsmodellen (SEM), met behulp van het programma AMOS 3.6. Vanuit een hypothetisch, initieel model zijn aanpassingen gemaakt, om zo te komen tot een passend model (Jöreskog & Sörbom (1993) spreken over 'model generating').

Het hypothetische, initiële model, is gepresenteerd in Figuur 1. De indeling in drie verschillende 'lagen' is gebaseerd op de in de inleiding gepresenteerde theorie. De nadere invulling van veronderstelde relaties is gebaseerd op de resultaten van Vermunt's regressie-analyses (1998). Hieruit bleek diepe verwerking vooral samen te hangen met zelfregulatie, en ook (in mindere mate) met het 'opbouwen van kennis'. Met 'opname van kennis' bestond een klein negatief verband. Een regressie-model met zelfregulatie als afhankelijke variabele liet zien dat deze (precies zoals diepe verwerking) samenhang met 'opbouwen van kennis', en negatief met 'opname van kennis'. In het hypothetische model is verondersteld dat de genoemde verbanden tussen de concepties en de diepe verwerking verklaard kunnen worden via de tussenliggende rol van zelfregulatie.

Stapsgewijze verwerking bleek volgens Vermunt samen te hangen met zowel externe regulatie, als zelfregulatie. Ook werd stapsgewijze verwerking verklaard door 'opname van kennis'. Een regressie-model met externe regulatie als afhankelijke variabele toonde aan dat 'opname van kennis' positief met deze variabele samenhang. Ook voor stapsgewijze verwerking is er in het hypothetische model vanuit gegaan dat de relaties met leerconcepties via de regulatiestrategieën lopen.

'Gebruik van kennis' verklaarde de concrete verwerkingsstrategie in aanzienlijke mate. Verder hing 'gebruik van kennis' licht positief samen met externe regulatie. Omdat deze leerconceptie geen sterke verbanden liet zien met regulatie-strategieën is ervoor gekozen om ook een rechtstreeks pad van 'gebruik' naar 'concreet' op te nemen in het initiële model. Relaties met

zelfsturing en externe sturing zijn ook opgenomen, maar dit was voornamelijk hypothetisch van aard. Tenslotte wezen Vermunt's regressie-analyses uit dat stuurloze regulatie positief samenhangt met 'opname van kennis'. Ook dit pad is vertegenwoordigd in het initiële model.

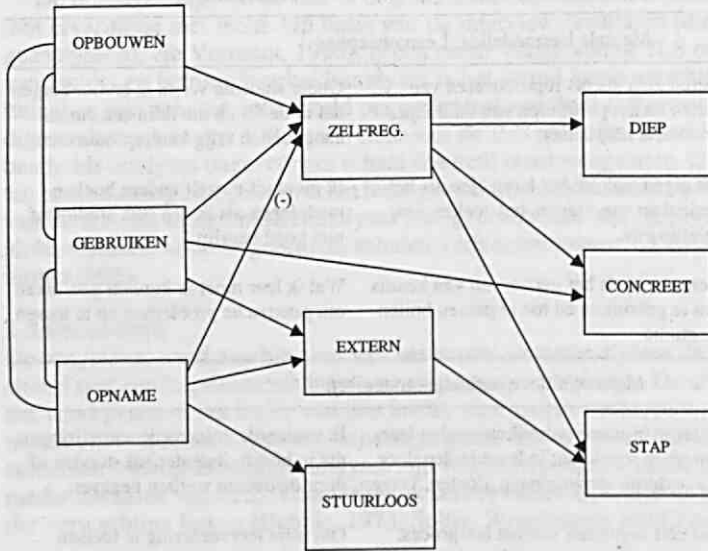
Tabel 1: Betekenis van de gebruikte ILS-schalen

Schaal en aantal items	Omschrijving van de schaal	Voorbeeld-item
Mentale leermodellen: Leerconcepties		
Opname van kennis; N=5	Leren zien als het reproduceren van feiten en het prefereren van duidelijke en exacte instructies.	Onder studeren versta ik ervoor zorgen dat ik de feiten die ik in een cursus aangeboden krijg kan reproduceren.
Opbouwen van kennis; N=5	De eigen taak bij het leren zien als het bedenken van vragen, het zoeken naar relaties enz.	Ik moet uit mezelf andere boeken raadplegen als ik een stuk studiestof niet goed begrijp.
Gebruik van kennis; N=5	Leren zien als het verwerven van kennis om te gebruiken en toe te passen buiten de studie	Wat ik leer moet ik kunnen gebruiken om praktische problemen op te lossen.
Metacognitieve regulatie-strategieën		
Externe regulatie; N=10	Externe bronnen gebruiken om het leerproces te reguleren; je laten leiden door b.v. externe aanwijzingen, doelen, vragen.	Ik studeerde volgens de aanwijzingen die in het studiemateriaal stonden of door de docent werden gegeven.
Zelfregulatie; N=10	Het zelf reguleren van het leerproces, door b.v. planning, jezelf toetsen.	Om mijn leervordering te toetsen probeerde ik een antwoord te formuleren op vragen over de studiestof die ik zelf bedacht.
Stuurloos; N=5	Het constateren van problemen met het eigen leerproces.	Ik constateerde dat ik het moeilijk vond om vast te stellen of ik de studiestof voldoende beheerste.
Cognitieve verwerkingsstrategieën		
Diepe verwerking; N=10	Het relateren en structureren van leerstof en leerstof kritisch benaderen.	Ik probeerde onderwerpen die bij een cursus afzonderlijk werden behandeld samen te brengen tot een geheel.
Stapsgewijze verwerking; N=10	Herhalen en uit het hoofd leren, stap voor stap verwerken en bestuderen van details.	Rijtjes met kenmerken van een bepaald verschijnsel prentte ik in mijn hoofd.
Concrete verwerking; N=5	Het gebruiken van leerstof buiten de studie en aandacht geven aan de praktische bruikbaarheid.	Ik probeerde gebeurtenissen in de dagelijkse werkelijkheid te interpreteren met behulp van de kennis die ik door de cursussen had verworven.

Voor het evalueren van het model, werden, naast een statistische toets, meerdere 'fit-indices' gebruikt (Hoyle, 1995). Ook is gebruik gemaakt van zogenaamde informatie-criteria (AIC en CAIC), die niet alleen een maat zijn voor de 'fit', maar die ook de spaarzaamheid (parsimony)

van het model in acht nemen. Als fit-indices zijn gebruikt: de GFI, NFI, IFI, CFI en RMR (zie Jöreskog & Sörbom, 1993). Al deze indices moeten hoger zijn dan .90 om een adequate fit aan te geven, behalve RMR, die lager moet zijn dan .05.

Figuur 1: Hypothetisch analyse-model.



Noot: *OPBOUWEN* = Opbouwen van kennis; *GEBRUIKEN* = Gebruiken van kennis; *OPNAME* = Opname van kennis; *ZELFREG.* = Zelfregulatie; *EXTERN* = Externe regulatie; *STUURLOOS* = Stuurloze regulatiestrategie; *DIEP* = Diepe verwerking; *CONCREET* = Concrete verwerking; *STAP* = Stapsgewijze verwerking.

3. RESULTATEN

Er zijn twee modellen gegenereerd, waarvan het tweede, door ons geaccepteerde model, is weergegeven in Figuur 2. In het initiële model (Model 1) bleken drie veronderstelde relaties niet significant te zijn. Dit betrof het pad van 'gebruik van kennis' naar zelfregulatie, van 'opname van kennis' naar zelfregulatie, en van 'gebruik van kennis' naar externe regulatie. Deze paden zijn in het tweede opgestelde model weggelaten. Daar de fit-indices voor het initiële model onder de maat waren, is gekeken naar de 'modification indices' die in de output van AMOS worden gegeven. Indien de gesuggereerde relaties met de hoogste waarden ook op inhoudelijke en theoretische gronden verklaarbaar waren, zijn deze toegevoegd in het tweede model. Dit gold voor drie paden: van zelfregulatie naar externe regulatie, van 'opname van kennis' naar stapsgewijze verwerking en een wederzijdse relatie tussen diepe verwerking en concrete verwerking. De statistische toets en de fit-indices behorend bij de modellen zijn weergegeven in Tabel 2.

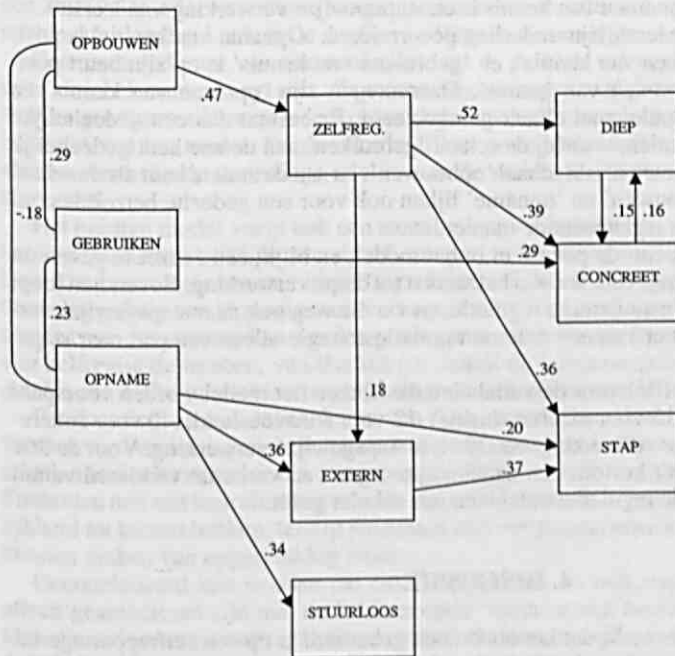
De χ^2 -toets liet een duidelijke verbetering zien ten opzichte van het initiële model. Deze toets is zeer gevoelig voor de steekproefgrootte en zal bijna nooit een model accepteren, hetgeen ook in onze studie het geval bleek. Echter, als de ratio χ^2/df kleiner is dan vijf, duidt dit volgens Marsh en Hocevar (1985) op een aanvaardbaar model. De ratio χ^2/df voor het tweede

model was 5.2. De fit indices, die nu alle hoger lagen dan .90 (en voor RMR: lager dan .05) duiden op een aanvaardbare fit. De informatie-criteria lieten ook een behoorlijk verbetering voor het tweede model zien. Ook in vergelijking met het nulmodel (waarin géén onderlinge relaties worden verondersteld) past het tweede model veel beter op de data. Gezien deze indicaties werd besloten model 2 als definitief model te aanvaarden.

Tabel 2: Statistische toetsing, fit indices en informatie-criteria voor het nulmodel en de geteste modellen

	Statistische toetsing			Fit Indices				Informatie-criteria	
	χ^2/df	p-waarde	GFI	RMR	NFI	IFI	CFI	AIC	CAIC
Nulmodel	1834.31/ 36 = 51	0.000	.61	.101	.00	.00	.00	1852.3	1902.6
Model 1	254.99/ 22 = 11.6	0.000	.93	.029	.86	.87	.87	300.99	429.60
Model 2	115.14/ 22 = 5.2	0.000	.97	.020	.94	.95	.95	161.14	289.75

Figuur 2: Maximum likelihood schatters (gestandaardiseerd) voor het geaccepteerde SEM model (Model 2).



Noot. OPBOUWEN = Opbouwen van kennis; GEBRUIKEN = Gebruiken van kennis; OPNAME = Opname van kennis; ZELFREG. = Zelfregulatie; EXTERN = Externe regulatie; STUURLOOS = Stuurloze regulatie-strategie; DIEP = Diepe verwerking; CONCREET = Concrete verwerking; STAP = Stapsgewijze verwerking.

3.1 Beschrijving van het definitieve model

In Figuur 2 is te zien dat de leerconceptie 'opbouwen van kennis' een indirect verband heeft met het op een diepe manier verwerken van leerstof. Het indirecte pad loopt via zelfregulatie, wat er op duidt dat het idee dat leren bestaat uit het opbouwen van eigen kennis samenhangt met de neiging om het leerproces zelf te sturen, wat op zijn beurt een positief effect heeft op diepe verwerking (beide paden hebben de hoogste regressie-coëfficiënt, duidend op de grootste effecten in het model).

De leerconceptie 'gebruiken van kennis' heeft een direct effect op de concrete verwerking van leerstof. Er is hierbij geen sprake van een indirect effect via een regulatiestrategie. Iemands neiging tot zelfregulatie (die samenhangt met 'opbouw van kennis') heeft ook een (iets sterker) effect op concrete verwerking van leerstof. Het concretiseren en personaliseren van leerstof heeft dus zowel te maken met de leerconceptie dat je kennis moet gebruiken, als met het idee dat je je eigen kennis opbouwt en de neiging tot zelfregulatie. Verder is er een tweezijdige relatie tussen diepe verwerking en concrete verwerking. Het bovenste gedeelte van het model (bestaande uit 'opbouwen', 'gebruiken', 'zelfregulatie', 'diepe verwerking' en 'concrete verwerking') lijkt een onderling (positief) samenhangend geheel te zijn. Dit blijkt verder nog uit de positieve correlatie tussen de twee leerconcepties 'opbouwen van kennis' en 'gebruiken van kennis'.

De opvatting dat leren bestaat uit het (passief) opnemen van kennis is gerelateerd aan het zich richten op externe bronnen voor de regulatie van het leerproces. Daarnaast heeft deze opvatting een effect op een stuurloze regulatie. Stuurloze regulatie is een soort eindpunt, in de zin dat dit niet samenhangt met het op een bepaalde manier verwerken van leerstof. Externe regulatie doet dat wel: de neiging van een student om externe regulatie te gebruiken is een indicatie voor een stapsgewijze verwerking van leerstof. Er is ook een directe samenhang tussen de opvatting dat leren het opnemen van kennis is en stapsgewijze verwerking van leerstof.

De drie opvattingen over leren zijn onderling gecorreleerd. 'Opname van kennis' is positief gerelateerd aan 'gebruiken van kennis', en 'gebruiken van kennis' is op zijn beurt positief gecorreleerd met 'opbouwen van kennis'. Daarentegen zijn 'opname van kennis' en 'opbouwen van kennis' negatief met elkaar geassocieerd. Er bestaat dus een gedeeltelijke overlap tussen deze drie schalen, waarbij de schaal 'gebruiken' aan de ene kant gedeeltelijk hetzelfde construct lijkt te meten als de schaal 'opbouwen', en aan de andere kant als de schaal 'opname'. De schalen 'opbouwen' en 'opname' lijken ook voor een gedeelte hetzelfde construct te meten, maar op een tegengestelde manier.

Zelfregulatie neemt een centrale positie in in het model, en blijkt een aanzet te geven tot alle soorten verwerkingsstrategieën, hoewel het sterkst tot diepe verwerking. Bovendien loopt er een pad van zelfregulatie naar externe regulatie, en via die weg ook naar stapsgewijze verwerking. Daarentegen loopt er van een externe regulatiestrategie alleen een pad naar stapsgewijze verwerking.

De verklaarde varianties (R^2) voor de variabelen die binnen het model worden voorspeld zijn .22 voor zelfregulatie; .15 voor externe sturing; .12 voor stuurloosheid; .40 voor concrete verwerking; .49 voor diepe verwerking; en .39 voor stapsgewijze verwerking. Voor de drie verwerkingsstrategieën wordt kortom een aanzienlijke mate van variantie verklaard vanuit het model, terwijl dit voor de regulatiestrategieën iets minder geldt.

4. DISCUSSIE

Als eerste moet worden opgemerkt dat het onderzoek gebaseerd is op een zelfrapportage-instrument, waaraan een belangrijk nadeel is verbonden. De ILS meet niet het daadwerkelijke leergedrag van studenten, maar alleen de ideeën die studenten zelf hebben over de door hen gehanteerde strategieën, opvattingen enzovoort. Hoewel zelfrapportages vaak gebruikt worden in (onderwijs)onderzoek, moet deze beperking worden onderkend.

De belangrijkste vraagstelling van het onderzoek betrof hoe verschillende verwerkingsstrategieën kunnen worden verklaard vanuit onderliggende leerconcepties en regulatiestrategieën. Het gevonden model blijkt goed aan te sluiten bij de resultaten die Vermunt (1998) rapporteerde, maar geeft een meer integratief beeld, dat mede de gelaagdheid van het systeemmodel uit de persoonlijkheidsleer weerspiegelt. De aanzienlijke verklaarde varianties van de verwerkingsstrategieën door het model geven aan dat de onderliggende variabelen belangrijke voorspellers zijn. Het lijkt er dus op dat verschillen die geconstateerd worden in het studiegedrag van studenten voor een groot deel verklaard kunnen worden vanuit de verschillende opvattingen die studenten hebben over wat leren is en vanuit hun regulatiestrategieën.

Dat de opvatting 'gebruiken van kennis' samenhangt met een concrete manier van verwerken van leerstof -die zoal inhoudt dat men probeert relaties te leggen met de praktijk- is een logisch verklaarbaar verband. Dit geldt ook voor het zien van leren als het reproduceren van feiten ('opname van kennis') dat gerelateerd is aan het herhalend en stap voor stap te werk gaan. Deze resultaten komen overeen met de resultaten uit de regressievergelijkingen van Vermunt, die ook vond dat 'gebruik van kennis' en 'opname van kennis' goede voorspellers zijn van de twee genoemde verwerkingsstrategieën. Vermunt vond geen sterke samenhang tussen 'opbouw van kennis' en diepere verwerking, maar wel tussen zelfregulatie en diepe verwerking. De indirecte relatie tussen 'opbouw van kennis' en diepe verwerking, via zelfregulatie, verklaart deze eerdere resultaten uit Vermunt's regressiemodellen.

De verbanden van 'opname van kennis', met externe sturing, en van externe sturing met stapsgewijze verwerking komen eveneens overeen met de resultaten van Vermunt. De relatie tussen 'opname van kennis' en stuurloosheid lijkt in het huidige onderzoek sterker dan in de resultaten van Vermunt. Dit kan mogelijk verklaard worden door een aantal variabelen die niet in het huidige model zijn opgenomen, maar wel in de regressievergelijkingen zaten (nl. een ambivalente leeroriëntatie, samen studeren en stimulerend onderwijs). Overigens voorspelde stuurloosheid ook in de studie van Vermunt nauwelijks verwerkingsstrategieën.

Een resterend punt van overeenkomst was de centrale rol van zelfregulatie, in de zin dat deze variabele samenhangt met alle soorten verwerkingsstrategieën. Dat de resultaten uit beide studies goed met elkaar overeenkomen (bij verschillende steekproeven) pleit voor de houdbaarheid van het model. Enige zekerheid kan zodoende worden gegeven over de generaliseerbaarheid van de onderlinge relaties van de variabelen binnen de groep van universitaire studenten.

Het huidige model voegt ook een aantal nieuwe inzichten toe aan de eerdere resultaten. In het structurele model blijkt een significant pad te lopen van zelfregulatie naar externe regulatie. Dit is een interessante uitbreiding van de resultaten van Vermunt, omdat dit suggereert dat zelfregulatie een strategie van een hogere orde is dan externe regulatie. Het verband tussen de twee constructen kan worden geïnterpreteerd in de zin dat studenten die een hoge mate van zelfregulatie hebben, van daaruit (of vanuit de leerconceptie 'opbouwen' die onderliggend is aan zelfregulatie) kunnen besluiten om ook op externe bronnen van regulatie te varen.

Dit doet denken aan de resultaten van Greene en Miller (1996) die een (onverwacht) pad vonden van betekenisgericht leren naar oppervlakkig leren. Zij verklaarden dit vanuit de verschillende doeloriëntaties die ten grondslag lagen aan de verschillende vormen van leren. Studenten met een leeroriëntatie zouden een breed repertoire aan strategieën, en dus een mogelijkheid tot kiezen hebben, terwijl studenten met een prestatieoriëntatie alleen gebruik zouden kunnen maken van oppervlakkig leren.

Geconcludeerd kan worden dat externe regulatie, en ook stapsgewijze verwerking, niet alleen geassocieerd zijn met de leerconceptie 'opname van kennis'. Beide kunnen ook verklaard worden vanuit zelfregulatie, en vanuit de leerconceptie 'opbouwen van kennis'. Hierdoor krijgen deze vormen van verwerking en regulatie een meer positieve betekenis, in de zin dat zij een bewust onderdeel kunnen zijn van een veelzijdige manier van leren.

Ook werpt het model meer licht op de in de inleiding genoemde onduidelijkheid van de betekenis van 'gebruik van kennis'. Er blijkt inderdaad een dubbele betekenis met deze schaal

verbonden te zijn, die gekenmerkt wordt door enerzijds een overlap met de schaal 'opname van kennis', en anderzijds met de schaal 'opbouw van kennis'. Het onderscheidende van deze conceptie ligt in het belang dat wordt gehecht aan het gebruiken en benutten van de geleerde kennis. Dit kan zowel gepaard gaan met een meer schoolse opvatting over leren, als met een meer constructivistische opvatting over leren.

De veronderstelde positie van de regulatiestrategieën op het controle- of afstemmingsniveau wordt door het model bevestigd. Er is echter geen uitputtende tussenliggende rol voor de regulatiestrategieën, daar er ook directe verbanden voorkwamen tussen leeropvattingen en verwerkingsstrategieën. Zulke directe verbanden werden gevonden tussen enerzijds de opvatting 'gebruik van kennis' en een concrete manier van verwerken, en anderzijds de opvatting 'opname van kennis' en een stapsgewijze manier van verwerken. Hiervoor bleek de samenhang voor een deel los te staan van hoe een student zijn leerproces reguleert. Dit kan er mogelijk op duiden dat het gebied van regulatie (de procedurele kant van metacognitie) met de huidige constructen nog niet volledig in kaart is gebracht. Het kan ook betekenen dat deze twee, meer 'objectivistische' (waarbij men denkt dat kennis objectief en absoluut is en leren een passief absorberend proces) opvattingen -waarbij 'gebruik van kennis' slechts in zekere mate die betekenis heeft-, met zich mee brengen dat er in mindere mate een bewuste, actieve afstemming wordt bewerkstelligd tussen de eigen overtuigingen en de interpretatie van de omgeving.

Wat een ongunstig scenario lijkt voor een student om in te verkeren, is als alleen het onderste gedeelte van het model 'in werking' is. Als het idee dat leren het opbouwen van kennis is niet of minimaal aanwezig is, en ditzelfde geldt voor zelfregulatie, is er weinig kans op een gevarieerde manier van leren, en zal er eerder sprake zijn van eenzijdige stapsgewijze verwerking, externe sturing, en/of stuurloosheid en het ondervinden van problemen. Aan de andere kant lijken zelfregulatie en de leerconcepties 'opbouw' en 'gebruik van kennis' (de laatste in mindere mate) voorboden te zijn van een veelzijdige manier van studeren.

4.1 Implicaties voor het onderwijs

Een belangrijke praktische consequentie van de resultaten is dat de kwaliteit van de leerprocessen van studenten in het hoger onderwijs waarschijnlijk alleen verhoogd kan worden door hierbij de onderliggende factoren te betrekken. Verschillende opvattingen over wat leren is, en over hoe men dit proces ziet, lijken namelijk in belangrijke mate te bepalen op welke manier de student te werk gaat bij het leren. Iemand proberen over te halen tot een ander soort gedrag, terwijl dit niet in overeenstemming zou zijn met zijn of haar overtuiging heeft weinig kans van slagen. De onderliggende opvattingen 'sturen' het feitelijke gedrag en daarom zal er ook op dit conceptuele niveau verandering moeten plaatsvinden, om de kwaliteit van leerprocessen te kunnen beïnvloeden.

Hiertoe kunnen de signalen die in het onderwijs worden afgegeven een bijdrage leveren. Op dit moment is het nog in te veel onderwijssituaties de gewoonte om tentamens af te nemen waarin de 'kwantiteit' van de kennis centraal staat (Dahlgren, 1997). Vragen naar 'de kenmerken van...', 'welke auteurs onderschrijven..', of 'welke drie niveaus zijn er te onderscheiden...' meten alleen de hoeveelheid geleerde feiten. Het zou de 'consumptieve' opvattingen van studenten over leren kunnen opschudden als tentamens het geven van betekenis en het begrip van theorieën centraal zouden stellen (bijvoorbeeld door te vragen naar omschrijvingen in eigen woorden, te vragen naar relaties met eigen ervaringen, of te vragen naar de betekenis van een afbeelding in een bepaalde context).

Volgens Smit, Hulshof en Willems (1994) kan een docent een sterke invloed hebben op de beeldvorming van de student over de studie en het vakgebied. De docent draagt bijvoorbeeld uit wie de verantwoordelijkheid voor het leerproces heeft, en hoe de 'taken' tussen docent en student verdeeld zijn. Studenten ontdekken de onderliggende normen, waarden en verwachtingen van de omgeving. De docent, maar ook de inrichting van het onderwijs, en zoals hierboven genoemd de aard van de tentamens, communiceren dit soort zaken.

Als docenten dus van belang zijn in het stimuleren van de ontwikkeling van leerconcepties zullen ook hun eigen denkbeelden van belang zijn. Parallel met het idee dat leerconcepties ten grondslag liggen aan leerstrategieën, hebben docenten en makers van onderwijs, bepaalde concepties en overtuigingen die ten grondslag liggen aan hoe zij onderwijs (vorm)geven (Kember, 1997). Als een docent de opvatting heeft dat onderwijs bestaat uit het overdragen van kennis, en daarbij voornamelijk inhoudelijk gericht is, zal dit zijn gedrag overeenkomstig beïnvloeden. Om het onderwijs te vernieuwen en meer studentgeoriënteerd te maken (gericht op het leren in plaats van alleen op inhoud overdragen), is het volgens Kember noodzakelijk dat ook de concepties van de docenten worden aangeprepen.

Als mogelijke methoden om lerenden en docerenden bewust te maken van, en te laten nadenken over hun leer- en onderwijsconcepties, opdat deze zich kunnen ontwikkelen, worden in de literatuur verschillende mogelijkheden genoemd. Kritische reflectie op de eigen ideeën, en expliciete evaluatie van denkprocessen, waarin de bewustwording van de onderliggende opvattingen centraal staan zijn directe manieren voor het ontwikkelen van leer- en onderwijsconcepties (Mezirow, 1991; Donald, 1992; Vosniadou, 1994). Wij staan voor de uitdaging om deze moeilijk grijpbare, maar belangrijk factoren aandacht te geven binnen de onderwijspraktijk.

REFERENTIES

- Babbie, E.R. (1973). *Survey Research Methods*. Belmont, CA: Wadsworth.
- Dahlgren, L. (1997). Learning conceptions and outcomes. In F. Marton, D. Hounsell, & N. Entwistle (eds.). *The Experience of Learning: Implications for teaching and studying in higher education (2nd ed.)* (p. 23-38). Edinburgh, Great Britain: Scottish Academic Press.
- Donald, J.G. (1992). The development of thinking processes in postsecondary education: Application of a working model. *Higher Education*, 24, 413-430.
- Elen, J., & Lowyck, J. (1998). Students' views on the efficiency of instruction: An exploratory survey of the instructional metacognitive knowledge of university freshmen. *Higher Education*, 36, 231-252.
- Geisler-Brenstein, E., Schmeck, R. R., & Hetherington, J. (1996). An individual difference perspective on student diversity. *Higher Education*, 31, 73-96.
- Greene, B.A., & Miller R.B. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology*, 21, 181-192.
- Hettema, P. J. (1979). *Personality and adaptation*. Amsterdam: North-Holland Publishing Co.
- Hofer, B.K., & Pintrich, P.R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88-140.
- Hoyle, R. H. (Ed.). (1995). *Structural equation modeling: Concepts, issues, and applications*. Thousand Oaks, Calif.: Sage Publications.
- Jöreskog, K., & Sörbom, D. (1993). *Lisrel8: Structural equation modelling with the SIMPLIS command language*. Chicago: Scientific Software International.
- Kember, D. (1997). A reconceptualisation of the research into university academics' conceptions of teaching. *Learning and Instruction*, 7, 255-275.
- Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, 97, 562-582.
- Marton, F., Dall'Alba, G., & Beaty, E. (1993). Conceptions of learning. *International Journal of Educational Research*, 19, 277-300.
- Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco: Jossey-Bass.
- Perry, W.G. (1970). *Forms of intellectual and ethical development in the college years - A scheme*. New York: Holt, Rinehart & Winston.
- Perry, W.G. (1981). Cognitive and ethical growth: The making of meaning. In A. Chickering (ed.). *The modern American college* (p. 76-116). San Francisco: Jossey-Bass.
- Säljö, R. (1979). Learning in the learner's perspective. I. Some common-sense conceptions. *Reports from the Department of Education, University of Göteborg*, No. 76.
- Schommer, M. (1994). An emerging conceptualization of epistemological beliefs and their role in learning. In R. Garner & P. A. Alexander (eds.). *Beliefs about text and instruction with text* (p. 25-40). Hillsdale, NJ: Lawrence Erlbaum Ass.

- Seltiz, C., Wrightsman, L. S., and Cook, S. W. (1976). *Research Methods in Social Relations, 3e*. New York: Holt, Rinehart & Winston.
- Shuell, T.J. (1988). The role of the student in learning from instruction. *Contemporary Educational Psychology, 13*, 276-295.
- Simons, P.R.J. (1993). Constructive learning: The role of the learner. In T.M. Duffy, J. Lowyck, & D.H. Jonassen (eds.). *Designing environments for constructive learning* (p. 291-313). Berlin: Springer.
- Smit, N.J., Hulshof, M.J.F. & Willems, J.M.H.M (red.) (1994). *Beïnvloeden van studeergedrag*. Utrecht: VSNU.
- Vermetten, Y. (1999). Consistency and variability of student learning in higher education. [Academisch proefschrift]. Tilburg: Katholieke Universiteit Brabant.
- Vermetten, Y., Lodewijks, J., & Vermunt, J. (1999). Consistency and variability of learning strategies in different university courses. *Higher Education, 37*, 1-21.
- Vermunt, J. D. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs - Naar procesgerichte instructie in zelfstandig denken*. Amsterdam: Swets & Zeitlinger.
- Vermunt, J. D. (1996). Metacognitive, cognitive and affective aspects of learning styles and strategies: A phenomenographic analysis. *Higher Education, 31*, 25-50.
- Vermunt, J. D. (1998). The regulation of constructive learning processes. *British Journal of Educational Psychology, 68*, 149-171.
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*, 45-69.
- Zimmerman, B.J. (1994). Dimensions of academic self-regulation: A conceptual framework for education. In D.H. Schunk & B.J. Zimmerman (eds.). *Self-regulation of learning and performance: Issues and educational applications* (p. 3-21). Hillsdale, NJ: Erlbaum.

Leerconcepties en leerstrategieën cross-sectioneel vergeleken

Sabine Severiens en Geert ten Dam
Instituut voor de Lerarenopleiding, Universiteit van Amsterdam

ABSTRACT

The present study investigates stability of learning strategies and learning conceptions in a cross-sectional design. Vermunt's learning style theory (1992) is used to examine these learning concepts. Stability is interpreted in two ways: the first refers to changes in mean scores, and the second interpretation refers to changes in the relation between conceptions and strategies. Three groups of students are compared: 1) students in their fourth year of secondary education, 2) students in their sixth and final year of secondary education and 3) a group of higher education students. The secondary school students are all adult education students, which means they are in age comparable to the higher education students. This makes it possible to disentangle the possible effects of level of education and age. The results show that age as well as level of education have independent effects on certain types of learning strategies and certain types of learning conceptions. This means that not all strategies and conceptions are equally dependent on the level of education or age. The same holds true for changes in the relation between strategies and conceptions (the second stability interpretation). In the conclusions, some important directions for future research are described.

1. INLEIDING

In de afgelopen decennia is veelvuldig de vraag gesteld in hoeverre de manier waarop studenten leren stabiel dan wel variabel is. Het antwoord dat hierop wordt gegeven hangt in het algemeen af van de leertheorie of het leerconstruct waarop men zich baseert (zie ook Curry, 1983). In het debat omtrent stabiliteit spelen de begrippen leerstijl en leerstrategie een centrale rol. De meeste deelnemers aan het debat zijn het erover eens dat leerstijlen relatief stabiel zijn, terwijl leerstrategieën daarentegen gevoelig zijn voor instructie en training (Messick, 1996, p.639).

In Nederland is de discussie over het leren van studenten voor een belangrijk gedeelte gevoerd aan de hand van de Inventaris Leerstijlen (ILS) van Vermunt (1992). Het concept leerstijlen is in de ILS opgebouwd uit vier componenten die onderling een samenhangend geheel vormen: cognitieve verwerkingsstrategieën (de feitelijke studeeractiviteiten), regulatiestrategieën (meta-cognitie of sturing), mentale leermodellen of leerconcepties (de visie op leren en onderwijs), en leeroriëntaties (de reden om naar school te gaan en te leren). In navolging van Entwistle (1981) worden leerstijlen door Vermunt gedefinieerd als het resultaat van de interactie tussen persoons- en omgevingsgebonden invloeden (1992, p.21). Deze conceptualisatie van leerstijlen lijkt een tussenpositie in te nemen in termen van stabiliteit en veranderbaarheid. Doordat leerstijlen gedefinieerd worden als een samenhangend geheel van verschillende componenten, kan stabiliteit bovendien op tenminste twee manieren geïnterpreteerd worden. Enerzijds kan de vraag gesteld worden in hoeverre lerenden hun manier van leren

veranderen gedurende een bepaalde periode: is het bijvoorbeeld zo dat het gebruik van bepaalde verwerkingsstrategieën toeneemt of juist afneemt gedurende die periode? Met andere woorden, is er een verandering te constateren in de gemiddelde scores op de verschillende componenten? Anderzijds kan stabiliteit geïnterpreteerd worden in termen van verbanden tussen de leerstijlcomponenten. In dit laatste geval is de vraag in hoeverre de verschillende componenten op dezelfde wijze en in dezelfde mate blijven correleren gedurende een bepaalde periode.¹ In dit artikel richten we ons op beide aspecten van stabiliteit in het leren van studenten. Daarbij wordt met name ingegaan op - de samenhang tussen - de componenten leerstrategieën en leerconcepties. Het onderzoek dat wordt gepresenteerd heeft betrekking op samenhangen in verschillende groepen studenten, er is geen vergelijking gemaakt binnen personen. De term 'stabiliteit' (en 'verandering') wordt dus gebruikt in de context van een cross-sectioneel onderzoek en dient ook in die context geïnterpreteerd te worden.

Alhoewel er een respectabele hoeveelheid onderzoek is verricht naar leergedrag of leerstrategieën en leerconcepties, richten niet veel studies zich op ontwikkelingen in het leren van leerlingen en studenten. Dit betreft zowel ontwikkelingen op het niveau van 'leerstijlen' als ontwikkelingen op het niveau van de afzonderlijke concepten. In Nederland hebben Busato, Prins, Elshout en Hamaker (1998) een cross-sectioneel en longitudinaal onderzoek in het hoger onderwijs verricht met behulp van de Inventaris Leerstijlen. In het cross-sectionele onderzoek zijn verschillen te zien in de betekenisgerichte leerstijl: eerstejaars en derdejaars psychologiestudenten scoren hoger op de betekenisgerichte leerstijl vergeleken bij de vierdejaars studenten. In het longitudinale onderzoek daarentegen wordt een positieve ontwikkeling gevonden in de betekenisgerichte en toepassingsgerichte leerstijl, en een negatieve ontwikkeling in de reproductiegerichte en ongerichte leerstijl. Waarom de resultaten uit het cross-sectionele onderzoek en longitudinale onderzoek met elkaar in tegenspraak lijken, blijft onbelicht. Omdat in deze studie leerstijl als een combinatie van onder andere verwerkingsstrategieën en leerconcepties is opgevat, blijft ook onduidelijk hoe de beide concepten zich afzonderlijk ontwikkelen. Op basis van het verrichte onderzoek kan wel de conclusie worden getrokken dat er veranderingen optreden in de manier waarop studenten leren gedurende hun academische loopbaan (zie ook Dippelhofer-Stiem, 1989).

Over de tijdspanne waarin zich veranderingen in leergedrag voordoen, merken Pinto en Geiger (1991) op basis van hun onderzoeksresultaten op dat er binnen één studiejaar nauwelijks ontwikkelingen te zien zijn in de zogenaamde leerstijlpreferenties. Ook de verschillen in drie jaar tijd zijn nog steeds marginaal. Dat veranderingen in leren zich over een relatief lange periode uitstrekken, stellen ook Wigfield, Pintrich en Eccles (1996) in hun review van onderzoek naar ontwikkelingen tussen het elfde en vijftiende levensjaar. Gedurende deze periode ontwikkelen leerlingen/studenten hun gebruik van strategieën en wordt een beter zicht op zichzelf als lerende ontwikkeld. Door het ontbreken van longitudinaal of cross-sectioneel onderzoek naar mogelijke ontwikkelingen in de overstap van het voortgezet onderwijs naar het hoger onderwijs, is het beeld evenwel weinig gedetailleerd.

Een zoektocht naar studies waarin ontwikkelingen in leerconcepties of aanverwante begrippen zijn onderzocht, levert voornamelijk studies op waarin kennisconcepties - of epistemologische opvattingen - centraal staan. Slechts enkele studies richten zich specifiek op ontwikkelingen in leerconcepties (bijv. Marton, Dall'Alba en Beaty, 1993). De basis voor deze studies is gelegd door Säljö (1979). Hij beschreef vijf kwalitatief verschillende leerconcepties die, globaal gezien, kunnen worden onderverdeeld in een reproductieve visie versus een constructieve visie op leren. Ook Vermunt's leerconcepties bouwen voort op deze studie van Säljö. De ontwikkeling loopt van de reproductie naar de constructieve leerconcepties. Zoals aangegeven zijn meer studies verricht naar ontwikkelingen in kennisconcepties. Hofer en Pintrich (1997) betogen dat epistemologische opvattingen - opvattingen over kennis, leren en weten - naar een breder begrip verwijzen dan leerconcepties. Beide concepten zijn volgens hen echter wel duidelijk met elkaar vervlochten. De conclusies die te trekken zijn op basis

van onderzoek naar ontwikkelingen in kennisconcepties, zouden daarom wellicht ook betrekking kunnen hebben op leerconcepties. Bij een dergelijke gevolgtrekking is wel enige voorzichtigheid geboden. In Severiens en Ten Dam (1998) wordt een analyse gemaakt van de verbanden tussen kennis- en leerconcepties: terwijl op conceptueel niveau kennis- en leerconcepties aan elkaar gerelateerd zijn, worden deze verbanden op empirisch niveau niet teruggevonden. Voor het empirisch uitblijven van de verwachte verbanden was evenwel een belangrijke alternatieve verklaring te geven (verschillende methoden voor het meten van kennisconcepties en leerconcepties). Met de nodige slagen om de arm presenteren we daarom toch een kort overzicht van onderzoek naar ontwikkelingen in kennisconcepties en epistemologische ideeën.

In Wigfield *et al.* (1996) wordt beschreven dat de epistemologische ontwikkeling van lerenden weliswaar wordt vervolgd in het hoger onderwijs, maar reeds start in het voortgezet onderwijs. Verreweg de meeste studies beperken hun scope echter tot het hoger onderwijs (bijv. Perry, 1970; Baxter Magolda, 1992). Uitzonderingen hierop vormen onder andere de studies van Schommer (1993; Schommer, Calvert, Garigliette & Baja, 1997) waarin wordt laten zien dat de ontwikkeling in opvattingen over kennis wordt ingezet gedurende de adolescentie. In een cross-sectioneel onderzoek vond Schommer (1993) dat een periode van vier jaar nodig lijkt om te komen tot een verandering in epistemologische opvatting. Dit beeld wordt bevestigd in een longitudinale studie (Schommer *et al.*, 1997). In algemene termen karakteriseren Wigfield *et al.* epistemologische ontwikkeling als een ontwikkeling van het geloven in de absolute zekerheid van kennis, naar het in twijfel trekken van kennis en het daarbij betrekken van de context als belangrijke factor. Severiens en Ten Dam (1998) relateren deze verschillen in opvattingen van kennis aan respectievelijk de leerconcepties 'opname van kennis' en 'opbouw van kennis' uit de theorie van Vermunt. Als er in navolging van epistemologische ontwikkelingstheorieën er dus ook wat leerconcepties betreft sprake is van een ontwikkeling, dan kan verondersteld worden dat lerenden gedurende hun onderwijsloopbaan leren steeds minder in termen van 'opname van kennis' gaan zien, en steeds meer in termen van 'opbouw van kennis'.

Behalve de ontwikkeling in (de gemiddelde scores op) respectievelijk leerstrategieën en leerconcepties als componenten van leerstijlen, verwijst stabiliteit tevens naar de ontwikkeling in de verbanden tussen leerstrategieën en leerconcepties. Ook wat deze tweede interpretatie van stabiliteit betreft, is er relatief weinig empirisch onderzoek beschikbaar. Op conceptueel niveau worden er regelmatig parallellen getrokken tussen enerzijds het onderzoek naar leergedrag, dat veelal terugrijpt op het werk van Entwistle (1981) en Schmeck (1983), en anderzijds het onderzoek naar leerconcepties, waarvoor het fenomenografische onderzoek van Marton en Säljö (1976a en b) meestal het uitgangspunt vormt. De parallel bestaat eruit dat de diepte-oppervlakte dimensie van leren aan zowel onderzoek naar leergedrag als aan onderzoek naar leerconcepties ten grondslag ligt. Deze constatering heeft echter nog maar mondjesmaat zijn weerslag gevonden in empirisch onderzoek naar het verband tussen leergedrag en leerconcepties. Volgens Hofer en Pintrich (1997) is de vraag in hoeverre kennisconcepties (die bij hen ten dele ook leerconcepties omvatten) van invloed zijn op leerstrategieën en zelfregulatie echter één van de meest relevante vragen voor toekomstig onderzoek naar kennisconcepties.

Eén van de weinige onderzoeken waarin het verband tussen de opvattingen van lerenden over leren en onderwijs en hun feitelijke leergedrag empirisch is nagegaan, is dat van Purdie, Hattie en Douglas (1996). Zij verrichtten een onderzoek in het voortgezet onderwijs in Australië en vonden geen verbanden tussen de verschillende leerconcepties en leerstrategieën van leerlingen. Wel vonden ze een globale correlatie tussen de betekenisgerichte leerconcepties enerzijds en de gehele groep leerstrategieën anderzijds. Dit zou erop kunnen wijzen dat verbanden tussen concepties en strategieën in beginsel aanwezig zijn, maar in deze groep lerenden nog niet zijn uitgekristalliseerd. Deze veronderstelling wordt ondersteund door Vermunt en Van der Sanden (1996), in hun overzicht van onderzoeken die gebruik maken van de ILS.

Terwijl in het hoger onderwijs vrijwel steeds dezelfde factorpatronen, en dus onderlinge verbanden, zijn gevonden (bijv. Prins, Busato, Hamaker & Visser, 1996; Schouwenburg, 1996), blijkt dit in het voortgezet onderwijs niet zonder meer het geval te zijn (bijv. Roosendaal & Vermunt, 1996; Severiens & Ten Dam, 1996; Boekaerts, Otten & Simons, 1997). Met name de betekenisgerichte en de toepassingsgerichte leerstijl leveren bij onderzoek naar leerlingen in het voortgezet onderwijs problemen op wat betreft de onderlinge verbanden tussen de leerstijlonderdelen. Zo zijn de verbanden tussen 'leerconcepties' en 'verwerkingsstrategieën' minder sterk in het voortgezet onderwijs dan in het hoger onderwijs (Klatter, 1995; Roosendaal & Vermunt, 1996). De opvattingen van lerenden in het voortgezet onderwijs over leren en onderwijs (hun leerconceptie) zijn niet direct gerelateerd aan hun leergedrag. In het hoger onderwijs daarentegen, lijkt dit wel het geval te zijn. Dit wordt ook bevestigd in het onderzoek van Lonka en Lindblom-Ylänne (1996) onder Finse studenten in het hoger onderwijs, waarin eveneens gebruik is gemaakt van de Inventaris Leerstijlen. Ook in dit onderzoek worden, vergelijkbaar met Vermunt (1998), verbanden gevonden tussen de manier waarop studenten hun studie aanpakken en hun leerconcepties. Vermunt (1995; zie ook Roosendaal & Vermunt, 1996) interpreteert de gevonden verschillen tussen het voortgezet en hoger onderwijs in termen van een ontwikkelingsproces: gaandeweg de opleiding wordt het verband tussen concepties en leergedrag sterker.

Alvorens de onderzoeksvragen te formuleren van de studie waarover in dit artikel wordt gerapporteerd, is er nog één aspect waar in het kader van deze inleiding aandacht aan besteed moet worden. In veel onderzoeken waarin veranderingen worden geconstateerd in de manier waarop studenten leren, blijft onduidelijk of het gaat om een relatief autonome (persoonsgebonden) ontwikkeling, om een ontwikkeling als gevolg van de (veranderende) onderwijscontext, of om een combinatie van beide factoren. Het belang van de leeromgeving voor de manier waarop studenten hun studie aanpakken, komt veelvuldig uit onderzoek naar voren (zie bijv. Wierstra & Beerends, 1996). Zo is het plausibel te veronderstellen dat de instructie in het hoger onderwijs tot een andere manier van leren uitnodigt vergeleken bij het voortgezet onderwijs, en daarmee ook tot een ander verband tussen concepties en strategieën. Hiermee is echter nog niet gezegd dat het aspect van de persoonsgebonden ontwikkeling er niet toe doet. Zowel Entwistle (1981) als Busato (1998) vinden verschillende verbanden tussen persoonsgebonden factoren en de manier waarop studenten leren. Bij onze analyse van de vraag naar de stabiliteit in leren besteden we daarom tevens aandacht aan de mogelijke effecten van enerzijds een meer algemene persoonsgebonden ontwikkeling en anderzijds de onderwijscontext.

Samenvattend staat in het huidige artikel de vraag naar stabiliteit met betrekking tot leerconcepties en verwerkingsstrategieën centraal. Eerst wordt een beeld geschetst van de eventuele ontwikkelingen in respectievelijk leerconcepties en verwerkingsstrategieën door middel van het vergelijken van een groep leerlingen in het vierde jaar van het voortgezet onderwijs, een groep leerlingen in het zesde jaar en een groep studenten in het hoger onderwijs (eerste interpretatie van stabiliteit). Vervolgens worden de verbanden tussen leerconcepties en verwerkingsstrategieën nader onderzocht in een vergelijking van deze drie groepen (tweede interpretatie van stabiliteit). Als indicaties voor de mogelijke effecten van 'algemene persoonsgebonden ontwikkeling' en 'onderwijscontext' op de geconstateerde ontwikkelingen in leren, is gebruik gemaakt van respectievelijk de variabele 'leeftijd' en de variabele 'onderwijsniveau'.

Het onderzoek is verricht in het volwassenenonderwijs. In dit type onderwijs is de variatie in leeftijd groot: de onderzoeksgroep bevat 16-jarigen, maar ook 'leerlingen' van 71 jaar. Deze variatie biedt de mogelijkheid, binnen één onderwijsniveau, verschillende leeftijds-groepen onderling te vergelijken. Daartoe wordt eerst gekeken naar stabiliteit in termen van de eerste interpretatie: is het zo dat oudere lerenden binnen één en hetzelfde niveau andere strategieën gebruiken en er andere leerconcepties op na houden vergeleken bij de jongere lerenden? Daarnaast wordt het verband tussen concepties en strategieën onderzocht (tweede interpretatie van stabiliteit). Indien dit verband sterker wordt naarmate de lerenden ouder zijn,

kan worden geconcludeerd dat leeftijd een deel van het verband verklaart. Om uitspraken te kunnen doen over de invloed van onderwijsniveau is vervolgens, met gebruikmaking van de onderzoeksresultaten van Vermunt (1992), een vergelijking gemaakt tussen drie onderwijsniveaus, binnen één leeftijdsgroep. Ook hier wordt gekeken in hoeverre de onderwijsniveaus verschillen voor wat betreft hun gemiddelde scores op concepties en strategieën en in hoeverre ze verschillen voor wat betreft de onderlinge verbanden.

Vermunt (1992) onderscheidt vier leerstijlen, die elk opgebouwd zijn uit een specifieke combinatie van verwerkings- en regulatiestrategieën, leerconcepties en leeroriëntaties. In de inleiding op dit themanummer (ten Dam, Severiens & Vermunt, 2000) is beschreven welke verwerkingsstrategieën en leerconcepties worden verbonden in de vier leerstijlen. In het kort betreft het de volgende verbanden: opbouw van kennis en de diepe verwerkingsstrategieën, opname van kennis en de stapsgewijze verwerkingsstrategieën en tot slot gebruiken van kennis en de concrete verwerkingsstrategieën.

Samenvattend wordt in dit artikel een antwoord gezocht op de volgende onderzoeksvragen:

1. In welke mate zijn leerconcepties en verwerkingsstrategieën gerelateerd aan onderwijsniveau, onafhankelijk van leeftijd?
2. In welke mate zijn leerconcepties en verwerkingsstrategieën gerelateerd aan leeftijd, onafhankelijk van onderwijsniveau?
3. In welke mate verschillen de verbanden tussen leerconcepties en verwerkingsstrategieën als we verschillende onderwijsniveaus vergelijken, onafhankelijk van leeftijd?
4. In welke mate verschillen de verbanden tussen leerconcepties en verwerkingsstrategieën als we verschillende leeftijden vergelijken, onafhankelijk van onderwijsniveau?

Op basis van reeds verricht onderzoek spreken we de volgende verwachtingen uit. Allereerst verwachten we een algemene toename van het strategiegebruik van de groep leerlingen in het vierde jaar van het voortgezet onderwijs naar de groep studenten in het hoger onderwijs (zie Wigfield *et al.*, 1996). Wat de leerconcepties betreft, verwachten we een ontwikkeling te zien van de leerconceptie 'opname van kennis' naar de leerconceptie 'opbouw van kennis': studenten in het hoger onderwijs zullen gemiddeld hoger scoren op de leerconceptie 'opbouw van kennis' dan leerlingen in het voortgezet onderwijs, waarbij de zesde klas weer gemiddeld hoger zal scoren dan de vierde klas (zie Schommer, 1993). Tenslotte verwachten we dat het verband tussen leerconcepties en leerstrategieën sterker is bij studenten in het hoger onderwijs dan bij leerlingen in de zesde klas van het voortgezet onderwijs, dat op zijn beurt weer sterker zal zijn dan bij leerlingen in de vierde klas. Over het relatieve effect van leeftijd versus onderwijsniveau op de gevonden ontwikkelingen spreken we geen verwachting uit.

2. METHODE

Respondenten

Vijf scholen voor volwassenenonderwijs in de randstad participeerden in het onderzoek: 432 lerenden uit vier en vijf havo en vier, vijf en zes vwo. Omdat vrijwel alle lerenden het vak Nederlands in hun pakket hebben, is de ILS ingevuld tijdens de lessen Nederlands door alle aanwezige lerenden.² De gemiddelde leeftijd van de respondenten is 33, varieert van 16 tot 71 jaar, en 75% van de steekproef is jonger dan 22.³

Om uitspraken te kunnen doen over de stabiliteit van leerconcepties en leerstrategieën is een cross-sectioneel onderzoek verricht. In totaal zijn de gegevens van drie groepen lerenden gebruikt. De eerste twee groepen betreffen lerenden uit vier havo/vwo, lerenden uit zes vwo en eerstejaars universitaire studenten. Naast de respondenten uit eigen onderzoek is gebruik gemaakt van data uit het onderzoek van Vermunt (1992, p. 115 en p. 117).

Analyses

Met betrekking tot de gegevens uit het onderzoek van Vermunt hadden we uitsluitend de beschikking over gemiddeldes, standaarddeviaties en factorladingen. Het gegevensbestand uit het onderzoek in het volwassenenonderwijs bestond uit de volledige dataset. Dit betekent een beperking in het arsenaal aan analysemethoden met betrekking tot twee van de vier onderzoeksvragen (1 en 3). Daar gaat het immers om het effect van onderwijsniveau en dient dus gebruik gemaakt te worden van de gegevens van Vermunt.

Om de effecten van onderwijsniveau te berekenen (vraag 1 en 3), is leeftijd constant gehouden. Dit is gedaan door een selectie te maken uit de groep volwassenenonderwijs-respondenten: alleen die volwassenenonderwijs respondenten die jonger zijn dan 26 jaar zijn vergeleken met Vermunt's eerstejaars universitaire studenten. Met betrekking tot de eerste vraag zijn de gemiddeldes en standaarddeviaties gebruikt om verschillen (Cohen, 1977) te berekenen. Verschilcores kunnen in deze situatie een betrouwbaar beeld van de ontwikkeling geven, mits de betrouwbaarheid van de ILS-schalen voldoende is. In tabel 1 staan deze betrouwbaarheden vermeld, de alpha's zijn hoog genoeg om de verschillencores als betrouwbaar te kunnen beschouwen.

Tabel 1: Betrouwbaarheden (Cronbach's alpha's) van de gebruikte ILS-schalen in het huidige onderzoek

	Volwassenenonderwijs gegevens	Vermunt gegevens
diepte verwerking	.81	.85
stapsgewijze verwerking	.79	.78
concrete verwerking	.71	.71
opbouw van kennis	.87	.78
opname van kennis	.89	.77
gebruik van kennis	.87	.70

De verschillencores zijn geïnterpreteerd aan de hand van Cohen's aanwijzingen: .20 wijst op een klein verschil, .50 op een redelijk verschil en .80 wordt geïnterpreteerd als een groot verschil.⁴

Met betrekking tot de derde onderzoeksvraag, waarin het gaat om verbanden tussen leerconcepties en verwerkingsstrategieën, is een vergelijking gemaakt tussen factorladingen. De factorladingen zoals Vermunt ze publiceerde (1992, p. 117) zijn vergeleken met de factorladingen zoals ze gevonden zijn in twee afzonderlijke factoranalyses verricht op de groep (jonger dan 26) uit vier havo/vwo en zes vwo.

De tweede en vierde onderzoeksvragen betreffen het effect van leeftijd. Om dit effect te berekenen, is onderwijsniveau constant gehouden door alleen de respondenten uit het volwassenenonderwijs mee te nemen in de analyses. Omdat we van deze groep respondenten wel de beschikking hadden over de volledige dataset is hier gekozen voor regressie-analyses. Met betrekking tot de tweede onderzoeksvraag zijn de scores op de verschillende schalen als afhankelijk variabele geanalyseerd, en leeftijd en onderwijsniveau als onafhankelijke variabelen. Dat onderwijsniveau ook hier als variabele meegenomen wordt, heeft te maken met het feit dat ook binnen deze groep sprake kan zijn van mogelijke niveauverschillen (vier havo/vwo versus zes vwo). Omdat we echter in vraag 3 niet geïnteresseerd zijn in onderwijsniveau, is deze variabele als eerste in de regressievergelijking opgenomen, en vervolgens leeftijd. Hier toe is een stapsgewijze regressie-analyse verricht. Eventuele effecten van leeftijd kunnen daarmee geïnterpreteerd worden als onafhankelijk van onderwijsniveau.

De vierde onderzoeksvraag is beantwoord door een regressievergelijking te berekenen waarin de verbanden tussen leerconcepties en leerstrategieën als afhankelijke variabelen zijn

opgenomen. Evenals bij de tweede vraag is onderwijsniveau (vier havo/vwo versus zes vwo) als eerste variabele in de vergelijking opgenomen en vervolgens de variabele leeftijd.

3. RESULTATEN

3.1 Stabiliteit in termen van gemiddelde veranderingen

In deze paragraaf komt stabiliteit aan de orde in termen van de gemiddelde scores van respectievelijk leerconcepties en leerstrategieën (eerste interpretatie van stabiliteit). Allereerst worden de mogelijke effecten van het onderwijsniveau onderzocht. In tabel 2 zijn de gemiddelde scores en standaarddeviaties vermeld van de 'jongere' respondenten uit het volwassenenonderwijs en de respondenten uit het hogere onderwijs afkomstig uit Vermunt's onderzoek.

Tabel 2: Gemiddelde schaalscores, bijbehorende standaarddeviaties en verschillen van lerenden uit 4 havo/vwo, 6 vwo uit het volwassenenonderwijs (leeftijd < 26) en eerstejaars universitaire studenten (Vermunt, 1992, p. 11 en aanvullende informatie).

	4 havo/vwo N=73 M(s.d.)	Verschil- scores 4 havo/vwo en 6 vwo	6 vwo N=108 M(s.d.)	Verschil- scores 6 vwo en ho	ho N=795, p. 115 M(s.d.)
diepte verwerking	2.71(.71)	.08	2.65(.72)	.69	3.16(.76)
stapsgewijze verwerking	3.16(.70)	.37	2.90(.70)	.17	2.78(.68)
concrete verwerking	2.68(.86)	.09	2.60(.87)	.25	2.81(.80)
opbouw van kennis	3.25(.71)	.39	2.98(.67)	.86	3.52(.59)
opname van kennis	3.74(.59)	.18	3.63(.64)	.15	3.53(.65)
gebruik van kennis	3.45(.78)	0	3.45(.77)	.69	3.91(.57)

We bespreken eerst de veranderingen met betrekking tot verwerkingsstrategieën en vervolgens die met betrekking tot leerconcepties. De diepe verwerkingsstrategieën worden ongeveer even vaak door vierde- als door zesdeklassers gebruikt (d-score is .08). In het hoger onderwijs neemt het gebruik van dit type strategieën echter duidelijk toe (d-score is .69). Daarnaast lijken zesdeklassers minder vaak gebruik te maken van stapsgewijze verwerkingsstrategieën dan de vierdeklassers (d-score is .37). In het hoger onderwijs treden er ten opzichte van de zesde klassers geen veranderingen meer op (d-score is .17). Het gebruik van de concrete verwerkingsstrategieën is ongeveer gelijk in de vierde en de zesde klas (d-score is .09), om vervolgens in het hoger onderwijs iets toe te nemen (d-score is .25). Samenvattend lijkt het erop dat de overgang van het voortgezet onderwijs naar het hoger onderwijs met name invloed heeft op het (toenemende) gebruik van diepe verwerkingsstrategieën.

De ontwikkeling in de leerconceptie 'opbouw van kennis' vertoont het volgende patroon. De leerlingen uit de vierde klas zien leren vaker als het opbouwen van kennis dan de leerlingen uit de zesde klas (d-score is .39). Vervolgens keert deze trend zich echter om: in het hoger onderwijs wordt leren juist weer vaker gezien als opbouwen van kennis (d-score is .86). Met betrekking tot de leerconceptie 'opname van kennis' zien we nauwelijks verschil tussen de vierde en de zesde klas (d-score is .18). Evenmin zijn er verschillen tussen de zesdeklassers en de studenten in het hoger onderwijs (d-score is .15). Tot slot blijven de scores van lerenden ten aanzien van de leerconceptie 'gebruik van kennis' tijdens het voortgezet onderwijs gelijk. In vergelijking met leerlingen van het voortgezet onderwijs zijn studenten in het hoger onderwijs daarentegen vaker van mening dat het bij leren om het gebruiken van kennis gaat (d-score is .69)⁵.

Om een indruk te krijgen van de mogelijke verschillen in stabiliteit, zijn de gemiddelde d-scores van leerconcepties en leerstrategieën berekend. Deze bedroegen respectievelijk .38 en .28. Het verschil tussen beide d-scores lijkt te klein om te kunnen concluderen dat er sprake is van meer stabiliteit in het ene dan wel het andere concept. De grootte van de veranderingen die geconstateerd zijn, binnen één en dezelfde leeftijdsgroep, lijkt de conclusie te rechtvaardigen dat onderwijsniveau in ieder geval een deel van de veranderingen kan verklaren. De vraag blijft of leeftijd, onafhankelijk van onderwijsniveau, ook gerelateerd is aan de manier waarop studenten leren. Om een antwoord op deze vraag te geven, zijn regressie-analyses verricht.

Tabel 3: Stapsgewijze regressie-analyses: betagewichten en bijbehorende t en p waarden van de onafhankelijke variabelen 'onderwijsniveau' en 'leeftijd' in de groep lerenden uit het volwassenenonderwijs, alle leeftijden.

afhankelijke variabelen	betagewicht van onderwijsniveau	t (p)	betagewicht van leeftijd	t
diepte verwerking	-.01	-.17 (p<.87)	.18	2.51 (p<.01)
stapsgewijze verwerking	-.18	-2.52 (p<.01)	.05	.69 (p<.49)
concrete verwerking	-.06	-.81 (p<.42)	.08	1.10 (p<.27)
opbouw van kennis	-.16	-2.24 (p<.03)	.22	3.17 (p<.01)
opname van kennis	-.12	-1.69 (p<.10)	-.02	-.25 (p<.80)
gebruik van kennis	-.02	-.29 (p<.77)	.08	1.06 (p<.29)

Tabel 3 laat zien dat leeftijd in twee gevallen gerelateerd is aan leerconcepties en leerstrategieën: het gaat om de leerconceptie 'opbouw van kennis' en om de diepe verwerkingsstrategieën. In beide gevallen laat leeftijd een positief effect zien. Blijkbaar definiëren ouderen leren vaker in termen van het opbouwen van kennis, en zetten ze ook vaker de diepe verwerkingsstrategieën in, dan jongeren. Deze effecten bestaan los van het onderwijsniveau van de respondenten. De overige verwerkingsstrategieën en leerconcepties zijn niet gerelateerd aan leeftijd.

3.2 Stabiliteit in termen van verbanden tussen leerconcepties en strategieën

In deze subparagraaf bespreken we de resultaten met betrekking tot de verbanden tussen leerconcepties en verwerkingsstrategieën (tweede interpretatie van stabiliteit). Eerst wordt een vergelijking gemaakt tussen de drie onderscheiden onderwijsniveaus en vervolgens worden de mogelijke leeftijdseffecten gepresenteerd.

De vergelijking tussen respectievelijk onderwijsniveaus en leeftijd wordt gemaakt met behulp van factoranalyses. Evenals in Vermunt's onderzoek, zijn in het volwassenenonderwijsonderzoek alle 20 schalen uit de ILS gebruikt in een vier-factoren analyse met varimax rotatie. Wat opvalt in tabel 4 is dat in de eerste factor (de betekenisgerichte leerstijl) het verband tussen leerconcepties en leerstrategieën in de groepen in het volwassenenonderwijs minder sterk is vergeleken bij eerstejaars universitaire studenten groep. Dit geldt voor het verband tussen opbouw van kennis enerzijds en de diepe verwerkingsstrategieën anderzijds. In de tweede factor (de reproductiegerichte leerstijl) zijn de ladingen vrijwel identiek. Voor deze leerstijl geldt dat in de hoger onderwijs-groep leerconcepties in dezelfde mate verbonden zijn met verwerkingsstrategieën als in de volwassenenonderwijs-groepen.⁶ De verschillen tussen het volwassenenonderwijs enerzijds en het hoger onderwijs anderszijds worden dus met name gevonden in de betekenisgerichte leerstijl.

Tabel 4: Factoranalyses (gebaseerd op 20 schalen en vier factoren met varimax rotatie): factorladingen van de eerste twee factoren van lerenden uit 4 havo/vwo, 6 vwo uit het volwassenenonderwijs (leeftijd < 26) en eerstejaars universitaire studenten (Vermunt, 1992, p. 11)

	F1 4 havo/vwo	F1 6 vwo	F1 ho	F2 4 havo/vwo	F2 6 vwo	F2 ho
Verwerkingsstrategieën						
Diepe verwerking						
Relateren en structureren	.78	.80	.73			
Kritisch verwerken	.78	.80	.71			
Stapsgewijze verwerking						
Memoriseren en herhalen				.79	.69	.72
Analyseren				.61	.76	.73
Concrete verwerking						
Leerconcepties						
Opbouw van kennis	.52	.46	.74			
Opname van kennis				.69	.61	.60
Gebruik van kennis						
Stimulerend onderwijs						
Samenwerken						
Percentage verklaarde variantie	22%	22%	21%	9%	17%	15%

Om het effect van leeftijd op de verbanden tussen leerconcepties en leerstrategieën te onderzoeken zijn twee regressie-analyses verricht. In de eerste regressievergelijking gold het verband tussen opname van kennis en de stapsgewijze strategieën als afhankelijke variabele. In de tweede vergelijking werd de afhankelijke variabele gevormd door het verband tussen opbouw van kennis en de diepe verwerkingsstrategieën. In tabel 5 staan de betagewichten en t's vermeld.

Tabel 5: Stapsgewijze regressie-analyses: betagewichten en bijbehorende t en p waarden van de onafhankelijke variabelen 'onderwijsniveau' en 'leeftijd' in de groep lerenden uit het volwassenenonderwijs, alle leeftijden.

afhankelijke variabelen	betagewicht van onderwijsniveau	t (p)	betagewicht van leeftijd	t (p)
diepte verwerking - opbouw van kennis	-.10	-1.39 (p<.17)	.22	3.20 (p<.01)
stapsgewijze verwerking - opname van kennis	-.18	-2.64 (p<.01)	.03	.44 (p<.66)

Uit de verrichte analyses blijkt dat het verband tussen diepte verwerking en opbouw van kennis gerelateerd is aan leeftijd. Correlaties laten zien dat de verbanden tussen enerzijds de leerconceptie 'opbouw van kennis' en anderzijds de diepe verwerkingsstrategie sterker worden naarmate de leeftijd van de respondenten stijgt. De verbanden tussen opname van kennis en de stapsgewijze strategieën zijn daarentegen niet gerelateerd aan leeftijd. Samenvattend kan

worden geconcludeerd dat de mate waarin lerenden zich in hun studiegedrag laten leiden door hun leerconcepties deels afhankelijk is van hun leeftijd.

4. CONCLUSIES

In dit artikel is een beeld geschetst van de stabiliteit van het leerstrategieën en leerconcepties van leerlingen en studenten in het voortgezet en hoger onderwijs. Stabiliteit is op twee manieren geïnterpreteerd en onderzocht: in termen van gemiddelde veranderingen in de afzonderlijke componenten en in termen van veranderingen in de onderlinge verbanden tussen de componenten. Beide interpretaties van stabiliteit vormden ook het uitgangspunt voor een beschrijving van het reeds verrichte onderzoek naar ontwikkelingen in leren. Op basis van deze beschrijving, formuleerden we in de inleiding de volgende verwachtingen: 1) het gebruik van verwerkingsstrategieën laat een algemene toename zien; 2) de leerconceptie 'opname van kennis' laat een afname zien en de leerconceptie 'opbouw van kennis' een toename; en 3) het verband tussen leerconcepties en leerstrategieën wordt sterker.

Het door ons verrichte onderzoek had tot doel een meer gedetailleerd beeld te schetsen van ontwikkelingen in leren van het voortgezet onderwijs naar het hoger onderwijs. Omdat mogelijke verschillen tussen leerlingen/studenten in beide onderwijstypen zowel in termen van leeftijd als in termen van onderwijsniveau geïnterpreteerd kunnen worden, zijn in de huidige studie deze factoren onafhankelijk van elkaar onderzocht. Als onderzoekscontext is hiervoor het volwassenenonderwijs gebruikt. Door de grote leeftijdsvariatie leent dit type onderwijs zich goed voor een vergelijking van verschillende leeftijden binnen één onderwijsniveau. Het wordt daardoor mogelijk na te gaan in hoeverre leeftijd bepalend is voor de sterkte van het verband tussen leerconcepties en leerstrategieën, los van het onderwijsniveau. Omdat de jongere lerenden in het volwassenenonderwijs qua leeftijd vergelijkbaar zijn met studenten in het hoger onderwijs, is het tevens mogelijk een vergelijking te maken tussen onderwijsniveaus, los van leeftijd. Voor deze laatste vergelijking is mede gebruik gemaakt van de eerstejaars universitaire studenten uit het onderzoek van Vermunt (1992).

De eerste onderzoeksvraag naar de relatie tussen het onderwijsniveau en de gemiddelde scores op de verwerkingsstrategieën en leerconcepties (stabiliteit in de eerste interpretatie) kan als volgt beantwoord worden. De drie onderscheiden typen verwerkingsstrategieën vertonen alle duidelijke veranderingen in de tijd. Als het gaat om het gebruik van de diepe verwerkingsstrategieën en de concrete verwerkingsstrategieën zien we vooral veranderingen tussen het voortgezet onderwijs en het hoger onderwijs: studenten in het hoger onderwijs maken meer gebruik van beide strategieën dan leerlingen in de zesde klas van het voortgezet onderwijs. De stapsgewijze strategieën daarentegen worden al tijdens het voortgezet onderwijs minder vaak ingezet; de gemiddelde score in zes vwo is ongeveer gelijk aan de score in het hoger onderwijs. Omdat de diepe en concrete verwerkingsstrategieën over het algemeen gezien worden als strategieën die leiden tot betere studieresultaten, merken we deze ontwikkelingen als positief aan. Samenvattend concluderen we dat onderwijsniveau, los van leeftijd, een effect heeft op het strategiegebruik van studenten. Hetzelfde geldt voor de leerconcepties. Terwijl de leerconceptie 'opname van kennis' nauwelijks verandert, laat de leerconceptie 'opbouw van kennis' wel veranderingen zien. Het gaat hier om een niet-lineair patroon: leerlingen in de vierde klas zien leren vaker in termen van het opbouwen van kennis dan leerlingen uit de zesde klas, maar in het hoger onderwijs wordt leren juist weer wel vaker gezien als het opbouwen van kennis. Een mogelijke interpretatie voor deze 'tijdelijke afname' vormt de specifieke indexamensituatie van de zes vwo-ers waarin over het algemeen weinig ruimte is voor het opbouwen van kennis.

De tweede onderzoeksvraag had betrekking op de mogelijke effecten van leeftijd, los van onderwijsniveau. Ook hier gaat het om de vraag naar stabiliteit in de eerste interpretatie. Uit de resultaten blijkt dat alleen de diepe verwerkingsstrategieën en de leerconceptie 'opbouw

van kennis' gerelateerd zijn aan leeftijd. De andere componenten zijn minder gevoelig voor leeftijd. Zowel de diepe leerstrategieën als de leerconceptie 'opbouw van kennis' maken dus, los van onderwijsniveau, een ontwikkeling door.

De vraag is of de geobserveerde veranderingen in leerstrategieën geïnterpreteerd kunnen worden in termen van een toename van het repertoire aan strategieën (Wigfield *et al.*, 1996). Terwijl studenten in het hoger onderwijs meer gebruik maken van de diepe en concrete verwerkingsstrategieën dan leerlingen in het voortgezet onderwijs, laten de onderzoeksresultaten daarentegen zien dat de stapsgewijze strategieën in de loop van de onderwijsloopbaan juist minder vaak ingezet worden. Dit hoeft evenwel niet te maken te hebben met het beschikbare repertoire van lerenden, maar kan ook samenhangen met - de eisen van - de onderwijsomgeving. Het feit dat we geen relatie hebben gevonden tussen stapsgewijze strategieën en leeftijd zou er op kunnen duiden dat de ontwikkeling van deze strategieën in een vroeger stadium plaatsvindt dan de vierde klas van het voortgezet onderwijs. We achten het, met inbegrip van deze opmerkingen, plausibel om te concluderen dat lerenden in de loop van hun schoolloopbaan een steeds groter repertoire aan strategieën tot hun beschikking hebben. De daarbij gemaakte veronderstelling dat jongere lerenden zich eerst de stapsgewijze strategieën eigen maken, en vervolgens de diepe en concrete strategieën, zou nader onderzocht dienen te worden in een longitudinaal onderzoek dat het gehele voortgezet onderwijs omvat.

De geformuleerde verwachtingen met betrekking tot de leerconcepties zijn ten dele uitgekomen. De verwachte afname van de leerconceptie 'opname van kennis' vond niet plaats. De verwachte toename van de leerconceptie 'opbouw van kennis' gedurende het voortgezet onderwijs vond niet plaats, maar er was wel een duidelijke toename van deze conceptie in het hoger onderwijs. De door ons getrokken parallel tussen de ontwikkeling in epistemologische opvattingen en leerconcepties lijkt voor een deel gerechtvaardigd te zijn. Het onderscheiden van kennis- en leerconcepties verdient echter veel meer onderzoeks aandacht: Wat betekent vervlochtenheid van deze twee begrippen precies, op empirisch niveau? En welke consequenties heeft dat voor verbanden met leerstrategieën?

De derde en vierde onderzoeksvraag hadden betrekking op verbanden tussen leerstrategieën en leerconcepties (stabiliteit in de tweede interpretatie). De ontwikkelingshypothese van Roosendaal en Vermunt (1996) lijkt in het huidige onderzoek alleen op te gaan voor de verbanden tussen de diepe verwerkingsstrategieën en de leerconceptie 'opbouw van kennis'. Het verband tussen deze twee specifieke componenten van leerstijlen wordt inderdaad sterker naarmate het onderwijsniveau stijgt, en eveneens naarmate de leeftijd van de lerenden stijgt (onderzoeksvraag 4). De geformuleerde ontwikkelingshypothese gaat daarentegen niet op voor de verbanden tussen de stapsgewijze verwerkingsstrategieën en de leerconceptie 'opname van kennis'. De verbanden tussen deze twee componenten zijn even sterk in het volwassenenonderwijs als in het hoger onderwijs. Ook is het gevonden verband niet gerelateerd aan leeftijd, los van onderwijsniveau. We concluderen op basis van deze gegevens dat de verbanden tussen strategieën en concepties binnen de betekenisgerichte leerstijl (diepe verwerking en opbouw van kennis) zich wel ontwikkelen van het voortgezet naar het hoger onderwijs en in de leeftijd boven achttien jaar. De verbanden in de reproductiegerichte leerstijl (stapsgewijze verwerking en opname van kennis) ontwikkelen zich daarentegen niet meer na het achttiende levensjaar of gedurende het volwassenenonderwijs en hoger onderwijs. Hierboven gaven we reeds aan, dat lerenden zich eerst de stapsgewijze strategieën eigen lijken te maken, alvorens ze leren de diepe en concrete strategieën in te zetten. Een zelfde type ontwikkeling lijkt te gelden voor verbanden tussen leerconcepties en leerstrategieën: jongere lerenden verbinden in eerste instantie de 'opname van kennis' leerconceptie met de stapsgewijze strategieën. Pas in een latere leeftijdsfase en op een hoger onderwijsniveau worden de componenten in de betekenisgerichte leerstijl met elkaar verbonden.

Op basis van dit onderzoek kunnen we geen uitspraken doen over wat precies de factoren in het voortgezet onderwijs en het hoger onderwijs zijn die bijdragen aan de gevonden veranderingen in het strategiegebruik en in de manier waarop leerlingen en studenten leren defi-

niëren. Wel laten de resultaten zien dat, behalve leeftijd, in ieder geval ook het onderwijsniveau een factor van betekenis is. Deze constatering geeft richting aan het zoeken naar mogelijk verklarende factoren voor de leerconcepties van lerenden en voor het op een gegeven moment inzetten van specifieke leerstrategieën. Wat de gevonden effecten van leeftijd betreft, is het belangrijk om te realiseren dat deze een indicatie kunnen vormen voor een onderliggende variabele die de verbanden tussen leerstrategieën en leerconcepties nader kan verklaren. Een voorbeeld van zo'n variabele is 'onderwijservaringen'. Lerenden in het volwassenonderwijs verschillen nogal op dit punt: sommigen respondenten genoten bijvoorbeeld nog onderwijs op de HBS of de MMS. 'Werkervaring' is eveneens een variabele die mogelijk differentieert tussen de drie onderscheiden groepen respondenten.

Bij wijze van afsluiting wijzen we er op dat de conclusies uit dit onderzoek gebaseerd zijn op een cross-sectioneel onderzoek. We hebben de aanname gedaan dat het in principe gaat om longitudinale ontwikkelingen, maar dit hoeft niet persé het geval te zijn. Zo zien we in Busato *et al.* (1998) verschillende resultaten in het cross-sectionele en het longitudinale onderzoek. Schommer *et al.* (1997) vinden daarentegen wel dezelfde resultaten in hun cross-sectionele en longitudinale onderzoek. Het verdient kortom aanbeveling de resultaten uit het huidige onderzoek te controleren in een longitudinale onderzoeksofzet. Hiermee kan worden nagegaan of andere verschillen tussen lerenden in het hoger onderwijs versus het volwassenonderwijs, los van de invloed van de onderwijscontext of de leeftijd van de betrokkenen, een verklaring vormen voor de gevonden veranderingen in leerconcepties en leerstrategieën.

REFERENTIES

- Baxter Magolda, M.B. (1992). *Knowing and reasoning in college. Gender-related patterns in students' intellectual development*. San Francisco: Jossey-Bass Publishers.
- Boekaerts, M., Otten, R., & Simons, R.J. (1997). Een onderzoek naar de bruikbaarheid van de ILS. Leerstijlen in de onderbouw van het voortgezet onderwijs. *Tijdschrift voor Onderwijsresearch*, 22, 15-36.
- Busato, V. (1998). *Leerstijlen nader geanalyseerd*. Universiteit van Amsterdam, Academisch proefschrift.
- Busato, V., Prins, F., Elshout, J. & Hamaker, C. (1998). Learning styles: A cross-sectional and longitudinal study in higher education. *British Journal of Educational Psychology*, 68, 427-441.
- Curry, L. (1983). *An organization of learning styles theory and constructs*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada. (ERIC No. ED 235 185).
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences (Rev. Ed.)*. New York: Academic Press.
- Dippelhofer-Stiem, B. (1989). The development of research-oriented learning in five European countries. *European Journal of Psychology of Education*, 4, 489-503.
- Entwistle, N.J. (1981). *Styles of learning and teaching*. New York: John Wiley & Sons.
- Hofer, B.K. & Pintrich, P.R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88-140.
- Klatter, E. (1995). Leerstijlen in de brugklas - onderzoek naar een vakspecifieke leerstijl. In H.C. Schouwenburg & J.T. Groenewoud (Red.), *Studievaardigheid en leerstijlen*. Proceedings van de 14e Landelijke Dag Studievaardigheden te Groningen (pp.169-191). Groningen: WoltersNoordhoff.
- Lonka, K. & Lindblom-Ylänne, S. (1996). Epistemologies, conceptions of learning, and study practices in medicine and psychology. *Higher education*, 31(1), 5-24.
- Marton, F., Dall'Alba, G., & Beaty, E. (1993). Conceptions of learning. *International Journal of Educational Research*, 19, 277-300.
- 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-Outcomes as a function of the learners conception of the task. *British Journal of Educational Psychology*, 46, 115-127.
- Messick, S. (1996). Cognitive styles and learning. In E. De Corte & F. Weinert, (Eds.), *International encyclopedia of developmental and instructional psychology* (pp. 638-641). Oxford: Elsevier Science Ltd.
- Perry, W.G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart & Winston.

- Pinto, J.K. & Geiger, M.A. (1991). Changes in learning-style preferences: a prefatory report of longitudinal findings. *Psychological Reports*, 68, 195-201.
- Prins, F., Busato, V., Hamaker, C., & Visser, K. (1996). Een bijdrage tot de validatie van het (meta)cognitieve deel van de Inventaris Leerstijlen. *Pedagogische Studiën*, 73, 108-122.
- Purdie, N., Hattie, J. & Douglas, G. (1996). Student conceptions of learning and their use of self-regulated learning strategies: A cross-cultural comparison. *The Journal of Educational Psychology*, 88(1), 87-100.
- Roosendaal, A. & Vermunt, J.D.H.M. (1996). Leerstijlen en zelfstandig leren in het voorportaal van het studiehuis. *Tijdschrift voor Onderwijsresearch*, 21(4), 336-347.
- Schommer, M. (1993). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology*, 85, 498-504.
- Schommer, M., Calvert, C., Garigliette, G. & Baja, A. (1997). The development of epistemological beliefs among secondary students: A longitudinal study. *Journal of Educational Psychology*, 89(1), 37-40.
- Severiens, S.E. (1997). Gender and Learning. Learning styles, ways of knowing and patterns of reasoning. Amsterdam: PhD thesis.
- Severiens, S.E. & Dam, G.T.M. ten (1996). Verschillen in leerstijlen. *Tijdschrift voor Onderwijsresearch*, 21, 294-305.
- Severiens, S.E. & Dam, G.T.M. ten (1998). Gender and learning: Comparing two theories. *Higher Education*, 35, 329-350.
- Schmeck, R.R. (1983). Learning styles of college students. In: R. Dillon & R.R. Schmeck (eds.), *Individual differences in cognition* (Vol. 1 pp. 233-79). New York: Academic Press.
- Schouwenaar, H.C. (1996). Een onderzoek naar leerstijlen. *Tijdschrift voor Onderwijsresearch*, 21, 151-161.
- 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. Schouwenaar & J.T. Groenewoud (Red.), *Studievaardigheid en leerstijlen*. Proceedings van de 14e Landelijke Dag Studievaardigheden te Groningen (pp. 5172). Groningen: WoltersNoordhoff.
- Vermunt, J.D.H.M. (1998). The regulation of constructive learning processes. *British Journal of Psychology of Education*, 68, 149-171.
- Vermunt, J.D.H.M. & Sanden, J. van der (1996). Inleiding op het thema leerstijlen. *Tijdschrift voor Onderwijsresearch*, 21 (4), 291-293.
- Wigfield, A., Pintrich, P. & Eccles, J.S. (1996). Development between the ages of 11 and 25. In C. Berliner & R. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 148-185). New York: Simon & Schuster McMillan.
- Wierstra, R. & Beerends, E. (1996). Leeromgevingspercepties en leerstrategieën van eerstejaars studenten sociale wetenschappen. *Tijdschrift voor Onderwijsresearch*, 21(4), 306-322.

NOTEN

- Deze twee interpretaties van stabiliteit verwijzen naar 'tijd': het gaat om mogelijke veranderingen in de tijd. Het is echter ook mogelijk om stabiliteit in termen van context op te vatten: is het zo dat de manier waarop studenten leren in elke context dezelfde is? Of passen studenten hun leergedrag aan, aan de leercontext waarin ze zich bevinden. In het huidige onderzoek staat stabiliteit in termen van tijd centraal.
- In een enkel geval zijn de vragenlijsten thuis ingevuld, posthoc vergelijking wees uit dat het hier geen 'speciale' groep lerenden betrof.
- In het volwassenenonderwijs zijn globaal genomen drie groepen lerenden te onderscheiden (Severiens, 1997). De grootste groep bestaat uit jong-volwassenen (rond 19 jaar) die gestopt zijn in het reguliere onderwijs omdat ze bijvoorbeeld twee keer zijn blijven zitten, het examen niet gehaald hebben enzovoorts. Deze zogenaamde 'drop-outs' trachten in het volwassenenonderwijs alsnog hun diploma te halen. De tweede en kleinste groep in het volwassenenonderwijs is rond de 30 jaar. Hun motivatie heeft voornamelijk te maken met het vergroten van carrièrekansen. De derde groep bestaat voornamelijk uit vrouwen van 40 en ouder; in feite is dit de oorspronkelijke doelgroep van het volwassenenonderwijs. Deze groep kiest meestal voor die vakken die voor hen interessant of uitdagend zijn.
- Omdat verandering in de componenten impliceert dat er op het niveau van 'leerstijlen' ook veranderingen in termen van gemiddelden optreden, onderzoeken we in het huidige artikel alleen de veranderingen op het niveau van de componenten.

5. Met behulp van de formule $t = (d \cdot \sqrt{df}) / 2$ zijn de d-scores omgezet naar t's om de statistische significanties te bekijken. De verschillen tussen 4 en 6 vwo zijn significant ($p < .05$) als het gaat om stapsgewijze verwerking (de d-score is .37) en om de opbouw van kennis (de d-score is .39). Alle verschillen tussen 6 vwo en ho zijn statistisch significant. Dit laatste is mede te danken aan de grote N in het Vermunt's onderzoek.

6. De mogelijke verbanden tussen de leerconceptie 'gebruiken van kennis' en de concrete verwerkingsstrategieën worden niet besproken, omdat deze in het volwassenenonderwijs niet met elkaar correleren (voor details verwijzen we naar Severiens, 1997).

De samenhang tussen leerconcepties en verwerkingsstrategieën bij leerlingen in het middelbaar beroepsonderwijs

Angelique Slaats, Johan van der Sanden, Hans Lodewijks
Katholieke Universiteit Brabant

ABSTRACT

The study reported in this article focuses on learning strategies and -conceptions of students in vocational education. Since this type of education has received much less attention among researchers than higher education, it was necessary to conduct an interview study to gather information on learning strategies, learning conceptions and motivational orientations that are characteristic for students in vocational education. Analyses of the resulting protocols was aimed at detecting the relation between conceptions and strategies, separately for theoretical and practical courses. Results indicated an expected relationship within the theoretical courses, but no relationship within the practical courses. The discussion offers possible explanations for these results and suggestions for future research.

1. INLEIDING

Het beroepsonderwijs in Nederland omvat zowel een theoretische als een praktische beroepsvoorbereiding (WEB, 1996). De onderwijsprogramma's binnen het beroepsonderwijs kennen daarom zowel theoretische als praktische componenten, die vaak duidelijk gescheiden zijn van elkaar wat betreft tijd, locatie en lesmateriaal. De benamingen 'theorieles' en 'praktijkles' zijn dan ook gangbaar onder leerlingen, vaak geassocieerd met respectievelijk 'denken' en 'doen'. Theorielessen kenmerken zich vaak door een nadruk op het verwerven van cognitieve kennis en vaardigheden, waarvoor veelal traditionele vormen van instructie worden toegepast en de leerstof geordend is in vakken en lessen (Onstenk, 1998). Traditionele frontale en klassikale instructievormen bieden een docent de mogelijkheid om een grote mate van sturing aan te brengen (Engeström, 1994) en zetten doorgaans de leerling slechts in geringe mate aan tot actief leren (Mulder, 1997). De aangeboden leerstof is in het algemeen abstract en context-arm van aard en behoeft vertaling naar de praktijk. De kans is groot dat de leerling door deze vorm van theorieonderwijs kennis vergaart die inert is en daardoor onbruikbaar blijft (Boaler, 1998; Renkl, Mandl & Gruber, 1998).

Bij praktijklessen ligt de nadruk minder op het verwerven van cognitieve kennis en meer op het verwerven van beroeps- en situatiespecifieke vaardigheden (Resnick, 1987). In vergelijking met de traditionele theorieles, die op veel scholen nog altijd overheerst (Onstenk, 1998), is een praktijksituatie meer open en complex. Er is meer interactie tussen leerlingen, collega's en begeleiders, er wordt vaker gebruik gemaakt van hulpbronnen en werktuigen en er moet informatie worden verwerkt naar aanleiding van concrete gebeurtenissen en situaties. Binnen de school bestaan praktijkleersituaties vooral in de vorm van practica en praktijksimulaties, waarbij (een deel van) de beroepspraktijk wordt nagebootst (Raizen, 1994; Resnick, 1987). Vooral wanneer het maken van fouten een te groot risico -financieel of anderszins-

met zich meebrengt, vormen simulaties een veilige omgeving waarbinnen de benodigde vaardigheden kunnen worden opgedaan. Buitenschools vindt de beroepspraktijkvorming voornamelijk plaats in de oriënterende en participerende stages, waarbij het verwerven van functiegerichte, vakspecifieke kennis en vaardigheden centraal staat (De Vries, 1988). Naast het gestuurde, formele en intentionele leerproces dat in het theorieonderwijs aan bod komt, vinden binnen praktijkleersituaties ook informele en incidentele leerprocessen plaats (De Vries, 1988; Kwakman, 1993). Dit informele leren wordt vaak gezien als een toevallig leren, als nevenproces of onderdeel van (arbeids)handelingen (Rietbroek, 1998). Voorts kenmerken praktijkleersituaties, vooral de buitenschoolse, zich door participatie- en socialisatieprocessen, die invloed hebben op het zuiver vakinhoudelijke leerproces (Mulder, 1997; Nieuwenhuis, 1991).

Binnen het universitaire onderwijs blijft de praktijkcomponent meestal beperkt tot een afsluitende stage aan het einde van de opleiding. Gedurende het grootste gedeelte van de universitaire opleiding is er nauwelijks sprake van leersituaties die te vergelijken zijn met de praktijklessen, practica en stages uit het beroepsonderwijs.

Het leerstijlenmodel van Vermunt (1992; zie ook elders in dit themanummer) is gebaseerd op onderzoek onder studenten aan de universiteit. De respondenten die deelnamen aan de onderzoeken van Vermunt bevonden zich in de propedeutische of beginfase van hun opleiding, waarin geen praktijkcomponent is opgenomen. Vanwege verschillen in onderwijscontext en studentpopulatie, mag niet zonder meer aangenomen worden dat binnen het beroepsonderwijs dezelfde leerprocessen plaatsvinden als in het hoger onderwijs. Een eenvoudige (taalkundige) aanpassing van de Inventaris Leerstijlen voor het Hoger Onderwijs is om diezelfde reden ongewenst. Onderzoek naar leerstijlen binnen een relatief weinig onderzochte onderwijscontext als het beroepsonderwijs, vraagt in eerste instantie om diepgaande kwalitatieve studies die zich richten op een beschrijving van de processen en fenomenen die daar een rol spelen. Op basis van dergelijke studies kan dan een meer kwantitatieve onderzoeksbenadering gehanteerd worden, zodat het onderzoek op grotere schaal voortgang kan vinden.

In dit artikel wordt een deel van een onderzoeksproject beschreven waarin leeractiviteiten en competentie-ontwikkeling in beroepsgerichte praktijksituaties centraal staan (zie ook Van der Sanden en Van Bussel, 1995; Van der Sanden, 1997). Uitgaande van het leerstijlenmodel van Vermunt (1992) is door middel van gesprekken met studenten gezocht naar een invulling van de leerstijlonderdelen, toegespitst op het beroepsonderwijs. Voor de huidige onderzoeksvraag zijn enkel de leerconcepties en de verwerkingsstrategieën van belang.

Met betrekking tot leerconcepties is aangesloten bij Vermunts beschrijving hiervan als "...een persoonlijke theorie over leren en instructie" (Vermunt, 1992, p.16). Verondersteld wordt dat deze persoonlijke denkbeelden de interpretatie van instructiemaatregelen beïnvloeden en kleur en richting geven aan uitgevoerde leeractiviteiten (Van Rossum & Schenk, 1984). Vermunt gaat er verder vanuit dat een persoonlijke theorie over leren en instructie wordt opgebouwd als gevolg van ervaringen binnen het onderwijs en ervaringen met zichzelf als lerende, en niet contextspecifiek is. Verwerkingsstrategieën daarentegen zijn volgens hem wel contextafhankelijk en -specifiek. Ze zijn gericht op het bereiken van een leerdoel (zichzelf gesteld danwel extern opgelegd) en kunnen betrekking hebben op verschillende soorten van leerinhouden.

Men kan zich echter voorstellen dat het verwerven van praktische vaardigheden in praktijksituaties om andere activiteiten vraagt dan het verwerven van declaratieve kennis in schoolse leersituaties. De vraag is derhalve of leerconcepties, waarvan aangenomen mag worden dat ze voornamelijk betrekking hebben op formele en cognitief gerichte leersituaties, het leren in deze verschillende soorten onderwijscontexten in even grote mate beïnvloeden. De interviewstudie op basis waarvan de leerstijlen van lerenden in het beroepsonderwijs in kaart werden gebracht (zie Slaats, Lodewijks en Van der Sanden, 1994), voorzag dan ook in afzon-

derlijke vragen over het leren tijdens theorie- en praktijkleersituaties. Met de leerlingen werd gesproken over hun leergedrag tijdens het deelnemen aan zowel theorie- als praktijkonderwijs. De uitspraken van deze leerlingen zijn gecategoriseerd op basis van overeenkomsten en verschillen in dit leergedrag, zonder een verder onderscheid te maken naar de situaties waarin het desbetreffende leergedrag plaats had. Hoewel in de interviews dus zowel de theorie- als de praktijkcomponent nadrukkelijk aan bod is gekomen, zijn in dit stadium geen afzonderlijke categorieën gevormd voor deze twee onderwijscontexten.

De antwoorden die studenten gaven tijdens deze interviews, werden gebruikt om items te formuleren voor een leerstijlenvragenlijst, de ILS-MBO (Slaats, 1997).

In het leerstijlenmodel van Vermunt wordt er vanuit gegaan dat verwerkingsstrategieën worden beïnvloed door zowel leerconcepties als leeroriëntaties (Vermunt, 1992). Door Vermunt uitgevoerde regressieanalyses met leerconcepties en leeroriëntaties als voorspeller van verwerkingsstrategieën lieten echter zien dat de leeroriëntaties nauwelijks van invloed zijn op de verwerkingsstrategieën (Vermunt, 1992, p. 120-121). Daarentegen vond hij wel een duidelijke en logische samenhang tussen leerconcepties (door hem mentale modellen van leren genoemd) en verwerkingsstrategieën. Het gebruik van de dieptestrategieën relateren en kritisch verwerken blijkt samen te gaan met de opvatting dat kennis het resultaat is van de eigen constructieve leeractiviteiten van de lerende. Het gebruik van de stapsgewijze, meer oppervlakkige strategieën memoriseren en analyseren gaat gepaard met een reproductieve opvatting over leren; het gebruik van concrete verwerkingsstrategieën gaat samen met de opvatting dat kennis moet kunnen worden gebruikt bij het oplossen van problemen of het uitvoeren van praktische taken in de alledaagse werkelijkheid.

Ook Busato (1998, p. 65) vindt op basis van vragenlijstonderzoek onder psychologiestudenten sterke correlaties tussen leerconcepties en verwerkingsstrategieën en zwakke correlaties tussen leeroriëntaties en verwerkingsstrategieën. De gevonden samenhangen tussen leerconcepties en verwerkingsstrategieën komen overeen met de samenhangen die door Vermunt zijn beschreven. In een interview-studie werd door Van Rossum en Schenk (1984) vastgesteld dat studenten met reproductieve leerconcepties geneigd zijn de eerder door Marton en Säljö (1976) beschreven oppervlakte aanpak te hanteren. Studenten met constructieve leerconcepties, daarentegen, bleken een voorkeur te hebben voor een diepte-aanpak. Binnen het beroepsonderwijs is een vergelijkbare interviewstudie uitgevoerd door Eklund-Myrskög (1997). Zij vond dat de helft van de geïnterviewde leerlingen blijk geeft van een reproductieve leerconceptie die samenhangt met een voorkeur voor oppervlakkige verwerkingsstrategieën. De meer constructieve leerconcepties hangen samen met diepere verwerkingsstrategieën. Hoewel deze studie uitgevoerd werd onder leerlingen in het beroepsonderwijs (richting autotechniek), zijn de leerconcepties en verwerkingsstrategieën vastgesteld aan de hand van een schoolse leestaak.

De hierboven aangehaalde studies wijzen op een consistente samenhang tussen de aard van de leerconceptie van een lerende, en de door hem uitgevoerde of geprefereerde verwerkingsstrategieën. In alle studies hangen constructieve leerconcepties samen met diepe verwerkingsstrategieën en reproductieve leerconcepties met oppervlakkige verwerkingsstrategieën. Daar waar onder de respondenten ook een toepassingsgerichte leerconceptie wordt aangetroffen, gaat deze samen met concretiserende verwerkingsstrategieën. Men kan verwachten dat dergelijke samenhangen ook bij leerlingen in het beroepsonderwijs zichtbaar zullen zijn. De aangehaalde studies hebben evenwel plaatsgevonden in de context van theorieonderwijs of betroffen een theorie-gerelateerde leertaak. Ten aanzien van het onderhavige onderzoek wordt daarom verwacht dat de samenhang tussen concepties en strategieën naar voren zal komen in dat gedeelte van de interviewprotocollen dat betrekking heeft op de theoretische component van het onderwijsprogramma. Daar waar het de praktijkcomponent betreft zijn geen vergelijkbare studies voorhanden. Indien men aanneemt dat de categorieën verwerkingsstrategieën ook te herkennen zijn binnen deze praktijkcomponent, waar dit onderzoek

betrekking op heeft, kan men evenwel verwachten dat ook hier consistente samenhangen tussen leerconcepties en verwerkingsstrategieën zullen bestaan.

Samengevat gaat het om de volgende onderzoeksvragen.

1. Welke verwerkingsstrategieën en leerconcepties worden aangetroffen bij lerenden binnen het middelbaar beroepsonderwijs?¹
2. Is er sprake van een consistente samenhang tussen leerconcepties en verwerkingsstrategieën met betrekking tot het leren in het kader van theorievakken?
3. Is er sprake van een consistente samenhang tussen leerconcepties en verwerkingsstrategieën met betrekking tot het leren in praktijksituaties?

2. METHODE

Respondenten

Voor deze studie is de medewerking gevraagd van twee colleges voor middelbaar beroepsonderwijs, gevestigd in Rotterdam en Tilburg. Deze colleges verzorgen zowel lange als korte opleidingen binnen de volgende sectoren: Dienstverlening (DGO), Techniek (MTS) en Economie (MEAO). Binnen elke sector werden 20 leerlingen uitgenodigd deel te nemen aan het interview: tien uit de lange en tien uit de korte opleiding. De totale groep geïnterviewden bestond aldus uit 60 tweedejaars leerlingen. De leeftijd varieerde van 17 tot 20 jaar; er waren 28 mannelijke en 32 vrouwelijke respondenten.

Opzet en interviewprocedure

Selectie voor deelname vond steeds in de klas plaats, op willekeurige basis. Wanneer een leerling niet wenste deel te nemen, wat twee keer voorkwam, werd de naast hem of haar zittende leerling uitgenodigd. De deelnemende leerling kreeg een summiere uitleg over het doel van het onderzoek en werd verzocht zo veel mogelijk te vertellen over zijn of haar studiedrag. Tevens werd tijdens deze introductie de vertrouwelijkheid van de gesprekken benadrukt. De interviews vonden plaats in een leeg klaslokaal of een andere beschikbare ruimte binnen de school en werden met toestemming van de geïnterviewde op geluidsband opgenomen. Deelname was steeds vrijwillig; de respondenten ontvingen vergoeding noch terugkoppeling over de resultaten van hun interviews. Elk interview nam ongeveer 40 minuten in beslag.

Tijdens het interview werd gebruik gemaakt van een speciaal voor dit doel opgezet semi-structureerd interviewschema, dat bestond uit een 14-tal zo open mogelijke hoofdvragen met variabele doorvraagmogelijkheden. De hoofdvragen hadden, met uitzondering van twee inleidende vragen over achtergrondgegevens, betrekking op alle vier de onderdelen van leerstijlen. Voor het achterhalen van de verschillende soorten strategieën (verwerkings- en regulatiestrategieën) is gebruik gemaakt van een door Zimmerman en Pons (1986, 1988) ontwikkelde interviewprocedure. Deze komt er op neer dat aan de geïnterviewde wordt gevraagd zich verschillende leersituaties voor te stellen en te vertellen over de manier waarop hij of zij in die situatie gehandeld heeft. Voorgelegde situaties zijn onder andere: het leren voor een toets, leren voor een moeilijk vak, huiswerk maken, leren tijdens praktijksituaties en het leren in het kader van een hobby. Respondenten werd gevraagd iedere voorgelegde leersituatie te concretiseren door voorbeelden aan te dragen vanuit de eigen ervaringen op school. Bij de vragen over het leren voor een toets bijvoorbeeld, werd eerst gevraagd wat de laatste toets was die de respondent had gemaakt en werd verder gesproken over die bepaalde toets die door de respondent aangedragen was. Nadat op deze manier de leerstrategieën tijdens theorielessen en praktijksituaties besproken waren, kwamen de leerconcepties aan de orde. Hiertoe werd de respondenten gevraagd te omschrijven wat zij onder leren en goed onderwijs verstaan.

Data analyse

Ten behoeve van onderzoeksvraag 1 werden de geluidsbanden letterlijk uitgewerkt door een medewerker die niet bij de interviews betrokken was geweest. De resulterende protocollen werden beschouwd als ruwe data en vormden de basis voor analyse. Met het oog op de beantwoording van de eerste onderzoeksvraag zijn de protocollen op fenomenografische wijze geanalyseerd. Voor deze methode is gekozen omdat het de bedoeling was te inventariseren welke kwalitatieve verschillen zich voordoen in de wijze waarop leerlingen uit het MBO een aantal op de opleiding en het leren betrekking hebbende verschijnselen waarnemen, ervaren en conceptualiseren (Marton, 1986). De uitspraken van de leerlingen zijn gebruikt om een classificatieschema te ontwikkelen waarmee deze kwalitatieve verschillen kunnen worden gecategoriseerd (vergelijk Säljö, 1988). De eerste stap in de analyse bestond uit het selecteren van relevante uitspraken. Uitspraken en uitwijdingen die niet van belang werden geacht voor de onderzoeksvragen, werden verwijderd uit de dataset. Wat overbleef is een set met uitspraken die, op zichzelf of in de context van het beprokene, relevant zijn voor het leerstijlenconcept (vergelijk Vermunt, 1992). Deze geselecteerde citaten vormden de basis voor de volgende stap in de analyse. In deze volgende stap werden de citaten als afzonderlijke eenheden beschouwd, zonder aanzien des persoons en zonder rekening te houden met wat een individuele respondent verder aan uitspraken gedaan had. De uitspraken werden verschillende keren gelezen en in een iteratief proces op basis van overeenkomsten en verschillen gegroepeerd in een beperkt aantal categorieën. Vervolgens werden de categorieën beschreven en geïllustreerd met uitspraken uit de dataset (Marton, 1986, 1990). Deze categorieën lagen dus niet van tevoren vast.

Ter beantwoording van de onderzoeksvragen 2 en 3 werden twee onderzoeks-assistenten aangetrokken. Zij kregen de taak de interviewprotocollen te coderen met betrekking tot de eerder door de onderzoeker afgeleide categorieën voor verwerkingsstrategieën en leerconcepties. Aangezien de protocollen, als gevolg van individuele verschillen in bespraaktheid, zeer wisselend van omvang zijn, is ervoor gekozen om de codering te beperken tot de aan- en afwezigheid van de categorieën in de data. In dit onderzoek wordt dus geen rekening gehouden met de mate waarin een leerling gebruik maakt van een bepaalde verwerkingsstrategie, noch kunnen uitspraken gedaan worden over hoe sterk een leerconceptie bij een persoon aanwezig is.

Ter voorbereiding van de codering ontvingen de onderzoeksassistenten een instructie en een omschrijving van de gevonden categorieën, echter zonder voorbeelduitspraken. Hen is gevraagd per antwoord op een bepaalde vraag een codering van 1 of 0 toe te kennen voor ieder van de categorieën. Deze codering werd ingevuld in codeertabellen, waarop de relevante vraagnummers en de categorieën waarop diende te worden gescoord waren aangegeven. Tijdens de interviews zijn drie vragen gesteld over verwerkingsstrategieën tijdens theorie-onderwijs (vraag 5, 6 en 7, zie bijlage A) en drie vragen over verwerkingsstrategieën tijdens praktijkleersituaties (vraag 8, 9 en 10). Ook over leerconcepties zijn drie vragen gesteld in het interview. Echter, vraag 11 is de enige vraag die daar rechtstreeks betrekking op heeft. In het huidige onderzoek worden daarom, waar het gaat om de leerconcepties, alleen de antwoorden op deze vraag gecodeerd.

Tijdens de codering mochten de onderzoeksassistenten niet met elkaar overleggen, ten einde tot een onafhankelijke beoordeling te komen die een indruk kan geven van de interbeoordelaar-betrouwbaarheid. Deze bleek zeer acceptabel voor zowel de verwerkingsstrategieën tijdens het theorie-onderwijs, de verwerkingsstrategieën tijdens praktijkleersituaties, als voor de leerconcepties (Kappa's respectievelijk .83, .84 en .84). Daar waar de twee coderingen van elkaar afweken werd in samenspraak met de onderzoeker naar consensus gezocht en een definitieve codering aangebracht. Iedere respondent kreeg zodoende uiteindelijk drie scores toegekend: voor de dominante leerconceptie, de dominante leerstrategieën tijdens theoretische vakken en de meest gebruikte leerstrategieën tijdens praktijksituaties. Deze definitieve scoring werd gebruikt in de statistische analyses ter beantwoording van de tweede en de derde onderzoeksvraag.

3. RESULTATEN

In het algemeen, een enkele uitzondering daargelaten, bleken de respondenten goed in staat te verwoorden waarom ze voor een bepaalde opleiding gekozen hadden en wat ze onder leren verstaan. Moeilijker bleek het echter om te achterhalen wat de daadwerkelijke strategieën waren die men gebruikte bij het leren. De respondenten beschreven hun leerstrategieën eerder in termen van overte handelingen -zoals samenvattingen maken, hoofdzaken onderstrepen, een opdracht uitvoeren- dan in termen van coverte denkstrategieën als verbanden leggen, een probleem vanuit verschillende gezichtspunten bekijken, of het herhaaldelijk overdenken van een deel van de lesstof. Omdat het stellen van suggestieve vragen voorkomen moest worden, kwamen in sommige interviews de actuele verwerkings- en regulatiestrategieën vrijwel niet naar boven en bleven beschrijvingen beperkt tot bijvoorbeeld de volgende omschrijving: *"Ik ga gewoon aan het werk. Ik pak mijn boek en werk het van voor naar achter af."* (#38 kMEAO) Met betrekking tot de praktijksituaties was dit probleem het grootst, getuige de volgende kenmerkende uitspraak van een van de respondenten: *"Ik leer eigenlijk alle vakken op dezelfde manier. Behalve de vakken die ik voor praktijk heb maar daar hoeft je niet echt voor te leren. (...) Een praktijkvak is niet leren, dat is gewoon doen."* (#16 DGO)

Hieronder worden de gevonden verwerkingsstrategieën en de leerconcepties beschreven aan de hand van de categorieën die uit de interviews naar voren kwamen. Tevens worden, ter verduidelijking, de betrokken leerstijlcomponenten zelf nog eens kort beschreven.

Verwerkingsstrategieën

Met de term 'verwerkingsstrategieën' worden alle strategieën aangeduid die door lerenden gebruikt worden om informatie te verwerken waarmee ze in leersituaties geconfronteerd worden. Het betreft hier denkactiviteiten die een bijdrage kunnen leveren aan het verwerven van tot inzicht, vaardigheden en kennis. Bij lerenden in het MBO werden de volgende drie categorieën verwerkingsstrategieën onderscheiden: onbewerkt opslaan van informatie, conceptuele integratie en procedurele integratie.

1) Onbewerkt opslaan van informatie.

Met onbewerkt opslaan worden alle activiteiten aangeduid waarbij de leerstof herhaald worden zoals ze aangeboden zijn. Bij theorielessen valt hierbij te denken aan het inprenten van informatie door deze herhaaldelijk te lezen, over te schrijven of hardop uit te spreken. Het gaat hierbij om het uit het hoofd leren of memoriseren van de onderdelen die men belangrijk acht, zoals bijvoorbeeld rijtjes kenmerken of formules.

Met betrekking tot praktijksituaties heeft letterlijk herhalen meer het karakter van imitatie of oppervlakkige nabootsing. Handelingen worden dan net zo vaak herhaald tot ze snel en accuraat uitgevoerd kunnen worden. Op deze manier kunnen ook vaardigheden ontwikkeld, getraind en geautomatiseerd worden.

De volgende uitspraken wijzen op deze strategie:

"Gezondheidskunde. Eerst alles doorlezen en dan opschrijven. Dingen die ik niet kan onthouden die herhaal ik door ze op te schrijven. Dan schrijf ik ze gewoon tien keer over. Dus lezen, stukjes opschrijven, en gewoon herhalen." (#15 DGO)

"Bij verpleegkunde bijvoorbeeld moeten we een kraamvrouw, dat is dan een pop, aankleden en wassen, en baby's wassen. Het wordt uitgelegd door de lerares en ze doet het voor en daarna moeten we het zelf doen. Het staat nergens op papier. Ik probeer het hetzelfde te doen zoals de lerares het zegt. Zij vertelt dingen waar je op moet letten en daar probeer ik dan op te letten." (#11 DGO)

"Nederlands, woordjes bijvoorbeeld. Eerst de eerste tien goed leren en dan nog een keer lezen en nog een keer en dan ga ik het in mijn hoofd prenten. Dan weer de volgende tien, enzovoort. Aan het einde ga ik nog een keer alles doen, maar dan in een keer." (#38 MEAO)

"Het is steeds herhalen en daar leer ik het door. Door het vaak te lezen onthoud je het makkelijk." (#59 MTS)

"Opnieuw! Tot je het onder de knie hebt." (#18 kDGO)

2) Conceptuele integratie.

Bij conceptuele integratie gaat het om een dieper verwerking van informatie, waarbij het accent ligt op het begrijpen van de concepten en principes waarmee de lerende wordt geconfronteerd. Daartoe worden activiteiten ontplooid als het relateren van verschillende onderdelen van een vak of het zoeken naar overeenkomsten en verschillen tussen vakken, evenals het relateren van de voorkennis en ervaringen die men reeds bezit aan de nieuwe informatie.

Ook het structureren van de leerstof, het samenbrengen van stukken informatie in een conceptueel kader en het zich een totaalbeeld vormen van de stof vallen hieronder. De strategieën worden uitgevoerd met het oog op het verkrijgen van een beter begrip van de concepten en principes die in een bepaald vakgebied centraal staan. Wat betreft praktijksituaties worden ervaringen uit praktijklessen, stages of andere buitenschoolse activiteiten vooral gebruikt om tot een beter begrip van concepten en principes te komen.

De volgende uitspraken betreffen activiteiten in deze categorie:

"Psychologie vond ik ook best moeilijk. De tweede module ging over levensfasen, dat vond ik makkelijker. Je hebt dan al enkele dingen meegemaakt en dan is het toch makkelijker dan wanneer je het niet hebt meegemaakt. Dan weet je gewoon veel meer dingen en begrijp je beter waar het over gaat." (#18 DGO)

"Ik vertel voor mijzelf ook vaak hardop, dan kan ik het beter. Dan kan ik het beter onthouden, dat gaat gewoon beter. Dan heb ik in het boek iets gelezen en dan vertel ik hardop na wat ik gelezen heb, tot ik het zelf kan zeggen wat daar staat. En ook snap. Want soms staat het moeilijk omschreven en dan kun je dat wel uit je hoofd gaan leren maar als je niet weet wat je leert... Dus ik moet het eerst zelf snappen." (#11 DGO)

"Als je van stage terugkomt en dan weer theorie krijgt dan denk je: "Dat heb ik in de praktijk ook meegemaakt" en dan kun je het veel beter onthouden." (#15 DGO)

"Vanmiddag hebben we meettechniek. En dan gaan we eigenlijk van de theorie een aantal metingen doen. Dan krijg je even heel eenvoudig te zien wat er nou gaande is. Dat is belangrijk om te begrijpen wat je doet." (#59 kMTS)

3) Procedurele integratie.

Bij procedurele integratie gaat het om een diepe verwerking van informatie, waarbij het accent ligt op het verkrijgen en opbouwen van kennis omtrent procedures, aanpakken en werkwijzen waarmee de lerende op verschillende plaatsen of momenten is geconfronteerd. Indien stage gelopen wordt kan gedacht worden aan het vergelijken van dergelijke informatie naar aanleiding van stage-ervaringen en ervaringen uit de praktijklessen, maar ook aan het zich afvragen wat verschillende procedures, werkwijzen of aanpakken met elkaar te maken gehad hebben, wat overeenkomsten en verschillen waren, en dergelijke. Abstracte kennis, praktijkkennis en ervaringen worden geïntegreerd met de bedoeling een persoonlijke aanpak te ontwikkelen.

Lerenden omschrijven dit als volgt:

"De opdracht staat op een velletje, een tekening. Het wordt helemaal uitgelegd hoe je het het makkelijkst kunt doen. Hoe je je gereedschap moet behandelen. Ik heb het liefst de opdracht op papier want dan kan ik voor mezelf werken. Zelf proberen en naar eigen inzicht werken, kijken wat het makkelijkst is. Wat je in theorie leert is in de praktijk toch weer anders. Wat theoretisch niet kan, kan praktisch weer wel. Dat je het iets anders kunt doen dan eigenlijk moet." (#51 kMTS)

"Op stage leer je dingen die je op school niet leert. Foejjes en zo. Vaak zie je dat je die op

school ook gaat toepassen. Dus dan volg je de beschrijving niet helemaal meer." (#56 kMTS)

"Je leert het meeste van je collega's. Je leert anders dan op school. In de praktijkles legde de juffrouw iets uit en iemand anders had het al gehad op de stage. Het hangt er maar vanaf hoe iemand het uitlegt. De een zegt: zo tillen, en de ander zegt: zo tillen. Het hangt er vanaf wie je moet geloven. Degene uit de praktijk werkt er dagelijks mee, dus dan neem je maar aan dat dat hetzelfde werkt als dat het voor jou moet werken." (#17 kDGO)

"Ik leer er het meeste van als de uitleg met voorbeelden gaat. Dat helpt mij bij het proefwerk. Vaak het materiaal dat je erbij kunt gebruiken. Bijvoorbeeld dakbedekken, de verschillende dakbedekkingsmaterialen, dat helpt heel erg." (#56 kMTS)

Leerconcepties

Lerenden verschillen niet alleen in de strategieën die ze uitvoeren, maar ook in de manier waarop ze tegen leren en lesgeven aankijken. De een denkt bij leren bijvoorbeeld aan rijtjes die uit het hoofd geleerd moeten worden, de ander denkt aan problemen waarvoor een oplossing gevonden moet worden.

Uit de interviews kwamen drie verschillende categorieën van leerconcepties naar voren: leren als het opnemen van informatie, leren als het opbouwen van kennis en leren als het gebruiken van kennis.

1) Leren als het opnemen van informatie.

Bij deze leerconceptie wordt kennis gezien als een verzameling buiten de persoon aanwezige waarheden die getransporteerd dient te worden naar (het hoofd van) de lerende. Leren wordt hierbij gezien als een reproductief proces. De verantwoordelijkheid voor het (kant en klaar) aanbieden van de stof ligt bij de docent. Daarmee krijgt deze allerlei taken toebedeeld, zoals duidelijk uitleggen, voorbeelden geven, en precies voordoen hoe iets gedaan moet worden.

De volgende uitspraken illustreren deze leerconceptie:

"Leren is dingen lezen en onthouden. Gewoon goed leren zodat je het onthoudt, zodat je het blijft weten. Ja, dat is het doel van leren hè? Belangrijk zijn meestal de schuingedrukte regels, die leer je. Of moeilijke woorden uit de tekst, die staan dan uitgelegd, die leer je ook. Ik vind het makkelijk het helemaal te leren, ik zou het nooit zo snel in mijn eigen woorden opschrijven." (#14 kDGO)

"Leren is iets onthouden. Zodat je het later kunt herhalen. Voor maatschappijleer kan ik vandaag leren en morgen een proefwerk maken en het overmorgen alweer vergeten zijn, maar dan heb ik het wel geleerd. Dan heb ik het even onthouden." (#19 DGO)

2) Leren als het opbouwen van kennis.

Bij deze leerconceptie geeft de lerende aan dat hij geen genoegen neemt met enkel het opnemen van de aangeboden informatie. Hij hecht er waarde aan iets eigens toe te voegen aan de stof, een eigen mening te vormen over en verbanden te zoeken tussen verschillende bronnen van informatie. In deze opvatting heeft de lerende een actieve rol. Een rol waarin hij zelf zijn eigen kennis vormgeeft en zodoende een persoonlijk kennisbestand opbouwt waarin niet alleen plaats is voor de formeel aangeboden lesstof, maar ook voor bijvoorbeeld episodische kennis die in de loop der jaren is opgedaan en voorkennis die niet direct verworven is op de gevolgd opleiding.

De volgende uitspraken illustreren deze leerconceptie:

"Leren is denk ik beter weten dan gewoon weten. Het hoeft ook niet op school. De theorie moet je tot je eigen theorie maken, zodat je zelf begrijpt wat er staat. Je moet je eigen mening kunnen gebruiken en die vermengen met die van de docent." (#17 kDGO)

"Leren is gewoon een beetje weten wat het ongeveer inhoudt wat de leraar verteld heeft. Weten waar het over gaat, waar hij het over heeft of waar je op dat moment mee bezig bent.

Begrijpen. Als je iets moet leren vind ik dat je ongeveer moet weten waar hij het over heeft en dat je begrijpt wat hij ermee bedoelt. Ik geloof niet in letterlijk alles leren. Dan leer je het wel letterlijk maar dan weet je niet waar hij het over heeft. Als je begrijpt waar hij het over heeft en je krijgt een beetje een rare vraag dan kun je altijd nog een beetje kijken van het gaat daarover en dat bedoelt hij er ongeveer mee. En dan kom ik altijd in de problemen met mijn definities. Ik ben heel slecht in letterlijk leren, ik kan niets letterlijk onthouden. Dus ik heb altijd moeite met definities. Ik denk dat je een definitie moet begrijpen en niet letterlijk moet leren.” (#31 MEAO)

3) Leren als het gebruiken van kennis.

Bij deze opvatting van leren staat het toepassen van het geleerde centraal. Lerenden met deze leerconceptie hechten grote waarde aan de gebruiksmogelijkheden van de leerstof en zijn voortdurend op zoek naar mogelijke toepassingen van het lesmateriaal en de bijbehorende concrete situaties waarbinnen een dergelijke toepassing plaats kan vinden. De opleiding en de inhoud van de opleiding (zoals neergelegd in verschillende vakken of leerstofonderdelen) worden gewaardeerd in het licht van de toekomstige beroepspraktijk. Vaak is er reeds een goed gevormd en relatief volledig beeld van deze beroepspraktijk aanwezig bij de lerende en wordt de lesstof voornamelijk geïnterpreteerd in dat kader. Hierbij kan het, in de meest extreme vorm van deze leerconceptie, voorkomen dat vakken waarvan de toepassingsmogelijkheden niet direct helder zijn, als volledig irrelevant beschouwd worden.

Deze opvatting werd als volgt weergegeven:

“Nieuwe dingen die je nog niet kent, die dan toch langer blijven hangen. Meer dingen die een indruk op je maken. Hier leer ik nieuwe dingen die ik als ik deze opleiding dus niet gedaan zou hebben nooit geleerd had. Het zijn voor mij nieuwe dingen die ik nodig heb om het vak te kunnen beoefenen. De bedoeling van leren is dat de dingen blijven hangen. Je komt nieuwe dingen tegen, die leer je. Het zijn nieuwe dingen en door middel van leren blijven ze langer hangen. Het belangrijkste is dat het blijft hangen. Als je hier op school dingen leert en je later je vak uit gaat oefenen en denkt ‘dat heb ik ooit geleerd’ maar je weet het niet meer, dan heeft het geen nut gehad. Je bepaalt zelf ongeveer wat je later nodig zult hebben, dat kun je vantevoren toch nooit zeggen.” (#56 kMTS)

“Leren is dat je van een vak iets opsteekt, dat je er iets van leert. Dat je het later in de praktijk kunt brengen.” (#15 kDGO)

De samenhang tussen leerconcepties en verwerkingsstrategieën

Na scoring van de verwerkingsstrategieën tijdens theorie-onderwijs blijkt dat alle 60 respondenten gebruik maken van de verwerkingsstrategieën ‘onbewerkt opslaan van informatie’. Daarnaast geven 20 respondenten tevens blijk van het gebruik van ‘conceptuele integratie’ als verwerkingsstrategie. Voor het leren tijdens theorie-onderwijs blijkt de strategie ‘procedurale integratie’ niet gebruikt te worden. Binnen de praktijkleersituaties is het ‘onbewerkt opslaan van informatie’ te herkennen in de antwoorden van 59 respondenten. De conceptuele integratie is daarnaast herkenbaar bij vijf respondenten die ook blijk geven van ‘onbewerkt opslaan’ en één respondent die enkel deze strategie noemt. De procedurele integratie wordt aangetroffen bij 14 respondenten, waarbij ook het ‘onbewerkt opslaan van informatie’ herkenbaar is (zie tabel 1). De leerconceptie ‘leren als het opnemen van informatie’ is gevonden bij 48 respondenten, waarvan er 25 enkel deze leerconceptie noemen. De overige 23 noemen daarnaast nog andere leerconcepties. De leerconceptie ‘leren als het opbouwen van kennis’ is bij 20 respondenten en de leerconceptie ‘leren als het toepassen van kennis’ bij 15 respondenten zichtbaar (zie tabel 2).

Tabel 1: Aantallen respondenten (N = 60) aan wie een score is toegekend voor verschillende verwerkingsstrategieën tijdens theorie-onderwijs en praktijkleersituaties.

Verwerkingsstrategieën	Theorie-onderwijs	Praktijkleersituaties
Onbewerkt opslaan van informatie	60	59
Conceptuele integratie	20	6
Procedurale integratie	-	14

Tabel 2: Aantallen respondenten (N = 60) aan wie een score is toegekend voor verschillende leerconcepties.

Leerconcepties	Aantal respondenten
Leren als het opnemen van informatie	48
Leren als het opbouwen van kennis	20
Leren als het toepassen van kennis	15

Om inzicht te krijgen in de samenhang tussen leerconcepties en verwerkings-strategieën is een hercodering toegepast, waarbij elke respondent uiteindelijk drie scores krijgt toegekend: een score voor de leerconceptie, een score voor de verwerkingsstrategieën tijdens theorie-onderwijs en een score voor de verwerkings-strategieën tijdens praktijkleersituaties. Deze hercodering houdt in dat respondenten die blijf geven van het gebruik van verschillende verwerkingsstrategieën, of het hanteren van verschillende leerconcepties, worden gekenschetst aan de hand van een kenmerkende categorie. Bij zowel het theorie-onderwijs, de praktijkleersituaties als de leerconcepties, is een meest voorkomende categorie te onderscheiden. Indien een respondent daarnaast geen score voor een andere categorie heeft ontvangen, wordt deze als kenmerkende categorie opgevat. Indien er daarnaast wel een score aanwezig is voor een andere categorie, wordt deze als kenmerkende categorie aangeduid. Tijdens het theorie-onderwijs gebruiken bijvoorbeeld alle respondenten de strategie onbewerkt opslaan van informatie, terwijl 20 respondenten daarnaast ook gebruik maken van de conceptuele integratie. Aan deze 20 respondenten wordt nu een score toegekend voor conceptuele integratie, terwijl voor de overige 40 de strategie onbewerkt opslaan wordt gecodeerd. Voor de verwerkingsstrategieën tijdens praktijkleersituaties is een vergelijkbare hercodering toegepast, omdat, op één na, ook hier alle respondenten gebruik maken van de strategie onbewerkt opslaan. Aan de respondenten die daarnaast blijf geven van conceptuele integratie wordt nu een score voor conceptuele integratie toegekend, terwijl de respondenten die daarnaast ook gebruik maken van procedurele integratie als zodanig worden gecodeerd. Ook met betrekking tot de leerconcepties is op deze manier een hercodering toegepast.

Tijdens het theorie-onderwijs gebruiken bijvoorbeeld alle respondenten de strategie onbewerkt opslaan van informatie, terwijl 20 respondenten daarnaast ook gebruik maken van de conceptuele integratie. Aan deze 20 respondenten wordt nu een score toegekend voor conceptuele integratie, terwijl voor de overige 40 de strategie onbewerkt opslaan wordt gecodeerd. Voor de verwerkingsstrategieën tijdens praktijkleersituaties is een vergelijkbare hercodering toegepast, omdat, op één na, ook hier alle respondenten gebruik maken van de strategie onbewerkt opslaan. Aan de respondenten die daarnaast blijf geven van conceptuele integratie wordt nu een score voor conceptuele integratie toegekend, terwijl de respondenten die daarnaast ook gebruik maken van procedurele integratie als zodanig worden gecodeerd. Ook met betrekking tot de leerconcepties is op deze manier een hercodering toegepast.

De samenhang tussen leerconcepties en verwerkingsstrategieën is af te lezen uit de tabel-

len 3 en 4, die betrekking hebben op respectievelijk het theorie-onderwijs en de praktijkleersituaties.

Tabel 3: Kruistabel voor leerconcepties (kolommen) en verwerkingsstrategieën tijdens theoretische vakken (rijen); kolompercentages; Phi.

THEORETISCHE VAKKEN	Leren als het opnemen van informatie	Leren als het opbouwen van kennis	Leren als het toepassen van kennis	Rijtotalen
Onbewerkt opslaan van informatie	22 (88%)	8 (65%)	10 (60%)	40 (67%)
Conceptuele integratie	3 (12%)	12 (10%)	5 (13%)	20 (10%)
Kolomtotalen	25 (100%)	20 (100%)	15 (100%)	60 (100%)

Tabel 4: Kruistabel voor leerconcepties (kolommen) en verwerkingsstrategieën tijdens praktijksituaties (rijen); kolompercentages; Phi.

PRAKTIJK SITUATIES	Leren als het opnemen van informatie	Leren als het opbouwen van kennis	Leren als het toepassen van kennis	Rijtotalen
Onbewerkt opslaan van informatie	18 (72%)	13 (65%)	9 (60%)	40 (67%)
Conceptuele integratie	2 (8%)	2 (10%)	2 (13%)	6 (10%)
Procedurele integratie	5 (20%)	5 (25%)	4 (27%)	14 (23%)
Kolomtotalen	25 (100%)	20 (100%)	15 (100%)	60 (100%)

Phi = .11 (niet significant)

In tabel 3 is, met betrekking tot het leren binnen het theorie-onderwijs, een onevenredige verdeling te zien van verwerkingsstrategieën over de verschillende leerconcepties. Deze is voornamelijk toe te schrijven aan verschillen tussen de groepen respondenten met de leerconceptie 'leren als het opnemen van informatie' en 'leren als het opbouwen van kennis'. Bij de eerste groep wordt de strategie 'onbewerkt opslaan van informatie' relatief vaak en de strategie 'conceptuele integratie' relatief weinig gevonden. Binnen de tweede groep ligt de verdeling precies andersom: daar ligt de nadruk op de conceptuele integratie. De verdeling van verwerkingsstrategieën bij de groep respondenten met de leerconceptie 'leren als het toepassen van kennis' komt overeen met de verwachte waarden voor de totale groep.

Uit tabel 4, die betrekking heeft op praktijkleersituaties, is af te lezen dat de verdeling van verwerkingsstrategieën over de verschillende leerconcepties nauwelijks afwijkt van de verwachte waarden voor de totale groep.

4. CONCLUSIES EN DISCUSSIE

Uit de interviewstudie onder MBO-studenten kwamen drie categorieën naar voren voor verwerkingsstrategieën en drie voor leerconcepties. De gevonden overeenkomst van 83-84% tussen de coderingen van twee onafhankelijke beoordelaars steunt het idee dat we hier te maken hebben met betekenisvolle, inhoudelijk van elkaar verschillende en duidelijk herkenbare categorieën. Ook de relatief snelle consensus die bereikt werd over dat deel van de uitspraken dat oorspronkelijk verschillend gecodeerd werd, vormt een bevestiging van de helderheid van de beschreven categorieën.

Vergeleken met de categorieën die voor het hoger onderwijs werden beschreven is het aantal gevonden categorieën echter laag. In de ILS-HO zijn bijvoorbeeld vijf subschalen opgenomen voor de verwerkingsstrategieën (verdeeld over drie hoofdschalen) en nog eens vijf hoofdschalen voor de leerconcepties. Het verschil in aantal beschrijvingscategorieën kan het gevolg zijn van verschillen tussen de doelgroepen. Vergeleken met het hoger onderwijs bijvoorbeeld, zijn studenten in het middelbaar beroepsonderwijs wat jonger en wellicht minder getraind in het reflecteren op en verbaliseren van hun denkprocessen. Dit kan er de oorzaak van zijn geweest dat de gedane uitspraken op globaler niveau geïnterpreteerd werden dan bij de studenten van het hoger onderwijs mogelijk was. Kijkt men bijvoorbeeld naar de drie categorieën van verwerkingsstrategieën die gevonden werden onder studenten in het beroepsonderwijs, dan is enige gelijkheid met de hoofdschalen stapsgewijze verwerking, diepteverwerking en concrete verwerking uit het hoger onderwijs te zien. Een verdere detaillering zoals die in de subschalen van Vermunt aanwezig is, bleek echter niet mogelijk.

Een tweede opmerkelijk punt is dat twee van deze verwerkingsstrategieën bij MBO-studenten zowel voor de praktijksituaties als de theorielessen aangetroffen werden. De oppervlakkige en de diepere verwerkingsstrategieën worden gebruikt bij zowel abstract als praktisch leermateriaal. Dit onderscheid geldt dus niet enkel cognitieve taken (bijvoorbeeld Entwistle, 1981; Schmeck, 1988; Vermunt, 1992), maar heeft ook betrekking op het leren van praktische vaardigheden.

Met betrekking tot de leerconcepties komen de drie gevonden categorieën (leren als het opslaan van informatie, leren als het opbouwen van kennis, leren als het toepassen van kennis) inhoudelijk overeen met de concepties die reeds voor het hoger onderwijs beschreven werden. Twee andere werden echter niet aangetroffen: stimulerend onderwijs en samen studeren. Strikt genomen zijn dit echter geen leerconcepties, maar betreft het hier concepties van de condities waaronder leren plaats dient te vinden. Deze twee correleerden in Busato's studie dan ook niet met verwerkingsstrategieën (Busato, 1998; p. 65) en hadden in Vermunts regressieanalyse slechts in enkele gevallen een -lage- voorspellende waarde (Vermunt, 1992; p. 120-121).

De resultaten van de analyses, zoals weergegeven in de kruistabellen, tonen aan dat er voor de theoretische vakken wel een relatie is gevonden tussen verwerkingsstrategieën en leerconcepties, maar voor de praktijkvakken niet. Met betrekking tot de theorievakken ligt de gevonden samenhang (op basis van kruistabellen kan niet gesproken worden van causaliteit) in de lijn der verwachtingen: de leerconceptie 'leren als het opnemen van informatie' gaat relatief vaak samen met de verwerkingsstrategie 'onbewerkt opslaan van informatie' en de leerconceptie 'leren als het opbouwen van kennis' met de strategie 'conceptuele integratie'. Van Rossum en Schenk (1984), Vermunt (1992) en Busato (1998) vonden dezelfde relaties bij studenten in het hoger onderwijs. De inhoudelijke samenhang tussen leerconcepties en uit-

gevoerde verwerkingsstrategieën geldt dus voor beide respondentengroepen en kan wellicht gegeneraliseerd worden naar andere populaties, voor zover die te maken hebben met leersituaties waarin het leren van schriftelijk en theoretisch materiaal centraal staat.

Opmerkelijker is het gegeven dat een dergelijke relatie niet vastgesteld kon worden voor de meer praktische leersituaties. Hoewel ook hiervoor een codering van dominante verwerkingsstrategieën aangebracht kon worden, bleek deze niet consistent samen te gaan met de leerconcepties. Men kan zich afvragen waar dit door veroorzaakt wordt, en of bijvoorbeeld de gehanteerde onderzoeksmethode hier een rol speelt. Tijdens de interviews werden de verwerkingsstrategieën namelijk afzonderlijk voor theorielessen en praktijksituaties afgevraagd, terwijl de vragen over leerconcepties algemeen gehouden werden, vanuit de veronderstelling dat leerconcepties overkoepelende ideeën over leren betreffen. Een mogelijk gevolg van deze open vraagstelling is dat niet altijd alle op leren betrekking hebbende gedachten geuit worden en de uitspraken beperkt blijven tot de eerste associaties bij het woord 'leren'. Het is niet ondenkbaar dat deze eerste associaties meer in de lijn van theorieonderwijs dan van praktijkonderwijs liggen (vergelijk de in de inleiding van de resultatensectie gegeven uitspraak "... Een praktijkvak is niet leren, dat is gewoon doen."). Echter, deze vraag naar leerconcepties werd gesteld aan het einde van het interview, nadat reeds uitgebreid gesproken was over onder andere het leren tijdens praktijkvakken, praktica, stages en buitenschoolse activiteiten. Verondersteld kan worden dat hierdoor een breder referentiekader over leren is gecreëerd. Bovendien was het mogelijk de algemene beschrijvingen van verwerkingsstrategieën te herkennen in zowel theoretische als praktische leersituaties. Het blijft onduidelijk waarom de algemene leerconcepties dan niet ook betrekking zouden hebben op praktijksituaties.

Het gebrek aan verbanden tussen algemene leerconcepties en verwerkingsstrategieën tijdens praktijkleersituaties, samen met het verwachte verband tussen leerconcepties en verwerkingsstrategieën tijdens het theorieonderwijs, doet vermoeden dat de leerconcepties minder algemeen zijn dan werd verondersteld. Blijkbaar liggen er andere ideeën ten grondslag aan leergedrag tijdens praktijksituaties dan aan leergedrag tijdens theorieonderwijs. Een vraag die hieruit voortvloeit is in hoeverre leerconcepties situatiespecifiek zijn; een tweede vraag is in hoeverre de gevonden leerconcepties generaliseerbaar zijn naar andere leersituaties dan geïnstitutionaliseerd theorieonderwijs.

Een alternatief voor het achterhalen van op praktijksituaties betrekking hebbende leeractiviteiten en leerconcepties is het in het geheel vermijden van de term 'leren'. Dit zou tevens het naar boven komen van eventueel sterke eerste "theorie" en "boek"-gerichte associaties kunnen voorkomen. Eraut (1994) heeft bijvoorbeeld een dergelijke methode gehanteerd in een onderzoek naar competentie-ontwikkeling in werksituaties. Hij vroeg aan respondenten wat er de oorzaak van is dat ze nu iets wel kennen of kunnen dat ze voorheen nog niet konden of konden. Ook de 'story-line' methode (Gergen, 1988; Sannen, 1997; Van der Sanden en Teurlings, 1998) biedt aanknopingspunten voor het in kaart brengen van leeractiviteiten en leerconcepties tijdens praktijksituaties. Voor het achterhalen van praktijkleerconcepties moet echter nog een goede methode ontwikkeld worden. Gezien het grote belang van leren in praktische situaties binnen alle vormen van beroepsopleidingen en het weinige dat er over bekend is, is dit zeker een relevant thema voor vervolgonderzoek.

REFERENTIES

- Boaler, J. (1998). Alternative approaches to teaching, learning and assessing mathematics. *Evaluation and Program Planning*, 21(2), 129-141.
- Busato, V.V. (1998). *Leerstijlen nader geanalyseerd*. Universiteit van Amsterdam: academisch proefschrift.
- Eklund-Myrskog, G. (1997). Students' views of learning in vocational education. *Scandinavian Journal of Educational Research*, 41(2), 179-188.
- Engeström, Y. (1994). *Training for change: new approach to instruction and training in working life*. Geneva: International Labour Office.

- Entwistle, N. (1981). *Styles of Learning and Teaching*. Chichester: Wiley.
- Eraut, M. (1994). *Developing professional knowledge and competence*. London: The Falmer Press.
- Gergen, M. (1988). Narrative structures in social explanation. In: C. Antaki (Ed.), *Analysing every day explanation. A casebook of methods* (pp. 94-112). London: Sage.
- Kwakman, K. (1993). *Informeel en intentioneel leren*. Katholieke Universiteit Nijmegen: Doctoraalscriptie Onderwijskunde.
- Marton, F. (1986). Phenomenography - A research approach to investigating different understandings of reality. *Journal of Thought*, 21, 28-49.
- Marton, F. (1990). The phenomenography of learning - a qualitative approach to educational research and some of its implications for didactics. In: H. Mandl, E. de Corte, S.N. Bennett & H.F. Friedrich (Eds.), *Learning and Instruction - European research in an international context. Volume 2.1* (pp. 601-616). Oxford: Pergamon Press.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning: 1 - Outcome and process. *British Journal of Educational Psychology*, 46(4), 4-11.
- Mulder, R.H. (1997). *Leren ondernemen: ontwerpen van praktijkleersituaties voor het beroepsonderwijs*. Erasmus Universiteit Rotterdam: Academisch proefschrift.
- Nieuwenhuis, A.F.M. (1991). *Complexe leerplaatsen in school en bedrijf*. Groningen, RION: Academisch Proefschrift.
- Onstenk, J.H.A.M. (1998). Discussiebijdrage aan thema Bedrijfs- en Beroepsopleidingen & Volwasseneneducatie. In: J. van den Akker, J. Pieters, I. Visscher-Voerman & A. Wald (Red.), *Studiehuis en onderwijs-onderzoek*. Leuven/Apeldoorn: Garant.
- Renkl, A., Mandl, H., & Gruber, H. (1998). Inert knowledge: Analyses and remedies. *Educational Psychologist*, 31(2), 115-121.
- Resnick, L.B. (1987). Learning in and out school. *Educational Researcher*, 16(9), 13-20.
- Rietbroek, J.C.M. (1998). *HEAO-studenten: leerstijlen, persoonlijkheidskenmerken en leeractiviteiten op stage*. Katholieke Universiteit Brabant, Onderwijs- en Opleidingspsychologie: Doctoraalscriptie.
- 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.
- Säljö, R. (1988). Learning in educational settings: methods of inquiry. In: P. Ramsden (Ed.), *Improving learning: new perspectives* (pp. 32-48). London: Kogan Page.
- Sanden, J.M.M. van der (1997). *Duurzame ontwikkeling van leervermogen. Leren leren in het technische domein*. Technische Universiteit Eindhoven: Intreerede.
- Sanden, J.M.M. van der, & Bussel, F.J.J. van (Eds.) (1995). *Atrium in Europe. The development of learning abilities in youth training*. Bielefeld: W. Bertelsmann Verlag.
- Sanden, J.M.M. van der, & Teurlings, C.C.J. (1998). *Developing competence during practice periods: the learners' perspective*. Poster, presented at the European COST conference. Universiteit van Newcastle, November 1998.
- Sannen, H. (1997). *De levensloop als verhaal. Jongeren en hun tijdsperspectief 1955-1995*. Tilburg: TUP.
- Schmeck, R.R. (1988). Individual differences and learning strategies. In: F. Marton, D. Hounsell & N. Entwistle (Eds.), *The experience of learning* (pp. 171-191). Edinburgh: Scottish Academic Press.
- Slaats, A. (1997). *Inventaris Leerstijlen voor het Middelbaar beroepsonderwijs. Handleiding en vragenlijst*. Tilburg: KUB/STAR
- Slaats, A. (1999). *Reproduceren & Construeren. Leerstijlen van leerlingen in het middelbaar beroepsonderwijs*. Academisch proefschrift. Tilburg: A. Slaats.
- Slaats, A., Lodewijks, H.G.L.C., & Sanden, J.M.M. van der (1994). *Leerstijlen in het middelbaar beroepsonderwijs: verslag van een interviewstudie*. Tilburg: Katholieke Universiteit Brabant. Intern onderzoeksverslag, 24 pp.
- Slaats, A., Lodewijks, H.G.L.C., & Sanden, J.M.M. van der (1999). Learning styles in secondary vocational education: disciplinary differences. *Learning and Instruction*, 9(5), 475-492.
- 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.
- Vries, B. de (1988). *Het leven en de leer: een studie naar de verbinding van leren en werken in de stage*. Nijmegen/TTS: Academisch proefschrift.
- WEB (1996). *Wet Educatie en Beroepsonderwijs*. Zoetermeer: OCW.
- Zimmerman, B.J., & Pons, M.M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23(4), 614-628.
- Zimmerman, B.J., & Pons, M.M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, 80(3), 284-290.

NOOT

1 De dataverzameling op basis waarvan deze vraag beantwoord wordt, had oorspronkelijk ook betrekking op regulatiestrategieën en motivationele oriëntaties (zie Slaats, 1999; Slaats, Lodewijks & Van der Sanden, 1999). In het kader van dit artikel beperken we ons echter tot een bespreking van de verkregen gegevens omtrent verwerkingsstrategieën en leerconcepties.

Epistemologische opvattingen of mentale leermodellen in het Hoger Onderwijs: een verkenning van het terrein¹

Gonny Schellings (*Onderwijs en Educatie; Katholieke Universiteit Nijmegen*)
Hans Lodewijks (*Onderwijs- en Opleidingspsychologie; Katholieke Universiteit Brabant*)
Johan van der Sanden (*Technische Universitaire Lerarenopleiding; Technische Universiteit Eindhoven*)

ABSTRACT

In reviews about epistemological beliefs and conceptions about knowledge acquisition or learning, a variety of theoretical frameworks or models is observed. Empirical studies in this area are characterized by multiple conceptualizations and instruments, as well. This article is exploring possible interplays between epistemological beliefs (what is knowledge about) and learning conceptions (what is learning about). Although, we are mainly aimed at theoretical issues, we will illustrate our considerations with the help of two explorative studies, in which epistemological beliefs were related to learning conceptions. Since the translated version of the Epistemological Questionnaire (Schommer, 1992) yielded rather inconvenient psychometric results, we have to be reserved in drawing conclusions. The results supported the idea that beliefs about knowledge are distinct from conceptions of learning. By merging the epistemological scales with learning conception scales, we found an interesting distinction between a reproductive and constructive view of knowledge and learning.

INLEIDING

Onder invloed van het constructivisme, waarbinnen persoonlijke kennisconstructies een voorname rol spelen, is er aandacht gekomen voor subjectieve opvattingen en theorieën die leren vormen over het leren op school (zie bijv. Wigfield, Eccles & Pintrich, 1996). Gedurende het leren vormen lerenden persoonlijke theorieën over wat leren/studeren in het algemeen is en wat leren/studeren voor hen persoonlijk betekent. Het idee is dat de lerenden leerinhouden, leerdoelen en verschillende didactische maatregelen vanuit een persoonlijk referentiekader interpreteren en van daaruit verschillende soorten leeractiviteiten ontplooiën. Een subjectieve leertheorie is op te vatten als een richtinggevend interpretatie- en handelingskader voor het leren. Niet de leersituatie op zich, maar de perceptie en opvattingen van de leerling omtrent de leersituatie zijn richtinggevend aan het leren (Van der Sanden, 1997). Aan dit persoonlijk referentiekader zijn diverse aspecten te onderscheiden die min of meer kunnen samenhangen, waaronder bijvoorbeeld epistemologische opvattingen en leerconcepties. Maar de vraag kan worden gesteld of het conceptuele onderscheid dat in de literatuur wordt gemaakt tussen epistemologische opvattingen en leerconcepties wel noodzakelijk of empirisch houdbaar is.

In een review over epistemologische theorieën concluderen Hofer en Pintrich (1997) dat er in de literatuur weinig overeenstemming bestaat of concepties over intelligentie, over leren

en over instructie wel of niet deel uit (zouden moeten) maken van een epistemologische theorie (i.e. het geheel van samenhangende concepties over kennis). Zelfs de inspirator van het epistemologisch onderzoek, William Perry, analyseerde kennisconcepties aan de hand van interviews met studenten over hun ervaringen in het onderwijs, waardoor de beschreven ontwikkeling toch niet helemaal gevrijwaard is gebleven van opvattingen over leren en onderwijs. Tijdens hun intellectuele en ethische ontwikkeling doorlopen studenten verschillende stadia (Perry, 1970; zie ook de inleiding van dit themanummer Ten Dam, Severiens & Vermunt, 2000). In de eerste stadia (dualisme) zien de studenten kennis als "juist" of als "onjuist". Om te weten te komen wat juist is, vallen ze terug op een autoriteit (bijv. docent). Als studenten zich ontwikkelen, zien ze in dat bepaalde uitspraken tegenstrijdig kunnen zijn (multipliciteit), maar ze zullen nog steeds blijven zoeken naar *het* juiste antwoord. Daarna komen studenten in een stadium dat ze meerdere standpunten als juist kunnen aanvaarden (relativisme). In de laatste stadia verbinden studenten zich aan het relativisme: ze nemen een standpunt in over de waarde van kennis, waarbij het vormen van een persoonlijke identiteit centraal staat (commitment). Hoewel Perry verbanden tussen deze epistemologische ontwikkeling met cognitieve stijlen, leerstijlen en het ontwikkelen van het leervermogen, zeer plausibel vond, heeft hij deze relaties nooit zelf onderzocht. Zijn werk heeft echter geleid tot tal van alternatieve epistemologische denkkaders, waarbinnen kennis- en leerconcepties meer en meer door elkaar gingen lopen (zie Hofer & Pintrich, 1997).

Hofer en Pintrich (1997) gaan er van uit dat leerconcepties waarschijnlijk wel samenhangen met kennisconcepties, maar dat beide constructen op conceptueel niveau toch onderscheiden moeten blijven. Volgens hen omvat het epistemologische construct twee "kern"-dimensies, namelijk ideeën omtrent de aard van kennis en omtrent het proces van het weten. Opvattingen over de aard van kennis kunnen variëren van het idee dat kennis vaststaat of dynamisch is, dat kennis bestaat uit absolute waarheden of uit conditionele 'werk'-hypothesen en dat kennis een aaneenschakeling is van op zich geïsoleerde feiten of bestaat uit een netwerk van onderling gerelateerde concepten. Opvattingen over het proces van het weten hebben vooral betrekking op hoe je kennis kunt opbouwen. Bijvoorbeeld de opvatting dat kennis wordt overgenomen van een externe autoriteit (bijv. docent) of geleidelijk ontstaat in samenwerking met anderen (juist door de interactie). Andere opvattingen omtrent het proces van het weten hebben betrekking op het rechtvaardigen van de opgedane kennis: worden "waarheden" onderbouwd met argumenten en/of bewijsstukken, worden deskundigen geraadpleegd en worden hun uitlatingen kritisch bekeken. Naar ons inzien is op deze "proces van het weten"-dimensie (hoe wordt kennis opgebouwd) het conceptueel onderscheid tussen kennis- en leerconcepties vertroebeld, terwijl Hofer en Pintrich zelf pleiten om dit onderscheid te blijven maken.

Schommer (1990, 1994) gaat voorbij aan het conceptueel onderscheid tussen kennis en leerconcepties en definieert epistemologische opvattingen breder, namelijk als opvattingen van lerenden over de aard van kennis *en over kennisverwerving* (i.e. "what knowledge is and how knowledge is acquired," p.27, 1994). Hierbij ontleent zij weliswaar het kennistheoretische concept 'epistemologie' aan de filosofie, maar ze hanteert de term toch vooral in ruime zin. Het denken over kennis en leren wordt samengevat in een systeem van min of meer onafhankelijke epistemologische dimensies. Aan de ene zijde van *elke* dimensie bevindt zich een naïeve opvatting over kennis en leren en aan de andere zijde bevindt zich een meer genuanceerd of verfijnd perspectief.

Geïnspireerd door het werk van Perry, Dweck en Leggett, en Schoenfield stelt Schommer vijf dimensies voor, maar ze sluit niet uit dat er meer dimensies bestaan. Een eerste dimensie betreft de *zekerheid* van kennis, variërend van het idee dat kennis absoluut is tot het idee dat kennis onzeker is. De dimensie is te vergelijken met het onderscheid dat Perry maakt tussen dualisme en relativisme. Ook een tweede dimensie van Schommer vertoont raakvlakken met deze ontwikkeling en betreft de organisatie of *structuur* van kennis: kennis kan worden opgevat als uiteenvallend in geïsoleerde stukjes, of als een samenstelsel van geïntegreerde con-

cepten. Een derde dimensie betreft de *bron* waaruit de lerende de kennis moet opdoen, of anders gezegd, of de kennis wordt overgedragen door een autoriteit (bijv. studiemateriaal, docent) of wordt het verworven door denkwerk van de lerende zelf. Een vierde dimensie van Schommer komt voort uit het werk van Dweck en Leggett en betreft de twee impliciete theorieën die lerenden kunnen hebben over intellectuele vaardigheid: de "entity" en de "incremental" theorie. De entity-theorie beschouwt intellectuele vaardigheid als een vaststaand gegeven: leerervaringen hebben geen invloed op intellectuele vaardigheid. De incremental theorie neemt aan dat de intellectuele vaardigheid zich verder ontwikkelt en bijvoorbeeld door leren vergroot kan worden. Schommer vertaalt deze theorieën in de dimensie "*innate ability*", met als het naïeve perspectief dat leervaardigheid een vaststaand gegeven is vanaf de geboorte en het andere perspectief dat de vaardigheid om te leren kan worden geleerd. De laatste dimensie, die Schommer noemt, heeft te maken met de *snelheid van leren*: leren is een snel en alles of niets fenomeen (i.e. als de lerende het niet snel leert, dan leert hij/zij het nooit) versus het leren verloopt moeizaam en geleidelijk. Uitgaande van dit kader heeft Schommer een vragenlijst ontwikkeld, waardoor het mogelijk wordt epistemologische opvattingen in kaart te brengen.

Volgens Hofer en Pintrich (1997) zijn drie van de dimensies van Schommer "zuivere" epistemologische dimensies, de overige betreffen meer "perifere" opvattingen over leren, instructie en intelligentie. De leerconceptie van Schommer (i.e., de snelheid van leren) vinden we echter niet terug in andere modellen die specifiek gericht zijn op leerconcepties. Voordat we epistemologische opvattingen kunnen vergelijken met leerconcepties, gaan we eerst kort in op de kwestie waaruit leerconcepties bestaan (zie ook Ten Dam, et al., 2000).

Lodewijks en Aarnoutse (1995) omschrijven een leerconceptie als een relatief stabiel en samenhangend geheel van opvattingen, overtuigingen, (meta)cognities en affecties met betrekking tot het leren: wat leren inhoudt, hoe het leren het best kan worden aangepakt en waartoe het leren dient of moet leiden. Binnen een leerconceptie kan onderscheid worden gemaakt tussen (meta)cognitieve en affectieve aspecten. Het cognitieve aspect betreft bijvoorbeeld opvattingen over wat leren inhoudt, wat het nut van leren is. Het affectieve aspect heeft te maken met opvattingen van de lerende over zichzelf in relatie tot het leren, de leersituatie, en dergelijke. Hierbij gaat het om opvattingen, overtuigingen, waarden, normen en attitudes over diverse aspecten van de werkelijkheid en van zichzelf. Een belangrijk onderdeel betreft het zelfbeeld: hoe competent schat een leerling zichzelf in?

Hoewel de definitie van leerconceptie diverse samenhangen veronderstelt, worden de afzonderlijke componenten vaak apart bestudeerd, waardoor leerconcepties eenzijdig dreigen te worden belicht. Marton, Dall'Alba en Beaty (1993), bijvoorbeeld, beperken zich strikt tot de cognitieve aspecten. De verschillende leerconcepties die zij voorstellen komen alle voort uit een algemeen beeld dat mensen hebben van leren: naast een tijdspanne, nemen lerenden in het algemeen aan dat leren een acquisitie- en een toepassingsfase omvat. Marton en zijn collega's suggereren dat er een ontwikkeling in leerconcepties bestaat net zoals de ontwikkeling in kennisconcepties van Perry. Deze ontwikkeling begint met de opvatting dat leren simpelweg het vergroten van kennis is en eindigt met de fase, waarin leren wordt opgevat als een proces van persoonlijke ontwikkeling.

Vermunt (1992) vond dat leerconcepties (in zijn theorie "mentale leermodellen") een belangrijke component zijn van het persoonlijke referentiekader waarmee lerenden leersituaties tegemoet treden. Elk mentale leermodel is karakteristiek voor een bepaalde leerstijl. Studenten met een reproductieve leerstijl zien studeren vooral als het zo nauwgezet mogelijk overnemen van aangeboden informatie. Studenten met een betekenisgerichte leerstijl vatten studeren vooral op als het zelf construeren van kennis. In het mentaal leermodel van studenten met een toepassingsgerichte leerstijl staat het toepassen van kennis centraal. De ongerichte leerstijl wordt gekarakteriseerd door twee mentale leermodellen: de studenten hechten veel waarde aan samenwerking met hun medestudenten en aan onderwijs dat hen stimuleert.

In tegenstelling tot Marton en zijn collega's beperkt Vermunt zichzelf niet tot "zuivere"

leerconcepties. Juist de twee laatste mentale leermodellen die de ongerichte leerstijl typeren zijn meer opvattingen over hoe het onderwijs vorm zou moeten krijgen. Vermunt vindt (net zoals Schommer haar epistemologische dimensies opvat) dat de vijf mentale leermodellen min of meer onafhankelijk van elkaar zijn.

Hoewel sommige onderzoekers zich vooral concentreren op kennisconcepties (bijv. Schommer) en anderen zich meer richten op leerconcepties (bijv. Vermunt), lijken de inhoudelijke verschillen tussen de concepties klein. Een lerende die van mening is dat kennis bestaat uit absolute waarheden (i.e., zekerheid van kennis) en uit geïsoleerde stukjes feitenkennis (i.e. structuur van kennis) zal waarschijnlijk ook het idee hebben dat kennis "klakkeloos" op een reproductieve manier overgenomen moet worden (i.e., leren is kennisopname), waarbij de rol van de deskundigen (i.e. autoriteit is bron van kennis) groot is. Een lerende, daarentegen, die van mening is dat kennis niet absoluut is en bestaat uit een geïntegreerd geheel van concepten, zal waarschijnlijk ook van mening zijn dat je kennis *zelf* moet opbouwen (i.e., leren is kennis opbouwen). Bij het opbouwen van een dergelijk netwerk zal het leren waarschijnlijk langzamer gaan dan wanneer de aangeboden feiten uit het hoofd moeten worden geleerd. De rol van de deskundige (i.e., bron van kennis) zal waarschijnlijk groot zijn als de lerende van mening is dat het onderwijs een belangrijke stimulerende rol moet innemen, terwijl de rol van de deskundige minder belangrijk lijkt te zijn als de lerende vooral wil leren in samenwerking met mede-lerenden.

Kortom, de meeste kennis - en leerconcepties lijken niet alleen aan elkaar gerelateerd te zijn, maar lijken ook met elkaar verweven. In theorievorming over subjectieve leertheorieën moeten we eerst nagaan of het gewenst is deze aspecten uit elkaar te halen (Van der Sanden, 1997). De instrumenten, zoals die door de onderzoekers zelf ontwikkeld zijn, geven een handvat bij het zoeken naar de overeenkomsten en verschillen tussen de concepten. Daarnaast kan bekeken worden of de instrumenten bruikbaar zijn in het onderzoek naar "subjectieve leertheorieën". In de onderstaande studies zijn onder meer twee vragenlijsten afgenomen. Zo is er gebruik gemaakt van een "getrouwe" vertaling (en een "vereenvoudigde" versie) van Schommer's lijst (Epistemological Questionnaire; Schommer, 1992) en het deel "opvattingen" uit de Inventaris LeerStijlen (ILS; Vermunt, 1992; 1997). De eerste studie bij studenten uit het wetenschappelijk onderwijs had vooral tot doel de bruikbaarheid van het epistemologische instrument na te gaan. De tweede studie bij HBO-studenten maakte deel uit van een groter project naar de leereffecten van stages (Teurlings & Van der Sanden, 1999). Onder de diverse instrumenten, die hiertoe werden afgenomen, bevonden zich ook de vragenlijst van Schommer en de ILS. Ook in deze studie kijken we eerst naar de psychometrische kwaliteiten van het epistemologisch instrument.

Studie 1: Epistemologische opvattingen in het Wetenschappelijk Onderwijs

Methode

Respondenten. De vragenlijst werd (op vrijwillige basis) door 120 studenten aan de Katholieke Universiteit Brabant ingevuld. Er waren 43 studenten van de economische faculteit, 52 studenten van de sociale faculteit en 25 studenten van de juridische faculteit (50 vrouwen en 70 mannen). De studenten waren overwegend ouderejaars (30% tweedejaars, 35% derdejaars en 17% vierdejaars).

Materiaal. De studenten kregen een vragenlijst voorgelegd die bestond uit twee gedeeltes: opvattingen over kennis (Epistemological Questionnaire) en opvattingen over leren (ILS-Versie HO).

1. Epistemological Questionnaire: WO-versie (Schommer, 1992). De vijf epistemologische dimensies heeft Schommer uitgewerkt tot een lijst met 63 uitspraken over kennis en leren die gegroepeerd worden in 12 sub-sets van vragen. Elke uitspraak wordt herleid naar een 'naïeve' kijk op kennis en leren, waar de respondent het wel of niet mee eens kan zijn. In empirisch onderzoek (Schommer, 1990, 1993, Schommer, Crouse & Rhodes, 1992) vindt Schommer evenwel vier factoren: "fixed ability", "simple knowledge", "certain knowledge", en

“quick learning”. Dit betekent dat de bron van kennis (de rol van deskundigen) als dimensie verdwijnt, en de twee schalen waaruit een bron-van-kennis-factor zou moeten bestaan (“vertrouw op deskundigen” en “oordeel niet over deskundigen”) laden op twee verschillende factoren. Schommer presenteert de schalen die laden op een factor als volgt: “fixed ability” (vaardigheid is vast): bij de eerste poging vindt leren meteen plaats, lerenden kunnen niet leren hoe ze moeten leren, succes is niet gerelateerd aan hard werken, leervermogen is aangeboren en geconcentreerd leren is tijdverspilling. De tweede factor is “simple knowledge” (kennis is simpel): ambiguïteit moet worden vermeden, leren is het zoeken naar eenduidige betekenissen, integratie moet worden vermeden, lerenden zijn afhankelijk van deskundigen. De derde factor is “quick learning” (leren is snel) en heeft slechts één subschaal: leren is snel. De vierde en laatste factor “certain knowledge” (kennis is zeker) omvat twee subschalen: kennis is zeker en lerenden moeten deskundigen niet bekritisieren of beoordelen. 27 items hebben gezien vanuit deze naïeve kijk een ‘negatieve’ strekking (i.e., verwoorden het genuanceerde perspectief). Voordat er enige analyses uitgevoerd kunnen worden moeten de negatieve items worden gehercodeerd. De respondenten geven bij elke uitspraak hun mening op een vijf-punten schaal (1= helemaal mee oneens, 2= voor een groot deel mee oneens, 3= neutraal, 4= voor een groot deel mee eens, 5= helemaal mee eens).

2. *ILS-HO* (Vermunt & van Rijswijk, 1997). De ILS is ontwikkeld om leerstijlen te inventariseren en omvat dus meer uitspraken dan die betrekking hebben op mentale leermodellen (zie ook de andere bijdragen in dit thema-nummer). Het opvattingen-deel van de ILS bestaat uit 5 schalen met uitspraken over leren en onderwijs (in totaal zijn er 40 items). De schalen komen overeen met de vijf mentale leermodellen: leren is kennis opnemen (=,71), leren is kennis opbouwen (=,67), leren is kennis toepassen (=,80), leren is vooral samenwerken met medelerenden (=,89) en het onderwijs moet stimuleren (=,85). Er wordt gewerkt met dezelfde vijf-punten schaal.

Procedure. De studenten vulden de lijsten (vrijwillig) in tijdens colleges. Deze colleges werden gekozen op basis van de medewerking van de docent die bereid was college-tijd af te staan voor het onderzoek. Het invullen van de lijst nam ongeveer 20-25 minuten in beslag.

Dataverwerking. De ingevulde vragenlijsten werden geanalyseerd met behulp van SPSS.

Resultaten: eerste studie

Voordat we de resultaten van de ‘Epistemological questionnaire’ vergelijken met het opvattingen-deel uit de ILS, willen we eerst kijken naar de psychometrische gegevens van de vertaalde lijst: vinden we vergelijkbare interne consistenties op de (sub)schalen en een zelfde factor-structuur als in het onderzoek van Schommer? Is deze lijst een handig en/of bruikbaar instrument?

Instrument van Schommer: WO-versie. In tegenstelling tot het onderzoek van Schommer vinden we lage betrouwbaarheden, zeker voor de sub-schalen en voor de hoofdschaal “snelheid van leren” (zie tabel 1). Schommer zelf rapporteert betrouwbaarheden in de range van .51-.78 (Schommer, 1993) en .63-.85 (Schommer en Walker, 1995). Jehng, Johnson en Anderson (1993) vonden met een aangepaste versie (in deze versie werden 29 items verwijderd en de beoordelingschaal bestond uit zeven punten) lage betrouwbaarheden voor de schalen (.42-.59). De betrouwbaarheden van de door ons gebruikte vertaling lijken echter zeer laag (zeker voor de schaal snelheid van leren).

Ondanks de laag intern consistente opbouw van de schalen, voerden we toch, in navolging van Schommer, een factor-analyse (principale componenten analyse met varimax rotatie) uit op de 12 schalen.² Niet onverwacht vinden we noch de theoretische factor-structuur (vijf dimensies), noch de empirische factor-structuur (i.e., de structuur die Schommer heeft gevonden in haar onderzoeken) terug. Overigens vond Schommer zelf in drie van haar studies (Schommer 1990, 1993; Schommer, et al. 1992) ook andere factor-oplossingen dan de verwachte (theoretische) structuur; bovendien verschilden de factor-oplossingen tussen de stu-

dies behoorlijk: drie schalen (leervermogen is aangeboren, geconcentreerd leren is tijdverspilling, oordeel niet over deskundigen) gedroegen zich niet op eenzelfde manier, anders gezegd, laadden op andere factoren. De schalen die wel op dezelfde factoren laadden waren: "niet leren leren", "leren in één keer" en "succes is niet gerelateerd aan hard werken" op de factor "vaardigheid is vast". De schalen "vermijd ambiguïteit", "zoeken van eenduidige betekenis", "vermijd integratie" en "vertrouw op deskundigen" laadden op de factor "kennis is simpel", de schaal "leren is snel" vormde de factor "leren is snel" en de schaal "kennis is zeker" laadde consistent als enige schaal op de factor "kennis is zeker".

Tabel 1: Betrouwbaarheden (Cronbach Alpha's) voor sub- en hoofdschalen Schommer WO-versie (n=110) en HBO-versie (n=119)

Schaal		aantal items	α
WO-versie			
Vaardigheid is vast	(5 subschalen)	18	.55 (range: .28-.36)
Kennis is simpel	(4 subschalen)	28	.63 (range: .20-.50)
Leren is snel	(0 subschalen)	5	.14
Kennis is zeker	(2 subschalen)	12	.41 (range: .31-.44)
HBO-versie			
Vaardigheid is vast	(5 subschalen)	18	.36 (range: .13-.48)
Kennis is simpel	(4 subschalen)	28	.32 (range: -.34-.12)
Leren is snel	(0 subschalen)	5	.22
Kennis is zeker	(2 subschalen)	12	.29 (range: .12-.21)

Zoals tabel 2 laat zien, die we overigens voorzichtig moeten interpreteren omdat de betrouwbaarheden laag zijn, komen in deze studie de factoren "leren is snel" en "kennis is zeker" niet direct terug. De eerste factor, die we vinden, is opgebouwd uit de schalen "leren is snel" en "leervermogen is aangeboren". In tegenstelling tot de bevindingen van Schommer, laadt de schaal "zoeken van eenduidige betekenis" ook op deze eerste factor (Schommer et al., 1992; Schommer, 1993). De tweede factor bestaat uit de schalen "vermijd ambiguïteit", "oordeel niet over deskundigen", en "vermijd integratie". De derde factor bestaat uit de schalen "niet leren leren" en "succes is niet gerelateerd aan hard werken". Een derde schaal "vertrouw op deskundigen" heeft een negatieve factorlading op de derde factor en daarnaast heeft deze schaal een hoge factorlading op de eerste factor. Tot slot, bestaat de vierde factor uit de schalen "geconcentreerd leren is tijdsverspilling" en "leren in een keer". De schaal "kennis is zeker" heeft niet alleen een negatieve factorlading op deze vierde factor, maar laadt ook op de eerste én tweede factor. Kortom, de gevonden factor-oplossing is weinig overtuigend: noch de theoretische, noch de empirische oplossing wordt gevonden, acht van de 12 schalen hebben een redelijke factorlading op twee of drie factoren, en op de eerste factor laden al acht van de 12 schalen.

Omdat de resultaten sterk afwaken van de gegevens van Schommer en ook de betrouwbaarheden laag waren, hebben we iteratief items uit de hoofd-schalen verwijderd tot de Cronbach alpha's niet meer verder toenamen. Dit leidde ertoe dat 25 items verwijderd werden, negen items van de schaal "vaardigheid is vast" ($=.60$), acht items van "kennis is simpel" ($=.71$), twee items van "leren is snel" ($=.40$) en zes items van "kennis is zeker" ($=.56$). De aldus "opgevaardeerde" betrouwbaarheden zijn voor exploratieve onderzoeksdoeleinden acceptabel.

Tabel 2: *Principale componenten factoranalyse (varimax-rotatie) op schalen van Schommer*

WO-versie	F1	F2	F3	F4
leren is snel	.68			
leervermogen is aangeboren	.68			
zoek eenduidige betekenis	.60		-.25	-.26
<i>vermijd ambiguïteit</i>		.79		
oordeel niet over deskundigen		.74		
<i>vermijd integratie</i>	.36	.58		
<i>niet leren leren</i>		.25	.79	
<i>succes is niet hard werken</i>	.28		.69	
vertrouw op deskundigen	.42		-.60	
<i>geconcentreerd leren is tijdverspilling</i>	.26			.73
<i>leren in één keer</i>		.35	.27	.63
kennis is zeker	.33		.29	-.48
Eigenwaarde	2.8	1.7	1.2	1.1
% verklaarde variantie	23.5	14.5	10.2	8.8
HBO-versie	F1	F2	F3	F4
<i>niet leren leren</i>	.70			
<i>succes is niet hard werken</i>	.68			
leren is snel	.54	.40	.38	-.33
geconcentreerd leren is tijdsverspilling	.52			
oordeel niet over deskundigen	.51		.31	.50
leervermogen is aangeboren		.70		
vertrouw op deskundigen		.66		
<i>vermijd integratie</i>		.60		
<i>vermijd ambiguïteit</i>			.74	
leren in een keer		-.35	.63	
zoek eenduidige betekenis	-.31		.54	
<i>kennis is zeker</i>				.84
Eigenwaarde	2.2	1.9	1.4	1.1
% verklaarde variantie	18.2	16.0	12.0	9.0

Factorloadingen tussen -.25 en .25 zijn weggelaten; cursieve schalen horen volgens de empirische bevindingen van Schommer oorspronkelijk bij elkaar.

De vertaalde versie van Schommers vragenlijst lijkt echter nog allesbehalve gebruiksgereed. Diverse extra analyses liggen hierbij voor de hand. Qian en Alvermann (1993) verwijderden, bijvoorbeeld, eerst de twee schalen die oorspronkelijk bedoeld waren voor de bron-van-kennis-dimensie. Vervolgens voerden zij een factor-analyse uit op item-niveau, waarbij nog eens 21 items werden verwijderd. De uiteindelijke oplossing bestond uit drie factoren (leren is snel; kennis is simpel en zeker; leervermogen is aangeboren). Jehng et al. (1993) verwijderden 29

items met behulp van de item-totaal-correlaties en discriminatie-vermogen van de items en vonden daarna met behulp van LISRELL-analyse de vijf epistemologische dimensies. Deze onderzoekers vonden overigens lage betrouwbaarheden (range: .42- .59) nog acceptabel om groepen hierop te kunnen vergelijken.

Studie 2: Epistemologische opvattingen in het Hoger BeroepsOnderwijs

Methode

Respondenten. Deze studie maakte deel uit van een grootschalig project bij derde-jaarstudenten van een opleiding Sociaal-Pedagogische Hulpverlening-opleiding. De studenten werden op het "stage-criterium" geselecteerd, zodat ook enkele tweede-jaars en vierde-jaars studenten participeerden. De bedoeling was dat alle studenten op drie momenten tijdens hun stage-periode een vragenlijst zouden invullen. Uitval onder de respondenten was relatief hoog: de eerste lijst werd ingevuld door 166 studenten, de tweede lijst door 138 studenten en de laatste lijst door 83 studenten. Op deze plaats zijn we echter vooral geïnteresseerd in het onderdeel over de kennis-opvattingen, ongeacht het aantal vragenlijsten dat de student heeft getourneerd. Dit onderdeel werd door 137 studenten ingevuld, waaronder 9 tweedejaars, 12 vierdejaars en 108 derdejaars. Er waren 125 vrouwen en 8 mannen, met een gemiddelde leeftijd van 21,6 jaar (range 18-32 jaar).

Materiaal De studenten kregen drie uitgebreide vragenlijsten, die verschillende onderdelen bevatten. In dit artikel beperken we ons tot de HBO-versie van de Epistemological Questionnaire (onderdeel van de tweede vragenlijst) en het opvattingen-deel uit de ILS (onderdeel van de eerste vragenlijst).

1. *Epistemological Questionnaire: HBO-versie.* De WO-versie (i.e., de getrouwe vertaling van de Epistemological Questionnaire) werd enigszins vereenvoudigd in taalgebruik en sommige items werden aangepast met behulp van termen die kenmerkend zijn voor het HBO, bijvoorbeeld, "cursussen" werd vervangen door "modulen". De data-verzameling van de eerste studie was nog niet afgerond toen de HBO-vragenlijst werd uitgezet. Daarom werd het aantal items en de strekking van de items niet aangepast.

2. *ILS-HO.* De ILS-HO werd volledig afgenomen: naast de opvattingen-schalen werden ook uitspraken over verwerking- en regulatie-strategieën en oriëntaties aan de studenten voorgelegd. De ILS-HO versie wordt zonder aanpassingen zowel in het WO als in het HBO afgenomen (Vermunt & van Rijswijk, 1997).

Procedure. Het afnemen van de vragenlijsten werd volledig georganiseerd door medewerkers van de HBO-opleiding. Zij zorgden voor de verspreiding, alsook voor het geven van reminders (i.e., studenten werden tijdens mentoraatsuren gevraagd de lijsten alsnog te retourneren). Hoewel er plenaire bijeenkomsten waren georganiseerd waarin de studenten de lijsten (vrijwillig) konden invullen, werden er ook veel lijsten buiten de georganiseerde uren ingevuld en per post getourneerd.

Dataverwerking. De ingevulde vragenlijsten werden geanalyseerd met behulp van SPSS.

Resultaten: tweede studie

Ook in deze studie hebben we ons allereerst geconcentreerd op de psychometrische gegevens van de HBO-versie van de "Epistemological Questionnaire".

Instrument van Schommer: HBO-versie. Onze verwachting was dat het vereenvoudigen van de items zou leiden tot meer consistente schalen, maar het tegendeel bleek waar: slechts één hoofdschaal en twee subschalen vertonen een hogere consistentie dan op de WO-versie; de overige sub- en hoofdschalen laten zelfs lagere alpha's zien (zie tabel 1).

Gezien de lage interne consistentie van de schalen was het niet te verwachten dat een factoranalyse (principale componenten analyse met varimax rotatie op schaalniveau) een duidelijk plaatje zou opleveren. Toch zijn er wel overeenkomsten te zien met de factor-oplossingen van Schommer zelf (zie tabel 2). Er is echter geen overeenkomst met de factor-oplossing die voor de WO-versie gevonden was. Weliswaar vinden we op de HBO versie ook vier fac-

toren, maar slechts twee schalen (niet leren leren en succes is niet hard werken) laden consistent op een factor.

Ook op de HBO-versie verwijderden we iteratief items uit de hoofdschalen tot de Cronbach alpha's niet meer toenamen om acceptabeler betrouwbaarheden te vinden. Nu moesten we 31 items verwijderen, zeven uit "vaardigheid is vast", 14 uit "kennis is simpel", twee uit "leren is snel" en acht uit "kennis is zeker". Hoewel we begonnen met lagere betrouwbaarheden voor de HBO-versie, vinden we uiteindelijk alpha's die redelijk te vergelijken zijn met de alpha's voor de WO-versie. Toch werden niet zelfde items verwijderd, en ook de volgorde van verwijdering was verschillend. Op de schaal "vaardigheid is vast" ($= .52$) werden vier zelfde items verwijderd als op de WO-versie, zeven items uit "kennis is simpel" ($= .71$), één uit "leren is snel" ($= .45$) en vier items uit "kennis is zeker" ($= .55$).

Kortom, de versies WO en HBO leiden tot verschillende resultaten, die onverwacht groot blijken. Dit zou erop kunnen wijzen dat niet alleen de aanpassingen in de taal, maar ook het veranderen van doelgroep (studenten in het HBO) tot deze verschillen geleid hebben. Het ILS-onderdeel met betrekking tot de mentale leermodellen resulteert voor de HBO-groep overigens wel in vergelijkbare interne consistenties als voor de WO-groep (kennis-opname: .77; kennis-opbouw: .77; kennis-toepassen: .62; samen studeren: .86; stimulerend onderwijs: .86). Dit zou erop kunnen wijzen dat, in vergelijking met de WO-studenten, kennisconcepties van HBO-studenten misschien minder uitgekristalliseerd zijn dan leerconcepties.

Studie 1 en 2: Vergelijking van kennisconcepties en mentale leermodellen.

Hoewel de resultaten van de vragenlijst van Schommer niet overtuigen, willen we toch de relaties tussen kennis- en leerconcepties verkennen. Hiertoe bezien we eerst de correlaties tussen de schalen, waarbij we er vooral op letten of de 'gereduceerde' schalen van Schommer correleren met de schalen van Vermunt. Dit doen we voor de WO- en HBO-versies afzonderlijk, simpelweg omdat de vragenlijsten zelf verschillen.

In de correlatie-tabel valt op dat de 'gereduceerde' Schommer-schalen onderling redelijk tot hoog correleren (zie tabel 3). Voor de WO-groep correleren alle schalen met elkaar, voor de HBO-groep vinden we vooral correlaties met de schaal "kennis is simpel". "Kennis is simpel" correleert positief met "leren is snel", maar negatief met de andere twee schalen "vaardigheid is vast" en "kennis is zeker". Deze resultaten lijken in tegenspraak met het uitgangspunt van Schommer dat de dimensies relatief onafhankelijk zijn. Positieve correlaties zijn echter wel te verklaren vanuit het idee dat een lerende over *alle* dimensies heen, of een naïef, of een genuanceerd denkbeeld heeft. Negatieve correlaties tussen de dimensies zijn moeilijker te verklaren.

Ook bij de mentale leermodellen worden correlaties tussen de ILS-schalen gevonden, en wel bijna dezelfde voor de afzonderlijke groepen: positieve correlaties tussen "leren is kennis opbouwen" en "leren is kennis toepassen"; "onderwijs moet stimuleren" en "onderwijs is samen leren"; "onderwijs moet stimuleren" en "leren is kennis opnemen". De positieve correlatie tussen de beide onderwijsopvattingen is niet onverwacht, want beide opvattingen zijn kenmerkend voor de ongerichte leerstijl (Vermunt, 1992). Voor de HBO-groep wordt ook een correlatie gevonden tussen "onderwijs is samen leren" en "leren is kennis opnemen".

We zijn echter vooral geïnteresseerd in de correlaties tussen de Schommer en ILS-schalen. Voor beide groepen vinden we drie verwachte correlaties: "leren is kennis opnemen" is positief gecorreleerd met "kennis is simpel"; "leren is kennis opbouwen" is negatief gecorreleerd met "kennis is zeker" en "leren is kennis gebruiken" is negatief gecorreleerd met "vaardigheid is vast". Voor de WO-groep hangt "leren is kennis opnemen" ook nog positief samen met "kennis is zeker". En voor de HBO-groep vinden we een negatieve correlatie tussen "leren is kennis opbouwen" en "vaardigheid is vast". Dus we vinden enkele van de verwachte samenhangen tussen de twee lijsten terug. Een onverwachte maar kleine correlatie vinden we voor de WO-groep tussen de onderwijsopvatting "onderwijs moet stimuleren" en de epistemologische opvatting "kennis is simpel".

Tabel 3: Correlatie-coëfficiënten tussen de schalen voor studie 1 en studie 2

WO	1	2	3	4	5	6	7	8
1. Vaste vaardigheid	-							
2. Kennis is simpel	.382**	-						
3. Leren is snel	.351**	.384**	-					
4. Kennis is zeker	.254*	.349**	.306*	-				
5. Kennis-opname	.109	.546**	.120	.302*	-			
6. Kennis-opbouw	-.164	-.124	-.190	-.308*	-.120	-		
7. Kennis gebruiken	-.279*	.018	-.140	-.218	.035	.468**	-	
8. Samen leren	-.187	.061	.151	-.008	.103	.150	.136	-
9. Stimulerend	-.087	.253*	.131	.011	.315*	.107	.083	.328**

HBO	1	2	3	4	5	6	7	8
1. Vaste vaardigheid	-							
2. Kennis is simpel	-.251*	-						
3. Leren is snel	.023	.413**	-					
4. Kennis is zeker	.184	-.240*	-.163	-				
5. Kennis-opname	-.033	.381**	.147	.092	-			
6. Kennis-opbouw	-.321*	.113	-.021	-.262*	-.202	-		
7. Kennis gebruiken	-.263*	.189	.070	-.139	-.110	.213*	-	
8. Samen leren	-.156	.087	.120	-.179	.214*	.064	.043	-
9. Stimulerend	.040	.045	.167	.028	.427**	-.076	.048	.427**

** p < 0.001

* p < 0.01

Tot slot, voerden we een principale componenten factoranalyse (met varimax rotatie) op schaalniveau uit. Omdat we vooraf vooral geïnteresseerd waren in de wederzijdse invloeden van epistemologische opvattingen met leeropvattingen, kozen we ervoor in de factor-analyses alleen die schalen te betrekken die *tussen* de lijsten een correlatie hadden. Voor de beide groepen vinden we twee factoren, die we op dezelfde wijze kunnen interpreteren. Eén factor duidt op een genuanceerd denkbeeld over leren en kennis, de tweede factor lijkt vooral te duiden op een meer naïef denkbeeld. Deze tweedeling in opvattingen doet denken aan het onderscheid dat Byrnes (1996) maakt tussen twee verschillende typen wetenschappelijke leertheorieën, namelijk "objectivistische en constructivistische" leertheorieën. In de discussie zullen we hierop terugkomen.

Opvallend is dat voor de HBO-groep de schaal "leren is kennis toepassen" laadt op twee factoren, terwijl in de WO-groep deze schaal deel uitmaakte van het genuanceerde denkbeeld. Misschien betekent "leren is kennis toepassen" toch iets anders voor HBO-studenten dan voor WO-studenten.

Conclusie en discussie

De belangrijkste vraag, die wij ons in deze bijdrage hebben gesteld, is of we epistemologische concepten en mentale leermodellen van elkaar kunnen en/of moeten scheiden. Bij het vergelijken van de lijsten van Schommer en Vermunt vinden we wel enkele overeenkomsten tussen specifieke schalen, maar deze resultaten zijn niet dusdanig van aard dat de concepten met elkaar verweven lijken te zijn. Wel is er een interessante tweedeling in denkbeelden te maken. Aan de ene kant staat een "objectivistisch" denkbeeld: kennis is objectief beschikbaar in de omgeving en de lerende maakt zich de kennis eigen door de externe informatie te transporteren naar het interne geheugen. Aan de andere kant staat een "constructivistisch" denkbeeld over kennis en leren, dat veronderstelt dat objectieve kennis niet bestaat. In de omge-

ving bestaat alleen informatie en leren wordt dan opgevat als het *transformeren* van informatie vanuit een persoonlijk referentie-kader tot interne kennis (Byrnes, 1996; Van der Sanden, 1997). Deze tweedeling lijkt overigens op het onderscheid dat Schommer maakt in naïef en genuanceerd denken over leren en kennis, waarbij Schommer deze verder differentieert in vier of vijf dimensies (zie inleiding). In onze gevonden tweedeling komt het onderscheid in *meer* dimensies echter niet meer terug. Ook de psychometrische gegevens van de Schommer-vragenlijst geven niet veel aanleiding om te blijven denken in vier of vijf dimensies, die min of meer onderelateerd zijn aan elkaar.

Tabel 4: *Principale componenten factoranalyses (varimax rotatie) op de gereduceerde Schommer en Vermunt-schalen voor studie 1 en studie 2*

WO	F1	F2
Leren is kennis opbouwen	.76	
Leren is kennis gebruiken	.76	
Vaardigheid is vast	-.58	
Kennis is zeker	-.57	.40
Kennis is simpel		.81
Leren is kennis opnemen		.81
Stimulerend onderwijs	.30	.62
Eigenwaarde	2.24	1.68
% verklaarde variantie	32.0	24.0
HBO	F1	F2
Leren is kennis opbouwen	.79	
Vaardigheid is vast	-.60	
Kennis is zeker	-.59	
Leren is kennis opnemen	-.28	.85
Kennis is simpel	.28	.71
Leren is kennis toepassen	.49	.44
Eigenwaarde	1.87	1.39
% verklaarde variantie	31.2	23.2

Factorloadingen tussen -.25 en .25 zijn weggelaten

Een belangrijke bevinding, die wij in deze bijdrage moeten melden, betreft juist de zorgwekkende psychometrische gegevens van de vertaalde versies van Schommer's vragenlijst. Het instrument is verre van gebruiksklaar en de bovenstaande exploratie moet met de nodige voorzichtigheid worden gehanteerd. Ook Schommer en andere onderzoekers rapporteren regelmatig afwijkende resultaten (bijvoorbeeld andere factor-oplossingen). Sterker, zoals een anonieme reviewer voor een Amerikaans tijdschrift opmerkte: "my understanding in talking to colleagues is that this is one of a growing number of studies that did not replicate Schommer's factor structure, yet these studies remain unpublished, a problem that may impede progress in the field".

De door ons gevonden problemen hebben dus niet alleen met de vertaling te maken (sommige items komen wel zeer Amerikaans over, zoals "nothing is certain, but death and taxes"). Pajares (1992), bijvoorbeeld, wijst op een paar specifieke problemen die vragenlijsten met betrekking tot opvattingen hebben. Opvattingen kunnen nogal context-specifiek zijn, waardoor "individual items fall prey to 'it depends'- thinking" (p.227). Daarnaast kunnen afzonderlijke opvattingen deel uitmaken van een breder geheel of een samengesteld denkbeeld, zonder dat de afzonderlijke opvattingen sterk met elkaar behoeven samen te hangen. Juist dit idee van "nonconsensuality" van opvattingen kan volgens Qian en Alvermann (1993) resulteren in lage interne consistenties tussen de items op schaal-niveau, zeker wanneer de items op voorhand al naar meer geïsoleerde opvattingen lijken te verwijzen (bijv. de items van de schaal "kennis is simpel": "Ik hou niet van films met een open einde" en "Goed studeren betekent over het algemeen feiten onthouden") in plaats van een parafrase van het item (bijv. de items van de schaal "leren is kennis opnemen": "Onder leren versta ik: ervoor zorgen dat ik de feiten die ik in de cursus aangeboden krijg, kan reproduceren" en "Zelf moet ik definities en andere feiten van buiten leren"). Voorts werd ons duidelijk dat de vertaalde vragenlijst van Schommer allesbehalve eenduidige reacties bij de respondenten oproept en ook niet altijd eenvoudig is in te vullen. WO-studenten gaven bijvoorbeeld aan veel meer moeite te hebben met de Schommer-uitspraken dan met de ILS-uitspraken. Als reden gaf men veelal aan dat de vragen van Schommer op een "vage", "zweverige" en "filosofische" manier geformuleerd waren, hetgeen een sterkere neiging tot 'it depends-thinking' oproep dan bij de meer concreet geformuleerde ILS-items.

Bovenstaande resultaten roepen natuurlijk om analyses op item-niveau (zie ook Hofer & Pintrich, 1997). Hierbij is het ook interessant om te kijken of de afzonderlijke items van de verschillende lijsten c.q. verschillende schalen conceptueel en empirisch te scheiden zijn. Sommige Schommer-items (bijvoorbeeld "When I study, I look for the specific facts") lijken nogal verwant aan de ILS-items ("Onder studeren versta ik ervoor zorgen dat ik de feiten die ik in een cursus aangeboden krijg kan reproduceren."). Het samenvoegen van de WO- met de HBO-groep zou ons wel in staat hebben gesteld om een factor-analyse op item-niveau uit te voeren, maar we hebben hiervoor niet gekozen, onder meer omdat de beide vragenlijsten verschillen in taalgebruik en formulering.

Naast epistemologische opvattingen en mentale leermodellen die op schaalniveau samenhangen, zijn er ook schalen die niet samenhangen. Het zonder meer buiten sluiten van deze schalen in vervolgonderzoek is ons inziens niet aan te bevelen, omdat dit een wellicht onverantwoorde beperking impliceert van de veronderstelde complexiteit van de subjectieve leertheorieën, die studenten zich vormen tijdens hun studie. Subjectieve leertheorieën omvatten meer aspecten dan alleen de rol en aard van kennis en de manier waarop kennis verworven wordt (Van der Sanden, 1997). In plaats van de lijsten van Schommer en Vermunt te reduceren tot een paar schalen, zouden we de bruikbaarheid van alle schalen moeten nagaan. Misschien krijgen we door de twee lijsten te integreren (waarbij we het onderscheid kennis- en leerconceptie laten vallen) meer inzicht in subjectieve leertheorieën. Een theorie wordt hierbij opgevat als een samenhangend geheel van ideeën en geïntegreerde opvattingen, waarbij 'strikte' scheidingen tussen de diverse losse ideeën moeilijk te maken zijn (zie ook Pajares, 1992).

Deze bijdrage laat zien dat op het gebied van subjectieve leertheorieën nog veel werk te verrichten is. Zowel op het theoretische als op het empirische vlak staan nog vele vragen open. Theoretische bijdragen helpen de dynamische processen en de individuele verschillen in het schoolse leren te begrijpen. Omdat individuele opvattingen richting geven aan de manier waarop leerlingen hun leertaken en leeromgeving benaderen, kan dit soort onderzoek helpen leerlingen in een gewenste richting na te laten denken over hun leren.

LITERATUUR

- Byrnes, J.P. (1996). *Cognitive development and learning in instructional contexts*. Boston: Allyn and Bacon.
- Hofer, B.K. & Pintrich, P.R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88-140.
- Jehng, J.J., Johnson, S.D., & Anderson, R.C. (1993). Schooling and students' epistemological beliefs about learning. *Contemporary Educational Psychology*, 18(1) 23-35.
- Lodewijks, J.G.L.C., Aarnoutse, C.A. (1995). *Ontwikkeling van leerconcepties en (effectief) studiegedrag in de onderbouw van het Voortgezet Onderwijs: een longitudinaal onderzoek*. Voorstel voor een aio-project. Nijmegen: Katholieke Universiteit - Vakgroep Onderwijskunde.
- Marton, F., Dall'Alba, G. & Beatty E. (1993). Conceptions of Learning. *International Journal of Educational Research*, 19, 277-300.
- Pajares, M.F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62, 307-332.
- Perry, W.G., jr. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart and Winston.
- Qian, G., & Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text. *Journal of Educational Psychology*, Vol 87(2) 282-292.
- Sanden, J.M.M. van der (1997). *Duurzame ontwikkeling van leervermogen. Leren leren in het technische domein*. Inaugurale rede. Eindhoven: Technische Universiteit Eindhoven.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82, 498-504.
- Schommer, M. (1992). *Epistemological Questionnaire* (Revised for High School Students). Wichita State University, USA.
- Schommer, M. (1993). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology*, Vol 85(3), 406-411.
- Schommer, M. (1994). An emerging conceptualization of epistemological beliefs and their role in learning. In R. Garner & P.A. Alexander (Eds.). *Beliefs about text and instruction with text* (pp. 25-40). Hillsdale, NJ: Lawrence Erlbaum Ass.
- Schommer, M., Crouse, A., Rhodes, N. (1992). Epistemological beliefs and mathematical text comprehension: Believing it is simple does not make it so. *Journal of Educational Psychology*, Vol 84, 435-443.
- Schommer, M., & Walker, K. (1995). Are epistemological beliefs similar across domains? *Journal of Educational Psychology*, Vol 87, 424-432.
- Ten Dam, G., Severiens, S., & Vermunt, J.D.H.M. (2000). Inleiding op het thema "Leerconcepties". *Tijdschrift voor Onderwijsresearch*.
- Teurlings, C.C.J., & Van der Sanden, J.M.M. (1999). "Het vormt je in wat je eigenlijk wilt worden". Onderzoeksrapport. Tilburg: STAR-centre. Katholieke Universiteit Brabant.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs. Naar procesgerichte instructie in zelfstandig denken*. Amsterdam: Swets & Zeitlinger.
- Vermunt, J.D.H.M., & Rijswijk, F. van (1997). *Inventaris Leerstijlen HO*. Tilburg: STAR-centre/ Katholieke Universiteit Brabant.
- Wigfield, A. Eccles, J.S., Pintrich, P.R. (1996). Development between the ages of 11 and 25. In: D.C. Berliner, R.C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 148-185). New York: Simon and Schuster MacMillan.

NOOT

1. Het hier beschreven onderzoek maakt deel uit van het onderzoeksprogramma "Development of Learning Abilities" van de Onderwijs- en Opleidingspsychologie van de KUB.
2. Het is statistisch gezien meer adequaat een factor-analyse uit te voeren op de afzonderlijke items, maar daarvoor is onze respondenten-groep te klein van omvang. Schommer voert zelf overigens ook factor-analyses op schaal-niveau uit.

Notities en Commentaren

VERANDERDEN LEERLINGAANTALLEN IN HET VOORTGEZET ONDERWIJS IN HET SCHOOLJAAR 1998-1999 DOOR DE PUBLICATIE VAN INSPECTIEGEGEVENS EN DE BEREKENING VAN HET SCHOOLCIJFER DOOR TROUW IN OKTOBER 1997? EEN NADERE ANALYSE.

J. Dronkers¹

De eerste *Trouw*-publicatie over de inspectiegegevens dateert van oktober 1997, gevolgd door een tweede *Trouw*-publicatie in juni 1998, de *kwaliteitskaart* van de inspectie in september 1998 en de *Keuzegids* in januari 1999. Een van de bezwaren tegen het publiceren door *Trouw* of inspectie van gegevens over het intern rendement van scholen was dat bekendheid met deze gegevens zou leiden tot verschuivingen van leerlingen tussen scholen. Met name zou de publicatie van de sociale compositie van scholen leiden tot onterechte verschuivingen in leerlingaantallen. Een ander bezwaar was dat het berekenen en publiceren van een samenvattend schoolcijfer voor het intern rendement van scholen tot dramatische verschuivingen in leerlingaantallen tussen scholen zou leiden (Veenstra, Dijkstra, Peschar & Sniijders 1998: 131-132). Hoewel deze bezwaren vaak onder verwijzing naar buitenlandse ervaringen (Engeland, Frankrijk) genoemd werden, is voor Nederland nog niet vastgesteld of dergelijke bezwaren terecht zijn. Op grond van persberichten, waarin voorlichtinggevend directeuren en rectoren aan het woord kwamen, zou men verwachten dat deze effecten gering zijn: "ouders vroegen zelden naar de *kwaliteitskaart*". Op grond van de hoge verkoopcijfers van de verschillende *Trouw*-publicaties zou men verwachten dat de effecten groot moeten zijn, daar er blijkbaar veel behoefte bestaat aan de gepubliceerde informatie.

De vraagstelling van deze notitie luidt derhalve: "Is er een aantoonbare samenhang tussen het publiceren van inspectiegegevens en het berekenen van een samenvattend schoolcijfer door *Trouw* in oktober 1997 en veranderingen in de aantallen eerste klas leerling van het daaropvolgende cursusjaar?"

Onder de aannemelijke veronderstelling dat ouders hun kinderen niet tussentijds naar een andere school sturen op grond van omstreden krantenpublicaties, betekent dit dat het mogelijke effect van het publiceren van de inspectiegegevens pas voor het eerst zichtbaar wordt in de leerlingaantallen van de eerste klas voor het cursusjaar 1998-1999. Pas bij de schoolkeuze van dat cursusjaar konden ouders rekening houden met de *Trouw*-publicaties.

De leerlingaantallen zijn door het Ministerie van OC&W ter beschikking gesteld aan *Trouw*. Van deze krant heb ik vervolgens deze aantallen gekregen om de samenhang nader te analyseren. Deze gegevens van de aantallen eerste klassers kunnen helaas niet uitgesplitst worden naar schooltype (vbo, mavo, havo, vwo) of naar locatie. De aantallen zijn omgezet in percentages groei ten opzichte van het voorafgaande schooljaar. Omdat de leerlingaantallen alleen per scholengemeenschap beschikbaar zijn, gebruik ik in deze analyse het gemiddelde schoolcijfer dat een scholengemeenschap (met al zijn verschillende onderwijstypen) in oktober 1997 van *Trouw* kreeg, als samenvattend oordeel over de toegevoegde waarde van die scholengemeenschap ten aanzien zijn intern rendement. Deze schoolcijfers varieerden van 1 tot 10 met een 6 als gemiddelde. Omdat het hier gaat om het effect van publicatie van deze schoolcijfers, gebruik ik de getallen zoals die toen zijn gepubliceerd (dus bijvoorbeeld niet het zogenaamd gecorrigeerde cijfer van de Da Vinci scholengemeenschap in Leiden). De middeling van schoolcijfers en leerlingaantallen op het niveau van scholengemeenschappen in plaats van locatie betekent dat de samenhang tussen schoolcijfers en leerlingaantallen eigenlijk niet op het meest aangewezen niveau wordt vastgesteld. Het betere niveau van analyse zou de

afzonderlijke locatie zijn. Belangrijke verschillen in schoolcijfers en leerlingaantallen binnen scholengemeenschappen verdwijnen door deze onvermijdelijke middeling. Daardoor zal de waargenomen samenhang waarschijnlijk lager zijn dan de werkelijke samenhang. De uitkomsten moeten dus als een conservatieve schatting van het effect van de publicatie van de inspectiecijfers door *Trouw* gezien worden.

De samenhang tussen de percentages groei in leerlingaantal in de eerste klas in een bepaald cursusjaar in verhouding tot het voorgaande jaar en de *Trouw* schoolcijfers wordt in tabel 1 in vijf stappen vergeleken. Allereerst wordt de samenhang voor alle scholengemeenschappen gezamenlijk geanalyseerd.

In de tweede stap gaat het alleen om de scholengemeenschappen met eenzelfde schoolcijfer voor alle onderwijstypen daarbinnen. Homogene scholen zijn óf scholengemeenschappen waarvan de schoolcijfers van de verschillende schooltypen binnen die scholengemeenschappen hetzelfde waren óf scholen die slechts 1 schooltype omvatten. Het gaat dus om scholen die door de *Trouw*-publicatie een ondubbelzinnig imago kregen toegeschreven. Heterogene scholen zijn in deze tweede stap buiten de analyse gelaten.

In de derde stap gaat het om de concurrentieverhoudingen tussen scholen in dezelfde gemeente. Bosker en Scheerens (1999) betoogden dat de publicatie van de inspectiegegevens door *Trouw* ouders weinig zou helpen omdat die ouders maar weinig keuzemogelijkheden binnen hun woongemeenten zouden hebben. Als deze bewering juist is, zou de samenhang tussen schoolcijfer en leerlingaantal hoger moeten zijn in die gemeenten waar meer keuzemogelijkheden bestaan en scholen onderling verschillen in schoolcijfers. Want als scholengemeenschappen binnen een zelfde gemeente weinig onderling verschillen in hun schoolcijfers zijn sterke veranderingen in leerlingaantallen door de *Trouw*-publicatie onwaarschijnlijk. Hetzelfde geldt als in een gemeente slechts één scholengemeenschap staat: ouders hebben in dat geval minder keuzemogelijkheden. Om het belang van de concurrentieverhoudingen tussen scholen binnen dezelfde gemeente te toetsen berekende ik het relatieve schoolcijfer. Dat relatieve schoolcijfer is het verschil tussen het schoolcijfer van een scholengemeenschap en het gemiddelde schoolcijfer van alle scholen in de gemeente van die school. Scholengemeenschappen, die binnen hun gemeente geen andere scholen naast zich hebben, zijn in deze derde stap buiten de analyse gelaten.

In de vierde stap gaat het ook om het relatieve schoolcijfer, maar nu alleen van de homogene scholen.

In de laatste stap ga ik voor alle scholengemeenschappen na of veranderingen in het percentage eerste klas leerlingen samenhangen met de sociale compositie van die scholengemeenschappen. De *Trouw*-publicatie maakte immers ook die compositie openbaar. Het percentage allochtone leerlingen is daarom aan dezelfde *Trouw*-publicatie ontleend en heeft betrekking op alle scholen. Indien ouders zich vooral laten leiden door de sociale compositie van scholen in plaats door de kwaliteit van scholen, die blijkt uit de toegevoegde waarde, moet de samenhang van dit kengetal met de veranderingen in het aantal eerste klas leerlingen hoger zijn dan de samenhang tussen het schoolcijfer en de veranderingen in het aantal eerste klas leerlingen.

De eerste rij van tabel 1 op de volgende pagina laat zien dat het percentage groei van het aantal eerste klasleerlingen in 1998-1999 ten opzichte van 1997-1998 positief samenhangt met het schoolcijfer. Dit is vooral duidelijk voor de homogene scholengemeenschappen, d.w.z. voor scholengemeenschappen waarvan de onderwijstypen hetzelfde schoolcijfer kregen. De samenhang is echter bij lange na niet perfect, want hij komt niet hoger dan .15. Dat betekent dat maar 2% van alle waargenomen verschillen in percentage groei van het aantal leerlingen verklaard kan worden uit het schoolcijfer. Het effect van de *Trouw*-publicatie is dus wel significant maar niet dramatisch.² Overigens is de hier gerapporteerde samenhang stevig te noemen in vergelijking met de gevonden effecten van andere belangrijke onderwijsmaatregelen

(OVb, basisvorming). De afwezigheid van samenhang tussen verandering in leerlingaantal en het percentage allochtone leerlingen laat zien dat ouders bij hun schoolkeuze zich niet hebben laten leiden door de gepubliceerde sociale compositie van een school maar wel door de kwaliteit van een school, zoals die blijkt uit het schoolcijfer.

Tabel 1: De bivariate en partiële samenhangen tussen de percentuele veranderingen in het aantal leerlingen in de eerste klas ten opzichte van het voorafgaande schooljaar en het schoolcijfer en het percentage allochtonen leerlingen

% verandering aantal leerlingen 1ste klas t.o.v. voorgaand jaar	1. Alle scholen	2. Homogene scholen	3. Relatief schoolcijfer voor gemeenten met meerdere scholen	4. Relatief schoolcijfer voor homogene scholen	5. % allochtone leerling van alle scholen
98-99	.07#	.14*	.07	.15*	-.02
97-98	.00	.00	-.03	.00	-.03
98-99, gecor. voor 97-98	.07#	.14*	.07	.15*	-.02
N	602	230	492	196	602

Legenda: ** $p < .01$; * $p < .05$; # $p < .10$

De tweede rij van de tabel laat zien dat het schoolcijfer niet of negatief samenhangt met veranderingen in leerlingaantallen in 97-98. Met andere woorden, de positieve samenhang voor het eerste schooljaar na de *Trouw*-publicatie kan niet verklaard worden uit het feit dat kwalitatief goede scholen reeds in het eerder schooljaar groeiden. Als wij aannemen dat de groei van het aantal eerste klas leerlingen in het cursusjaar 1998-1999 niet uitzonderlijk was, dan heeft de publicatie van de inspectiecijfers door *Trouw* en de vaststelling van het schoolcijfer dus een significante verschuiving in de leerlingaantallen tussen de scholen teweeg gebracht.

In de derde rij toets ik de significantie van de samenhang tussen het percentage groei van het aantal eerste klasleerlingen in 1998-1999 ten opzichte van 1997-1998 en het schoolcijfer door te controleren voor de percentage groei in 1997-1998 ten opzichte van 1996-1997. Deze controle levert geen afwijkend resultaat.

Het relatieve schoolcijfer vertoont geen hoger samenhang dan het absolute schoolcijfer (vergelijk in de tabel kolom 1 met 3 en kolom 2 met 4). Dat wijst er op dat ouders zich weinig aan de gemeentegrenzen bij hun schoolkeuze gelegen laten liggen en dat Bosker en Scheerens op dit punt ongelijk hadden. Hier niet getoonde analyses laat zien dat de samenhang tussen schoolcijfers en groei van aantal eerste klassers bij bijna alle onderwijsrichtingen (openbaar, katholiek, protestant) voorkwam. Ook blijkt uit een andere analyse dat deze samenhang zich voordeed bij witte en zwarte scholen.

Referenties

- Bosker, R. J. en J. Scheerens, 1999. "Openbare prestatiegegevens van scholen; nuttigheid en validiteit." *Pedagogische Studiën* 76:61-73.
- Veenstra, D. R., A. B. Dijkstra, J. L. Peschar en T. A. B. Snijders, 1998. "Scholen op rapport. Een reactie op het *Trouw*-onderzoek naar schoolprestaties." *Pedagogische Studiën* 75:121-134.

Noten

1. Adres: SCO-Kohnstamm Instituut van Amsterdam, Postbus 94208, 1090 GE Amsterdam. Voor de goede orde wijs ik er op dat ik betrokken was bij de hier genoemde *Trouw*-publicaties door het berekenen van het schoolcijfer.

2. Ten einde na te gaan of de correlaties in de eerste rij van de tabel niet ontstaan zijn door kunstmatige hoge groeipercentages ten gevolge van scholenfusies met behoud van een oud BRIN-nummer heb ik in een afzonderlijke analyse alle scholen met groeipercentages groter dan +51.4% verwijderd. De correlaties in de eerste 4 cellen van de eerste rij van tabel 1 zijn dan aanzienlijk hoger: .17, .19, .20, .26. Indien wij ook de scholen met minder dan 51 leerlingen zouden verwijderen (met de onderstelling dat het hier om marginale scholen zou gaan) worden de correlaties: .18, .22, .18 en .23. Omdat ik niet weet ik deze twee veronderstellingen juist zijn en ik mijzelf niet wil rijk rekenen, gebruik ik de analyses toch alle scholen.

DIMENSIES, DILEMMA'S EN DIDACTIEK

Aantekeningen bij het *Eindrapport MOOJ-onderzoek*

F. Goffree

Heuvel-Panhuizen, M. van den en H.J. Vermeer (1999).

Verschillen tussen meisjes en jongens bij het vak rekenen-wiskunde op de basisschool. Eindrapport MOOJ-onderzoek. CD- β Press/Freudenthal instituut: Utrecht

1 Vooraf

Dat leerlingen op vele punten verschillen, wekt in onderwijsland al lang geen verbazing meer op. De aandacht voor differentiatieproblematiek, zorgverbreding en vrij recentelijk voor adaptief onderwijs laat zien dat men ook aan die verschillen niet zomaar voorbij wil gaan (Boland, 1996). Men ziet ook een ontwikkeling in de benadering van de verschillen. Bij het differentiëren dacht men vooral aan het onderscheiden van verschillende (niveau-)groepen, bij zorgverbreding ging het om het erkennen van en tegemoetkomen aan de verschillen, met het perspectief om in de hele groep juist de heterogeniteit te exploiteren. In feite is dat perspectief adaptief onderwijs. In adaptief onderwijs wordt aandacht gevraagd voor goed onderwijs over de hele linie, waarin ingespeeld wordt op de verschillen zodat men preventief werkt om het ontstaan van achterblijvers en afhakers te voorkomen. De laatste tijd mag daarbij ook weer gedacht worden aan (hoog-)begaafde leerlingen. Men spreekt in termen van onderwijs op maat, hetgeen in kringen van reken-wiskundendidactici, waar leren als sociale activiteit hoog in het vaandel staat, heeft geleid tot de gedachte aan neo-klassikaal onderwijs (Treffers, 1997). Dit betekent samen werken aan reken-wiskundige problemen, met respect voor ieders inbreng en aandacht voor de verschillende oplossingsaanpakken. Dit staat in schrille tegenstelling tot de traditionele klassikale aanpak, waarbij het voorkwam dat door het optreden van de leraar (zich vaak niet bewust hiervan) bestaande verschillen juist versterkt werden.

In het MOOJ-onderzoek wordt nagegaan of meisjes en jongens in de rekenen-wiskunde verschillen. Dit 'verschillen' heeft twee kanten. In de eerste plaats kan men op de prestaties van de leerlingen afgaan. Dat betreft de output van het reken-wiskundeonderwijs. Maar men kan aan de andere kant ook de input onderzoeken. In dat geval let men op de wijze waarop de leerlingen bezig zijn in het reken-wiskundeonderwijs. Kort gezegd: presteren jongens en meisjes verschillend in het RW-onderwijs en zijn er verschillen tussen jongens en meisjes zichtbaar tijdens het RW-onderwijs?

In het MOOJ-onderzoek zijn output- en inputverschillen in twee fasen onderzocht. De in fase 1 geconstateerde outputverschillen leidden tot de tweede fase. Maar omdat de onderzoekers tevens geïnteresseerd zijn aan het ontwikkelwerk ten behoeve van de verbetering van het RW-onderwijs op de basisschool, bleef het niet bij een constatering van verschillen in fase 2. In toenemende mate ging men op het microniveau van de schoolklassen op zoek naar mecha-

nismen, die op de leerkansen van meisjes (en jongens) van invloed kunnen zijn. Zo kon de opbrengst van MOOJ ook een geloofwaardige vertaling krijgen naar aanbevelingen voor een RW-onderwijs, waarin meisjes de leerkansen krijgen, die hen mogelijk voorheen onthouden werden. Het rapport eindigt zelfs met een voorstel voor een nascholingsmoduul.

De nu volgende bespreking van het rapport begint met de opbrengst van het onderzoek. Vervolgens wordt getracht een beeld te schetsen van het onderzoeksproces. Hierin worden de dimensies van het fenomeen 'verschillen tussen meisjes en jongens in het RW-onderwijs op de basisschool' steeds genuanceerder naar voren gebracht en komen de dilemma's van de onderzoekers in beeld. De didactiek van het RW-onderwijs vormt voortdurend de context waarin de genoemde dimensies en dilemma's begrepen dienen te worden.

Tenslotte worden enkele vragen over het onderzoek opgeworpen en kan misschien de vraag beantwoord worden: hoe mooi was MOOJ?

2 De opbrengsten van het MOOJ-onderzoek

De aanleiding voor dit onderzoek werd gevonden in de opbrengsten van twee periodieke peilingen van het onderwijsniveau in Nederlandse basisscholen (PPON 1987 en 1992), waarin sekse-specifieke prestatieverschillen werden geconstateerd. Beperken we ons even tot de scores van groep acht, dan blijken de jongens het op bijna alle 27 (1987)/28 (1992) (vaardigheids)schalen beter dan de meisjes te doen en slechts op 21 ervan significant. De meisjes scoren hoger dan de jongens alleen op de schalen voor cijferen (optellen, aftrekken, vermenigvuldigen en delen), op delen significant hoger. Dit is verontrustend, temeer daar de noodzaak van cijferen in de samenleving van vandaag in twijfel getrokken kan worden. Staatssecretaris Netelenbos deelde die zorg en verstrekte een opdracht en MOOJ kon, anno 1995, van start gaan.

In de eerste fase trachtte men een bevestiging van het voorgaande te vinden in de beschikbare data van de CITO Eindtoetsen Basisonderwijs uit de jaren 1993-1995. Hieraan doet 70 procent van de basisscholen mee, hetgeen neerkomt op 100 000 leerlingen van ca 5000 scholen. Op deze data werden secundaire analyses toegepast vanuit drie invalshoeken: analyse op leerlingniveau, op opgavenniveau en op schoolniveau.

Globaal kan men stellen dat de aanwezigheid van sekse-specifieke prestatieverschillen (leerlingniveau) bevestigd werd, dat er in de opgaven onderscheid gemaakt kon worden tussen j-opgaven (waarop de jongens beter scoorden dan de meisjes) en m-opgaven. Op de karakteristiek van beide soorten opgaven komen we terug bij de bespreking van 'dimensies'. Op schoolniveau kwam men tot een wat meer bescheiden resultaat omdat een in de data gesuggereerd onderscheid tussen j-scholen en m-scholen niet significant stabiel bleek te zijn over de (drie) jaren heen.

Met deze gegevens en de inspiratie opgedaan in een literatuurstudie, ging men de tweede fase van het onderzoek in. Daar werden in de eerste plaats 14 scholen (7 j-scholen en 7 m-scholen) onderzocht op de variabelen 'leerlingenprestaties', 'motivatie', 'leerkracht' en 'school'. De onderzoeks-instrumenten werden behalve de drie motivatie-vragenlijsten, zelf ontwikkeld en bestonden uit een MOOJ-reken-wiskundetoets en vragenlijsten voor leerlingen (motivatie), leraren (professionele inzichten en opvattingen) en scholen (algemene kenmerken).

Het tweede deel van fase 2 werd besteed aan observaties in de groepen 8 van vier speciaal in het eerste deel van fase 2 geselecteerde scholen (2 m-scholen en 2 j-scholen). Men spreekt hier van een casestudie. Er zijn dan twee onderzoeksteams aan de slag. Het Utrechtse team (Fi) bestaat uit 4 vakdidactisch onderlegde onderzoekers, die gericht zijn op het in kaart brengen van de leersituaties, met bijzondere aandacht voor de verschillen tussen leerkansen die jongens en meisjes in de reken-wiskundeles krijgen aangeboden. Ze maken geen gebruik van een observatie-instrument, wat een open benadering en mogelijk nieuwe dimensies zou kun-

nen blokkeren. Er is wel een conceptueel kader met aandachtspunten en theoretische ingangen beschikbaar. Naast de 4 Utrechtse onderzoekers is er één observator uit Leiden (Onderwijsstudies), die op basis van een gesloten, computer gestuurd observatie-instrument van Dolle-Willemsen en Rodenburg-Smit (de FROG), de interactiepatronen tussen leraar en leerlingen in kaart tracht te brengen.

Beide onderdelen van fase 2 worden gekenmerkt door een zeer zorgvuldige aanpak van het onderzoek en een heldere en door en door eerlijke beschrijving ervan. Het ontwikkelen van de instrumenten, het afnemen en analyseren van de data, de bespreking van de statistische uitkomsten en de verklaringen van hun tot standkoming, de formulering van de conclusies en de noodzakelijke relativeringen, alles is tot in details te vinden in dit rapport. Bovendien leest het rapport als een spannend boek.

De resultaten van het onderzoek in deel 2 van fase 2 (casestudie in 4 klassen) moeten begrepen worden als verfijningen van de in fase 1 en deel 1 van fase 2 gevonden resultaten, die hiervoor aangeduid zijn.

De groepen, waar meisjes niet minder hoog scoren dan de jongens (m-klassen) kenmerken zich door een veilig klassenklimaat, zowel van sociale aard als van vakinhoudelijke. In het eerste geval hebben de leerlingen respect voor elkaar, heerst er een ordelijke werksfeer en is er sprake van overzichtelijke, sociale regels. De vakinhoudelijke aspecten van het klassenklimaat betreffen onder meer de wijze waarop met oplossingen en met fouten van de leerlingen wordt omgegaan en het geduld dat wordt opgebracht bij het luisteren naar de inbreng van een andere leerling. Het gevolg van een goed klassenklimaat is dat er meer vragen worden gesteld. De m-klassen onderscheidden zich ook in dit opzicht van de j-klassen.

Duidelijkheid over verschillende interactiepatronen in m- en j-klassen, is niet verkregen. In alle vier groepen kregen jongens vaker een beurt bij rekenen en wiskunde en gaven jongens ook meer antwoorden op gestelde vragen. Wel kwam naar voren dat in de m-klassen meer denkpauzes werden ingelast.

De didactiek in de m-klassen is te karakteriseren met de aandacht voor instrumentele uitleg (alleen vertellen hoe je iets moet doen) en in samenhang daarmee de behoefte van leerlingen om louter instrumenteel te begrijpen (snappen hoe je iets moet doen). In die klassen werd ook minder appél gedaan op de eigen inbreng van de leerlingen. In j-klassen vertoonde de uitleg ook vaker didactische tekortkomingen (daarover zijn intrigerende observaties beschreven en men geeft een opmerkelijke verklaring van het feit dat dit meer in het voordeel van jongens zou kunnen werken). In die j-klassen werden onvoldoende leerkansen geschapen.

Het feit dat men door het kiezen van evidente j- en m-klassen genoodzaakt werd om genoeg te nemen met enerzijds klassen die acceptabel maar mechanistisch reken-wiskunde-onderwijs verzorgden, en anderzijds in groepen terecht kwam, waar de realistische reken-wiskundemethode nog niet 'als-bedoeld' was geïmplementeerd, geeft aan deze opbrengsten een bepaalde kleur. Ook ontbreekt daardoor de mogelijkheid om met de opbrengst ver in de toekomst van het (realistische?) reken-wiskunde-onderwijs te kijken.

Toch geven de onderzoekers aan het eind van hun rapport 5 aanbevelingen voor verbetering, die, zoals ze zelf zeggen 'verder gaan dan de MOOJ-onderzoeksgegevens'. Eerst wordt een mogelijke invulling van een '*sociaal veilig klassenklimaat*' gegeven, vervolgens van een '*cognitief veilig*' klimaat. Men beveelt een '*dubbelsporige manier om met fouten om te gaan*' aan (leerlingen mogen fouten maken en er wordt gezamenlijk gezocht naar een verklaring van de gemaakte fouten), noemt een '*actieve leerlingenparticipatie*' en werkt enkele '*vakdidactische en vakinhoudelijke verbeteringen*' in dit kader uit.

3 Dimensies en dilemma's in fase 2

Zoals hiervoor al werd opgemerkt, heeft het fenomeen 'Verschillen tussen meisjes en jongens

in het reken-wiskundeonderwijs op de basisschool' globaal twee dimensies: 'de verschillen in prestaties aan het eind van het basisonderwijs' (D_1) en 'de verschillen in de wijze, waarop het reken-wiskundeonderwijs in de basisschool wordt geconsumeerd' (D_2).

Hoe verder het MOOJ-onderzoek vorderde, des te genuanceerder men deze dimensies kon gaan opvatten. In fase 1 (PPON-gegevens) werd D_1 als het ware in 27 (28) leerstofdomeinen (vaardigheidsschalen) opgedeeld. Bijvoorbeeld $D_1^{\text{score}(1)}$ = 'basisoperaties' en $D_1^{\text{score}(2)}$ = 'gehele getallen'. Op elk van deze dimensies kunnen de prestatieverschillen van jongens en meisjes worden vergeleken. De gezuiverde scoreverschillen (men geeft ze op 5% niveau) dragen bij aan de kennis van het onderzochte fenomeen.

De ook al genoemde analyse op opgavenniveau ging uit van de p-waarden voor jongens en die van meisjes per opgave. De verschillen per opgave dragen eveneens bij aan ons inzicht in het verschijnsel. Zo kon er gesproken worden van m-opgaven en j-opgaven. Bijvoorbeeld $D_1^{\text{m-opgave}(1)}$ = 'met vraag naar bekende standaardprocedure' en $D_1^{\text{m-opgave}(2)}$ = 'waarbij alleen de oplossingsaanpak moet worden aangegeven'.

In fase 2 komt de andere dimensie (D_2) in onderzoek. Voordat men daaraan toekwam dienden evenwel enkele dilemma's het hoofd worden geboden. Het eerste dilemma betrof de constatering dat de onderscheiding in j- en m-scholen niet stabiel bleek te zijn. Wat te doen? Enerzijds kon de constatering (terecht?) genegeerd worden omdat bij de indeling van de scholen met betrekking tot de CITO Eindtoets een grove maat gehanteerd wordt. Anderzijds kon men ook overgaan op m- en j-klassen, die maar een bestaan van 1 jaar kennen. Men gebruikte het scholenonderscheid om de 14 scholen voor fase 2 te selecteren en ging daar vervolgens vooral met de groepen (8) aan de slag.

Een tweede dilemma dient zich aan als in de resultaten van het internationale TIMMS onderzoek in de meeste landen geen significante verschillen tussen de prestaties van jongens en meisjes worden gevonden (en in Nederland wél) maar met de Nederlandse versie van TIMMS nu juist weer geen significante verschillen kunnen worden aangetoond. Stoppen met MOOJ, of een acceptable verklaring vinden. Het laatste wordt gedaan.

Het derde dilemma betreft de maatschappelijke relevantie. Zijn de verschillen tussen allochtone en autochtone leerlingen momenteel niet praktisch relevanter? Ja, maar MOOJ is ook op weg om bij te dragen aan de domeinspecifieke theorievorming, waarin de verschillen tussen jongens en meisjes (nog) niet zijn opgenomen. Toch hiermee doorgaan dus.

Fase 2 begint eigenlijk met een methodologisch dilemma. In plaats van een theorie-toetsend onderzoek, waarbij mogelijke resultaten in de vorm van hypothesen al bekend moeten zijn, wilde men zich hier juist openstellen voor nieuwe elementen van de theorie. Men laat zich inspireren door de onderzoeks-aanpak van Glaser en Strauss uit 1967 en komt tot theorie-generend onderzoek. Deze onderzoekers geven een rationale en ook aanwijzingen voor het kiezen van veelbelovende onderzoeksplekken, zodat men hier ook m-scholen kan selecteren, al vormen die maar een klein deel van de populatie.

Zoals eerder gesteld wordt in deel 1 van fase 2 de aandacht gericht op 14 scholen. Nu wordt dimensie D_2 nader genuanceerd in $D_2^{\text{leerlingen}}$, D_2^{leraren} , D_2^{school} en D_2^{groep} . Ook onderzoekt men de prestaties van de leerlingen van de 14 scholen op een MOOJ-toets, met de bedoeling na te gaan of voor deze scholen de eerder gevonden data (dimensies) ook gelden. (Afspiegelings-onderzoek). Men voegt er ook een dimensie aan toe: strategieonderzoek. Hier probeert men na te gaan of jongens en meisjes verschillen in hun voorkeuren voor bepaalde werkwijzen in rekenen en wiskunde. Dit leidt ertoe dat in de MOOJ-toets expliciet kladblaadjes zijn opgenomen, de leerlingen wordt gevraagd hoe ze aan een antwoord zijn gekomen en er een keuzemogelijkheid voor leerlingen is tussen met name genoemde werkwijzen. (de auteurs spreken hier voort-

durend over 'strategieën', wat mijns inziens reken-wiskundedidactici op een verkeerd been zet).

De toets is hoofdzakelijk samengesteld uit de meest 'extreme' opgaven (ze accentueren verschillen) uit de CITO-eindtoets en enkele opgaven uit de MORE-toets. Bij het nadenken over de vorm van de opgaven moet zich het dilemma van de open/gesloten opgaven hebben aangediend. In lijn met de gegeven opgaven kiest men voor de multiple choice vorm, waarmee een accent wordt gelegd op het *product* van het rekenwerk, terwijl de onderzoekers ook meer willen weten over het *proces*, de wijze van rekenen. Maar onmiddellijk dient zich het volgende dilemma aan. Omdat men toch de rekenwijze van de leerlingen bij het onderzoek wil betrekken, stelt men M.C. vragen waarin de leerlingen voor een bepaalde, uitgewerkte rekenwijze op basis van de eigen voorkeur uit vier mogelijkheden, moet kiezen. Het dilemma voor de onderzoekers bestond dus uit de moeilijkheid om een keuze te maken uit de twee mogelijkheden: 'laat kinderen *kiezen* uit gegeven rekenaanpakken' of 'laat de leerlingen de eigen rekenwijze *tonen* in concrete gevallen'.

De dimensie D_2 ^{leerlingen} krijgt een stevige nuancering als de 'motivatie' van de leerlingen in beschouwing wordt genomen. Eerst onderscheidt men vier indicatoren, die dan vervolgens gemeten worden op de variabelen (1) 'taak-/ego-oriëntatie', (2) 'zelfbeeld/belang', (3) 'succes-falen/inzet-aanleg' en (4) 'opvatting'. De leerlingen moeten vragenlijsten invullen, hen is medegedeeld dat ze anoniem zullen blijven. Zouden de onderzoekers hier ook met een dilemma zijn geconfronteerd? Enerzijds de ethische achtergrond van de toegezegde anonimiteit, anderzijds de onbekende effecten van een dergelijke mededeling op het werk van jonge kinderen.

Ook de leraren van de veertien scholen vullen een vragenlijst in en dragen daarmee bij aan de dimensie D_2 ^{leraren}. De vragen betreffen het hoe en waarom van de wijze waarop zij het reken-wiskundeonderwijs inrichten. In feite wordt hier gezocht naar signalen van realistisch tegenover mechanistisch reken-wiskundeonderwijs. Daarnaast wordt ook gevraagd naar ervaren sekse-verschillen bij de prestaties van leerlingen in het vak rekenen en wiskunde.

Door middel van een schoolvragenlijst (D_2 ^{school}) tracht men zich een algemeen beeld van de school te vormen. Hier wordt ook gevraagd naar de in gebruik zijnde reken-wiskundemethode, en of men daar tevreden mee is. Zo wordt op school 2 de traditionele (vaak als mechanistisch getypeerde) methode Naar Zelfstandig Rekenen gebruikt, naar men zegt niet naar volledige tevredenheid en men ziet de eigen school als 'een gezellig vat vol tegenstrijdigheden'.

De leerlingen-scores op de MOOJ-toets bevestigen opnieuw dat jongens in rekenen en wiskunde beter zijn dan meisjes. De motivatie-vragenlijst levert weinig significante resultaten op, behalve op de sub-schalen 'zelfbeeld' en 'inzet/aanleg'. Voor curriculum-ontwikkelaars, die qq een eclectische instelling hebben, leveren de verkregen data wel indicaties voor een nadere specificatie van het fenomeen motivatie binnen het reken-wiskundeonderwijs. Jongens blijken meer taakgericht en zelfbewust, meisjes zijn ego-georiënteerd en achten zichzelf niet sterk in het vak. Meisjes volgen het liefst strakke regels om tot oplossingen te komen en, jongens schromen er niet voor af en toe eens een gokje te wagen. Meisjes vragen eerder hulp. De laatste uitspraak werd evenwel gedaan door de leerkrachten. De antwoorden op de aan hen gestelde vragen, bedoeld als data om de aard van de beschikbare inzichten, opvattingen en ervaring vast te stellen, laten echter een indruk van het ontbreken van professionaliteit achter. Dit laatste als men tenminste uitspraken die een mechanistische opvatting uitstralen in strijd vindt met opvattingen die een realistische kijk op reken-wiskundeonderwijs laten zien.

Zou het MOOJ-onderzoek aan het eind van deel 1 van fase 2 afgesloten zijn, dan was er weinig meer bereikt dan een enkele harde uitspraak over meisjes in het reken-wiskundeonderwijs en veel aanwijzingen en richtingwijzers om een vervolgonderzoek in te richten. Nu is er (gelukkig) nog een tweede deel van fase 2 waar een en ander concreet kan worden aangepakt.

Uit de veertien scholen worden eerst 2 extreme m- en j-scholen geselecteerd. Het belangrijkste selectie criterium vormen de scores op de MOOJ-toets. Alle andere gegevens uit het eerste deel van fase 2 worden gebruikt voor een finishing touch. Volkomen in lijn met de Glaser-Strauss aanpak kiest men de vier scholen waarvan men verwacht dat zij de meest bruikbare data zullen opleveren. Die data worden verworven door middel van observaties in vier groepen 8. Er zijn vier observatoren uit Utrecht (Freudenthal instituut) en één uit Leiden (Onderwijsstudies van de Leidse Universiteit). In het rapport wordt gesproken over 2 teams, hoogstwaarschijnlijk omdat Utrechters en Leidenaars nogal verschillend denken over de methodologie van klassenobservaties.

Utrecht denkt aan kwalitatief onderzoek, binnen een globaal theoretisch kader, niet al te systematisch geblokkeerd, met een open mind voor nieuwe inzichten en op basis van ervaring en deskundigheid.

Leiden daarentegen acht de systematiek hoog, maakt dan ook gebruik van een gesloten observatiesysteem (de FROG) waarin kwalitatief (in de observatiepunten) en kwantitatief (rekenen met scores) beide aanwezig zijn. Men staat niet open voor impulsen die FROG niet betreffen en op vakinhoudelijke en/of vakdidactische deskundigheid op het gebied van rekenen en wiskunde wordt geen appel gedaan.

De onderzoekers staan hier al weer voor een dilemma. Moeten beide teams op één spoor komen of kunnen de twee verschillende onderzoeksaanpakken als complementair opgevat worden? Let wel, het betreft hier niet een futiel verschil tussen twee opvattingen. Het gaat hier om fundamentele verschillen van mening over waarden en normen in de wereld van onderzoek.

Misschien is het feit dat het Leidse onderzoek slechts significante waarden vond op het gebied van 'denkpauzes' (in m-klassen werd significant meer tijd besteed aan denkpauzes) in het voordeel geweest van MOOJ. Dit resultaat bleek de uitkomsten van de Utrechtse observaties te ondersteunen zodat het nu niet moeilijk was om tot een gezamenlijk rapport te geraken.

Inmiddels hadden de MOOJ-onderzoekers opnieuw twee dilemma's moeten verwerken. Wat te doen in het geval van de absoluut niet volgzone expert, die zich niet als een aio het veld laat insturen met een gesloten observatielijst en waarvan bekend is dat beschikbare theoretische kaders (zoals de interactiekenmerken van Jungwirth) niet wezenlijk onderdeel uitmaken van de expertise. Wat kies je, een volgzone en goedkope aio of een lastige en dure expert?

Het tweede dilemma ontstond bij het lezen van de reflectieve notities als resultaat van de Utrechtse klasseobservaties. Daarin kwamen bepaalde leraren er heel goed af terwijl de professionaliteit van anderen zeer in twijfel werd getrokken, gestaafd door niet mis te verstane voorbeelden. Wat kun je nog verantwoord opschrijven over de gang van zaken in de groepen, is dan de vraag. Aan de andere kant is dit geen 'gewoon' rapport. In verband met de didactische en innovatieve doelstelling van het onderzoek kunnen de reflectieve notities en letterlijke observatieverslagen niet weggelaten worden. Wie een onderzoek zo toespitst op een kwalitatieve afronding, mag de kwalitatieve data namelijk niet vervangen door een globale interpretatie.

Het opnemen van de observatieverslagen en reflectieve notities geeft aan het onderzoek ook nog een maatschappelijke relevantie. Onderzoekers worden erdoor in de gelegenheid gesteld over de schouder mee te kijken naar de analyses. Ontwikkelaars kijken vanuit een bepaalde invalshoek met de expert mee naar hetgeen er in de groepen met hun ideeën gebeurt. Leraren-opleiders kunnen hun theoretisch werk onderbouwen met casussen uit de onderwijspraktijk en nascholers hebben concrete aandachtspunten voor het ontwerpen van nascholingsmodulen. Met dit laatste hebben ook de MOOJ-onderzoekers zich, in het rapport reeds, bezig gehouden.

4 Wat is zo MOOJ?

1. In tegenstelling tot gebruikelijke onderzoeksdesigns, gaan de onderzoekers van MOOJ de weg van grootschalig, kwantitatief onderzoek naar kleinschalig, kwalitatief onderzoek. De oorzaak hiervan is te vinden in het didactisch perspectief en de innovatieve context.
2. Om zich open te kunnen stellen voor nieuwe bijdragen aan de theorie van het reken-wiskundeonderwijs, koos men een methode van onderzoek, gericht op de generering van theorie. Dit leidde in het bijzonder tot een andere methode van 'sampling', beschreven in het werk van Glaser-Strauss (1967).
3. In plaats van jonge volgzaam onderzoekers, voorzien van 'researcher-proof' instrumenten, werd gebruik gemaakt van dure experts in het vakgebied. Wat Adri Smaling opmerkte, 'de kwalitatieve onderzoeker is zijn eigen instrument', paste geheel op het onderzoeksdeel in deel 2 van de tweede fase.
4. Men ging niet alleen aan het werk om verschillen op te sporen, maar in het didactisch perspectief van dit onderzoek trachtte men de mechanismen te achterhalen, die (ongemerkt) aanleidingen vormden voor het ontstaan en/of versterken van de verschillen.
5. Het onderzoeksproces zat vol dilemma's waarvoor een directe oplossing moest worden gevonden. De onderzoekers werden daardoor steeds weer op een ander been gezet.
6. Interessant is ook de vraag aan de vier observatoren van de 4 x 4 reken-wiskundelessen in de 2 j-scholen en 2 m-scholen, om een inschatting te maken van de aard van elk van de vier klassen. En, belangrijker nog, de argumenten te noemen die tot de klassificering hebben geleid. Het ging de onderzoekers om de argumenten, maar in zekere zin werd hiermee ook een (weliswaar zwakke) indicatie voor de interne consistentie van de geïnterpreteerde observaties tot stand gebracht.

5 Vragen die onbesproken bleven en mogelijkheden die niet zijn opgepakt

1. Waarom zijn de leerkrachten niet sekse-specifiek in beschouwing genomen?
2. Waarom is de riantte mogelijkheid op school 12, waar zowel een j-klas als een m-klas was geïdentificeerd (p 93), om specifieke dimensies in een vergelijkend onderzoekje op het spoor te komen, niet benut?
3. Zijn de door Jungwirth gevonden (sekse-specifieke) interactie patronen 'systeem-onafhankelijk'? Of anders geformuleerd, treden in Oostenrijkse schoolklassen gevonden interactiepatronen ook in Nederlandse schoolklassen op?
4. Heeft men nagedacht over de effecten (op het werk) van de mededeling aan jonge kinderen, dat hun antwoorden anoniem bekeken zullen worden?
5. Weet men zeker dat 'reflectie' niet sekse-specifiek is? In de MOOJ-toets worden namelijk opgaven gegeven, die een reflectie op een mogelijk eigen handelen, vereisen. (pag. 54).
6. Kun je op basis van twee klassen met (goed) mechanistisch reken-wiskundeonderwijs en twee klassen met niet geïmplementeerd realistisch reken-wiskundeonderwijs, conclusies trekken voor verbetering van het huidige reken-wiskundeonderwijs? En mag men louter op basis daarvan een nascholingsmoduul ontwerpen? De onderzoekers weten dit laatste zelf ook wel. In paragraaf 9.3 beveelt men in het verlengde hiervan voortgezet onderzoek aan.

Literatuur

- Boland, Th., (1996). *Handen en voeten aan adaptief onderwijs - van idee naar praktijk*. Enschede: Instituut voor de Leerplanontwikkeling (SLO).
- Treffers, A., (1997). Rekenonderwijs naar menselijke maat. Gedifferentieerd en klassikaal: een onjuiste tegenstelling. *Willem Bartjens. Tijdschrift voor reken-wiskundeonderwijs in de basisschool*. Jrg.17, nr.1.
- Smaling, A., (1990). Enige aspecten van kwalitatief onderzoek en het klinisch interview. *Tijdschrift voor nascholings- en onderzoek van het reken-wiskundeonderwijs*, jrg. 8, nr.3 (4-10).

Boekbesprekingen

Jeroen J. G. van Merriënboer (1997)

Training Complex Cognitive Skills. A four-component instructional designmodel for technical training.

Englewood Cliffs, New Jersey: Educational Technology Publications.

ISBN 0-87778-298-9

Om maar met de deur in huis te vallen: van Merriënboer heeft met dit handboek voor het ontwerpen van onderwijs in cognitieve vaardigheden een indrukwekkende prestatie geleverd. Zijn opzet is de presentatie, in hechte onderlinge samenhang, van drie onderwerpen: in de eerste plaats de psychologie van complexe cognitieve vaardigheden, hun aard en ontwikkeling; in de tweede plaats de methoden die ter beschikking staan om deze voor onderwijsdoelinden op een principiële manier te analyseren en ten derde de methodologie van het ontwerpen van daarop gebaseerd onderwijs. Met deze driedelige doelstelling corresponderen de delen A, B en C van het boek. Een belangrijke verdienste is het systematische, breed opgezette overzicht in deel A van de voor het onderwijs in het algemeen belangrijkste onderwerpen uit de cognitieve psychologie. Ik herinner me nog mijn verbazing toen ik voor de eerste keer iemand uit onderwijskundige kring hoorde verklaren dat deze een bepaald 'theoretisch kader had gekozen'. Alsof het om een jas in een winkel ging! Dergelijk relativisme ontbreekt hier gelukkig geheel. Een theorie bestaat uit het geheel van de betrouwbare kennis dat als uitgangspunt voor toepassing of verder onderzoek kan dienen; zulke kennis bestaat gelukkig en te kiezen valt er alleen in zoverre iets, dat men meer of minder in detail kan treden over punten die nog onduidelijk zijn. Uitgangspunt van de theoretische uiteenzetting in deel A is het ACT-model van John Anderson, met o.a. de bekende scherpe onderscheiding tussen declaratieve en procedurele kennis. Het ACT-model vat heel goed samen wat in de cognitieve psychologie met betrekking tot leren en kennisverwerving de hoofdpunten zijn, met een accent op cognitie als activiteit. De schrijver heeft er van afgezien Anderson te volgen in diens latere herzieningen van dit model en dat is verstandig, omdat dit wat de hoofdzaken betreft sinds 1987 niet is veranderd. Van Merriënboer combineert het ACT-model heel effectief met de ideeën van Sweller over de belangrijke rol die cognitieve belasting speelt bij het leren. Centraal in het leermodel dat in deel A wordt gepresenteerd is het punt dat complexe vaardigheden uit verschillende deelvaardigheden bestaan waarvan sommigen door hun herhaalde gebruik routines kunnen worden terwijl andere door hun 'non-recurrent' karakter moeten worden uitgevonden (in het uiterste geval), dan wel uit eerder gevormde schema's worden afgeleid, gaande het probleemoplossingsproces. De wijze waarop deze twee typen deelvaardigheden worden verworven verschilt sterk en dat geldt ook voor verschillende typen declaratieve kennis waarop zij steunen. In vier hoofdstukken wordt de lezer vertrouwd gemaakt met de hedendaagse cognitieve theorie over expertise-ontwikkeling. Dit gebeurt op een didactisch perfecte manier, uitstekend in het gekozen niveau van detail, bewonderenswaardig helder in de uitleg. De salvo's van literatuurverwijzingen waarmee psychologen te vaak hun deskundigheid demonstreren ontbreken geheel, evenals elk ander geleerdheidsvertoon. De schrijver brengt het verhaal hier en daar wat te stellig, dat is een verwijt dat men zou kunnen maken, maar dat beschouw ikzelf niet als hinderlijk. De bedoeling van deel A is om de lezer zich een samenhangend geheel van voor het ontwerpen relevante kennis eigen te doen maken. Mitsen en Maren zijn daarbij eerder storend dan behulpzaam. Deel A is aanbevolen literatuur voor veel meer lezers dan alleen diegenen wiens werk het maken van lesmateriaal voor technische vaardigheden (e.g. het opsporen van mankementen) is. De implicaties van de cognitieve psychologie voor het onderwijs zijn zoals Van Merriënboer laat zien belangrijk en talrijk.

In deel B komen de methoden en technieken aan de orde waarmee de te onderwijzen vaardigheid en de daarbij te gebruiken kennis geanalyseerd en beschreven kunnen worden. Van

M noemt deze analyse 'principled' omdat het gaat om een uiteenleggen in samenwerkende componenten die volgens het theoretische uitgangspunt als principieel verschillend beschouwd moeten worden. Herhaald in te zetten deelvaardigheden die geautomatiseerd kunnen worden en de daarvoor benodigde parate kennis moeten worden onderscheiden van de strategische kennis en de cognitieve schema's die probeemoplossend handelen ondersteunen. Het opstellen van een Gewenst Handelingsverloop of SAP (Systematic Approach to Problem solving) bekroont deze fase. Net als in deel A passeert een groot aantal belangrijke begrippen de revue. De lezer verliest echt zijn onschuld. Om een indruk te geven volgt hier het lijstje van in een willekeurig hoofdstuk behandelde sleutelbegrippen: Fact analysis, feature lists, concept analysis, physical models, plan analysis, goal structures, templates, beacons, building blocks, principle analysis, change relationships, hierarchical analysis, associative analysis, supportive knowledge, misconceptions, naive and intuitive concepts, buggy plans, naive or intuitive plans, misunderstandings, cluster analysis, job aids, helpsystems. Dit lijstje besluit hoofdstuk 8. Dergelijke lijstjes besluiten ieder hoofdstuk. Ze behoren met de zeer goed geschreven 'advanced organizers' en samenvattingen tot de elementen die dit boek behalve als opzoekboek ook heel geschikt maken als leerboek voor 'educationalists' in al die situaties waarin men nu een inleidend boek in de cognitieve psychologie voorschrijft.

Deel C beschrijft de onderwijsmethoden die gebruikt kunnen worden in een te ontwerpen leeromgeving voor complexe cognitieve vaardigheden. Hoofdstuk 11 behandelt methoden om het ontwikkelen van expertise in een leren door te doen omgeving ('whole task practice') te ondersteunen. Leren door te doen, door inductie van probleemoplosschema's kan makkelijk mislukken zonder krachtige ondersteuning in de leeromgeving. De bijdrage van de 'cognitive load theory' is in dit hoofdstuk zeer overtuigend. Herhaald uit te voer taakcomponenten kunnen volgens de analyse in deel A profiteren van deeltaak oefening. De hoofdstukken 12 en 13 behandelen het opzetten van dit soort training en van de manier waarop de voor de deeltaak benodigde kennis kan worden aangeboden zonder dat de voortgang van de activiteit hindert: 'just in time information presentation'. Het volgende hoofdstuk behandelt methoden om de lerende te helpen te elaboreren en om begrip te bevorderen. De schrijver toont hier een lichte voorkeur voor de constructivistische notie dat kennis die men zelf 'gemaakt' heeft, bijvoorbeeld door 'guided discovery' op de een of ander manier van betere kwaliteit is dan kennis die...en dan wordt het meestal te moeilijk en laat men zich daarom vaak verleiden tot een karikatuur van de tegenpartij. Van Merriënboer is gelukkig ook hier pragmatisch en laat zich niet tot absolutismen verleiden. Deel C wordt besloten met een hoofdstuk over het tot een geheel smeden van de geselecteerde instructiemethoden en behandelt praktijkvoorbeelden van leeromgevingen die volgens de principes van dit handboek zijn opgezet. De boodschap is hier weer duidelijk pragmatisch: het werkt dus kan de theorie niet helemaal ondeugdelijk zijn. In het laatste hoofdstuk wordt het in dit boek gepresenteerde model voor het ontwerpen van onderwijs in complexe vaardigheden - de schrijver noemt dat het 4C/ID-model - geëvalueerd en met andere benaderingen vergeleken. De schrijver zegt er toe te neigen zijn model op te vatten als een voorbeeld van 'moderate constructivism'. Het boek besluit met een uitgebreide begrippenlijst waarin men bijvoorbeeld kan opzoeken wat 'moderate constructivism' behelst (dat is niet zoveel) en met twee behulpzame registers.

Het zal duidelijk zijn dat ik 'Training Complex Cognitive Skills, a four-component instructional design model for technical training' als een geslaagde tour de force beschouw. Er valt bijzonder veel uit op te steken. Laten we hopen dat ook de chauvinistische Amerikaanse onderwijsmarkt er zo over denkt. Aan het Engels zal dat niet liggen. Laat U niet ontmoedigen door het feit dat deze aanpak is ontwikkeld voor technische trainingen, bijvoorbeeld in de industrie. Evenmin door het uitgangspunt dat de doelgroep het HAVO 4 -niveau geacht wordt ver te zijn ontstegen. Doelgerichte instructie - en daar gaat het hier natuurlijk om, de modieuze flirt met het constructivisme ten spijt - die aansluit bij wat over cognitief leren en probleemoplossen bekend is komt nu eenmaal het beste tot zijn recht wanneer motivatie geen probleem meer is.

J. te Nijenhuis

Comparability of test scores for immigrants and majority group members in the Netherlands
 Enschede, Printpartners Ipskamp, 1997, ISBN 90-9011113-1

Inleiding

In het proefschrift van Jan te Nijenhuis wordt verslag gedaan van onderzoek naar de validiteit van tests voor allochtonen in Nederland. De gemiddelde testresultaten van minderheidsgroepen zijn vrijwel altijd lager dan die van de dominante bevolkingsgroep. Internationaal is reeds decennia de vraag gesteld, zeker in de Verenigde Staten, of de lagere scores van minderheidsgroepen werkelijk een weerspiegeling zijn van te meten vaardigheid of dat ze een artefact zijn van de meetprocedure. Als dit laatste het geval is dan zou de lage maatschappelijke positie die minderheidsgroepen doorgaans innemen mogelijk deels verklaard kunnen worden door het gebruik van voor hen ongeschikte evaluatieprocedures. In dit soort onderzoek worden zowel schoolvorderingstoetsen als tests op hun validiteit voor subgroepen beoordeeld.

Het onderzoek naar de validiteit van tests kan betrekking hebben op de vraag of het construct dat de test beoogt te meten voor bijvoorbeeld allochtonen en autochtonen dezelfde betekenis heeft (constructvaliditeit), maar het kan ook gaan om de vraag of de test de positie van allochtonen en autochtonen op een extern criterium even trefzeker schat (predictieve validiteit). Het onderzoek naar constructvaliditeit voor subgroepen kent eigenlijk drie fasen. In de eerste fase wordt vastgesteld of de test bij twee subgroepen dezelfde latente trek meet, in de tweede fase wordt per testitem nagegaan of testdeelnemers uit onderscheiden subgroepen maar met dezelfde vaardigheid op de latente trek een gelijke kans hebben om een bepaald item goed te beantwoorden. Het soort onderzoek uit de tweede fase wordt meestal met de term 'Differential Item Functioning' of kortweg 'DIF' aangeduid. Tot slot kan in de derde fase vastgesteld worden of er bij een item sprake is van bias door te achterhalen welk element van een item de DIF veroorzaakt.

Opzet van het onderzoek

Jan te Nijenhuis doet in zijn dissertatie in hoofdstuk twee verslag van onderzoek naar de constructvaliditeit van de General Aptitude Test Battery (GATB, een intelligentietest die gericht is op de algemene intelligentie) voor verschillende groepen allochtonen en autochtonen. Hij gaat na hoe de gemiddelde scores van autochtonen en een aantal groepen allochtonen op de acht subtests zich tot elkaar verhouden, zo ook voor de betrouwbaarheid, de correlatiematrix en de factoroplossingen. De items uit drie subtests worden met de Mantel-Haenszel-procedure onderzocht op DIF en er wordt onderzocht in welke mate het verschil tussen de gemiddelde score van allochtonen en autochtonen te verklaren is door één dimensie (algemene intelligente of g) in de subtestscores.

Vervolgens wordt in hoofdstuk drie voor een aantal groepen allochtonen de constructvaliditeit van een aantal persoonlijkheidstests vergeleken met die voor autochtonen. De persoonlijkheidstests zijn de Amsterdamse Biografische Vragenlijst (ABV) en de ICIP Rigiditeits Test (IRT). Naast het vergelijken van de gemiddelde scores en de betrouwbaarheid van de subtests voor verschillende groepen allochtonen en autochtonen, worden ook de factoroplossingen van allochtone groepen onderling vergeleken. Tevens wordt naar kwalitatieve en kwantitatieve DIF gezocht. Bij een item is sprake van kwalitatieve DIF wanneer de factorladingen van dat item bij subgroepen niet vergelijkbaar zijn. Er is kwantitatieve DIF wanneer van twee groepen met dezelfde positie op een latente trek er een groep is die een systematische hogere of lagere score op een bepaald item behaalt vergeleken met de andere groep. Omdat bij de persoonlijkheidstests de antwoorden polytoom zijn, wordt hier DIF gerapporteerd met een partiële correlatie index.

In hoofdstuk vier geeft de auteur aan hoe hij de constructvaliditeit van een aantal tests, die de geschiktheid van personen meten om zich ten opzichte van anderen (bijvoorbeeld: in het

verkeer) veilig te gedragen, voor allochtonen en autochtonen heeft onderzocht. Centraal staat bij deze tests de attitude om risico's te vermijden in de praktische situaties, waarin functionarissen bij de Nederlandse Spoorwegen en busmaatschappijen functioneren. Bij de tests is vastgesteld in hoeverre de gemiddelde resultaten van de onderscheiden groepen verschillen en er is gekeken naar de vergelijkbaarheid van de factoroplossingen. Verder is onderzocht in welke mate de verschillen tussen de gemiddelden van de subgroepen verklaard kunnen worden door één algemene factor of dat de taal- en/of rekenscores van de proefpersonen ook een bijdrage leveren.

Naast de reeds langer bestaande, hierboven beschreven, 'veiliggedrag'-tests bestaat er sinds 1994 een versie die met behulp van de computer afgenomen wordt. In hoofdstuk vijf gaat de auteur in op de constructvaliditeit van deze tests voor allochtonen en autochtonen. De methode die gebruikt wordt om de validiteit te vergelijken is voor een belangrijk deel vergelijkbaar met de methode die bij de traditionele test gebruikt is. Bovendien wordt nagegaan in welke mate de gemiddelde verschillen tussen de onderscheiden subgroepen verklaard kunnen worden door de *g*-factor. Verder wordt onderzocht of de traditionele versie van deze test vergelijkbaar is met de computerversie.

In het zesde hoofdstuk krijgen we een verslag van onderzoek naar de differentiële predictie van een aantal tests. Nagegaan wordt in welke mate de regressielijnen van predictoren op een extern criterium van allochtonen en autochtonen verschillen. De eerder beschreven GATB-intelligentietest, de persoonlijkheidstests en de 'veiliggedrag'-tests worden in dit onderzoek gebruikt als predictoren. Daarnaast wordt het oordeel dat een psycholoog zich over de kandidaat tijdens een interview heeft gevormd, ook gebruikt als predictor. Als extern criterium worden de resultaten gebruikt die de deelnemers bij een opleiding tot vrachtwagenchauffeur behaalden.

Resultaten

De uitkomsten van de door de auteur uitgevoerde analyses tenderen vrijwel allemaal in dezelfde richting. De resultaten die de allochtonen op de verschillende tests halen, zijn altijd minder gunstig dan die van autochtonen. Uit de analyses is nauwelijks gebleken dat de constructvaliditeit van de tests voor allochtonen anders is dan voor autochtonen. De meeste subtests meten wat ze beogen te meten en als er bij een item al sprake is van DIF dan is het nadeel dat de scores van allochtonen daarvan ondervinden, klein te noemen. Als het gaat om het voorspellen van de positie van allochtonen en autochtonen op een extern criterium dan blijkt dat de regressielijnen van de beide groepen niet altijd gelijk lopen. Dit is niet in het nadeel van de allochtonen want als in de praktijk dezelfde regressielijn geldt voor beide groepen, hetgeen veelal het geval is, dan werkt dat in het voordeel van allochtonen.

De conclusie van de auteur is dat de onderzochte tests niet perfect zijn, maar dat ze redelijk betrouwbaar zijn voor gebruik onder allochtonen. Het zou een kleine verbetering zijn als de allochtonen voorafgaand aan de testafname meer voorbeelditems zouden kunnen krijgen en als de (Nederlandstalige) instructie voor en tijdens de afname ook voor allochtonen volledig duidelijk zou zijn. Volgens de auteur kan over het algemeen niet gesteld worden dat de lage maatschappelijke positie die veel allochtonen innemen deels toegeschreven kan worden aan het gebruik van voor hen invalide tests. De tests geven aan dat de cognitieve vermogens van allochtonen gemiddeld lager zijn dan die van autochtonen en dat is een goede weergave van de werkelijkheid. Er kan gesteld worden dat juist deze conclusie uit het proefschrift niet ongemerkt aan de media is voorbij gegaan.

Beoordeling

Het proefschrift moet antwoord geven op de vraag of de onderzochte tests even bruikbaar zijn voor allochtonen als voor autochtonen.

Het meest valt op dat niet of nauwelijks melding wordt gemaakt van de inzichten die in binnen- en buitenland zijn verworven met onderzoek naar de validiteit van schoolvorderin-

gentoetsen voor minderheidsgroepen. Het lijkt wel of er op dit punt gescheiden werelden bestaan: tests en toetsen. Juist in de wereld van het analyseren van toetsen heeft het gebruik van de itemresponsstheorie (IRT) furore gemaakt. Als toetsitems met behulp van IRT een eendimensionale schaal vormen, is de schaal inhoudelijk doorgaans goed te duiden. Bovendien is een met IRT geconstrueerde schaal in bepaalde zin ook steekproefonafhankelijk. Het onderzoek naar DIF bij toetsen heeft ook geleerd dat het hanteren van één statistische procedure om DIF te constateren een hachelijk onderneming is en dat kruisvalidatie geboden is. In het onderhavige proefschrift wordt geen gebruik gemaakt van IRT en de intelligentie- en de persoonlijkheidstests worden met slechts één statistische procedure op DIF onderzocht, de 'veiliggedrag'-tests worden in het geheel niet op DIF gecontroleerd.

De auteur worstelt met een vraag die zowel bij onderzoek naar de constructvaliditeit als naar de predictieve validiteit voor subgroepen gesteld moet worden: vertoont de eendimensionale schaal, respectievelijk het extern criterium zelf geen bias? Bij onderzoek naar DIF ga je eerst op zoek naar proefpersonen van twee subgroepen die dezelfde positie op een eendimensionale schaal innemen en vervolgens ga je na of bij een bepaald item de kans op een goed antwoord voor beide subgroepen gelijk is. De assumptie is dat de eendimensionale schaal valide is voor beide groepen, als dit niet het geval is, is het valideringsonderzoek niet zinvol. De auteur is zich bij zijn onderzoek naar DIF bewust van dit probleem. Om er voor te zorgen dat zijn eendimensionale schaal zelf geen items met DIF bevat, verwijdert hij via een iteratieve procedure uit deze schaal de items die DIF bevatten en vervolgens gaat hij met de gezuiverde schaal (eendimensionaal? betrouwbaar?) alle afzonderlijke items op DIF beoordelen. Deze procedure lijkt in technisch opzicht elegant, maar inhoudelijk zitten er wel haken en ogen aan, omdat onduidelijk is welk construct de resterende items representeren. Ook moet opgemerkt worden dat niet is nagegaan of de DIF bij een item uniform is, dat wil zeggen of de DIF bij hoge en lage scoorders in dezelfde mate aanwezig is.

Bij onderzoek naar de differentiële predictie van tests moet je aannemelijk maken dat het extern criterium (in dit geval de opleiding tot vrachtwagenchauffeur) zelf geen bias vertoont voor de subgroepen en inhoudelijk verwant is met de predictoren. De auteur stapt ten onrechte over deze vraag heen. Overigens kun je je afvragen waarom je succes in een opleiding voor vrachtwagenchauffeurs zou willen predicieren met een intelligentietest, persoonlijkheidstests en 'veiliggedrag'-tests als je al weet dat de verwantschap tussen predictoren en extern criterium inhoudelijk moeilijk te onderbouwen is en als je uit de literatuur al weet dat een verklaarde variantie van boven de 10% tot de uitzonderingen behoort.

De auteur vindt dat de onderzochte tests gebruikt kunnen worden voor de beoogde selectiedoelen, ook als het gaat om allochtonen. Toch kan gesteld worden dat het onderzoek zeker op het gebied van DIF diepgaander had moeten zijn om deze conclusie te rechtvaardigen.

*Henny Uiterwijk
sector basis- en speciaal onderwijs
Instituut voor toetsontwikkeling (Cito)*

T.M. Janssen (1998)

Literatuuronderwijs bij benadering; een empirisch onderzoek naar vormgeving en opbrengsten van het literatuuronderwijs Nederlands in de bovenbouw van het havo en vwo. Proefschrift UU. Amsterdam: Thesis Publishers. ISBN 90-5170-446-1/NUGI 724/951. f 42,50

(Voor)oordelen over literatuuronderwijs

Publicaties dit voorjaar in de landelijke pers toonden een forse commotie rond het beoogde literatuuronderwijs in de tweede fase voortgezet onderwijs: los van het vak Nederlands; integratie met andere talen; beperking van het aantal les- en studie-uren, te lezen boeken en de omvang van de literatuurgeschiedenis. Er werden standpunten verwoord als: 'leraren kunnen leerlingen alleen maar literatuurgeschiedenis geven maar ze niet leren lezen', 'leerlingen lezen

alleen nog maar bankbiljetten', 'de kwantiteit van het lezen zegt niets over de kwaliteit', 'leesplezier staat voorop ten koste van cultuuroverdracht', 'het literatuuronderwijs verwordt tot zap-cultuur'. Deze 'discussie' lijkt me niet meer dan het uitwisselen van vooronderstellingen. Nu is dat niets nieuws. Het literatuuronderwijs in Nederland (en in landen om ons heen) staat al honderd jaar ter discussie. Het is eigenlijk één lange retorische discussie, niet in het minst omdat empirische gegevens over literatuuronderwijs schaars zijn.

De laatste tien, vijftien jaar wordt meer en meer onderzoek verricht. Historisch onderzoek verheldert dat er tenminste vier van elkaar te onderscheiden opvattingen over literatuuronderwijs bestaan: historisch/auteursgericht, tekstgericht, contextgericht, lezersgericht. Dat onderzoek wijst ook uit die opvattingen vermoedelijk meer in de retoriek, dan in de praktijk te onderscheiden zijn (Van de Ven 1996). Datzelfde blijkt uit empirisch doelstellingenonderzoek, dat ook aangeeft dat leraren veel belang hechten aan literatuuronderwijs. Enkele case-studies (n=1) geven meer inzicht in de praktijk ervan, maar zonder generaliserende prententie. Daarnaast is er verhoudingsgewijs veel onderzoek naar leesgewoontes en leesmotivatie (voor het meest recente overzicht van empirisch onderzoek zie Hoogveen & Bonset 1998).

Vijf deelonderzoeken

Tegen deze achtergrond bezien presenteert Tanja Janssen een uniek onderzoek. Ze constateert dat historisch onderzoek en doelstellingeninventarisatie uitwijzen dat er in de retoriek vier benaderingen van literatuuronderwijs kunnen worden onderscheiden. In aansluiting daarop zijn haar onderzoeksvragen (in het kort):

- Hoe zien de inhoud en de vormgeving van het huidige literatuuronderwijs Nederlands in de bovenbouw van het havo en vwo er in grote lijnen uit?
- Zijn er in de praktijk van het literatuuronderwijs Nederlands sporen te vinden van de vier benaderingen?

Bij een bevestigend antwoord op vraag b gelden de volgende vragen:

- Hoe krijgen de vier benaderingen vorm in de praktijk van het literatuuronderwijs?
- Wat leren leerlingen van het literatuuronderwijs - zijn er verschillen in de leerresultaten van de vier benaderingen?
- Wat hebben oud-leerlingen van het literatuuronderwijs - onderscheiden naar de benaderingen - geleerd?

Tussen 1990 en 1996 voerde Janssen vijf empirische deelstudies uit. Deelonderzoek 1 betreft een enquête onder docenten Nederlands. De vier bovengenoemde benaderingen vormden het uitgangspunt ervan. De vragenlijst was omvangrijk: 33 pagina's met 62 hoofdvragen en 140 deelvragen. De docenten zijn gekozen in een tweetrapssteekproef: eerst een steekproef van scholen, daarna daarbinnen van docenten. De respons van de scholen bedroeg 69%. Dat leverde een lerarenpopulatie op van 2000, van wie uiteindelijk 1165 docenten konden worden benaderd. Van hen reageerde in totaal 51%. Nonresponsonderzoek wees uit dat veel leraren de vragenlijst te omvangrijk vond. Het gaf ook aan dat de responsgroep representatief was.

De gegevens wijzen uit dat leraren literatuur het belangrijkste onderdeel van hun vak vinden. Ze willen er meer tijd aan besteden dan aan de afzonderlijke taalvaardigheden of aan taalbeschouwing. De statistische analyse (chikwadraatberekening, vaststellen van multivariante en het univariante significantieniveau e.d.) wijst uit dat leraren binnen deze voorkeur nogal van elkaar verschillen. Deze verschillen zijn niet toe te schrijven aan achtergrondkenmerken van docenten (leeftijd, ancienniteit, opleiding, geïnformeertheid). Ze kunnen worden getypeerd met de vier benaderingen. Deelonderzoek 1 mondt uit in vier, wat betreft inhoud (leerstof, teksten) en didactische aanpak (werkvormen, leeractiviteiten) significant van elkaar verschillende groepen docenten. Janssen brengt hen onder in vier profielen van literatuuronderwijs: Culturele vorming (48%), Literair-esthetische vorming (11%), Maatschappelijke vorming (12%) en Individuele ontplooiing (25%), een kleine restgroep van 4% viel niet te categoriseren. Het enquête-onderzoek toont echter ook grote standaarddeviaties die een grote

variatie tussen docenten op vrijwel alle onderzochte kenmerken indiceren: tijdsbesteding, doelstellingen, oordeel over de beginsituatie van leerlingen, tekstaanbod, wijze van werken met teksten, keuze van leerstofinhouden, gehanteerde werkvormen, gebruikte leerboeken, wijzen van evalueren en toetsen. Zo varieert het aantal gelezen en besproken teksten of tekstfragmenten van 0 tot meer dan 100; de omvang van de boekenlijst voor het examen varieert van 4 tot 35 titels. Tegelijkertijd bleken docenten uit verschillende profielen overeen te komen in de gebruikte literatuurmethode. De gekozen methode is dus geen goede graadmeter voor de onderwijspraktijk.

Deelonderzoek 2 gaat na in hoeverre de verschillen tussen de profielen zich in de lespraktijk manifesteren. Uit de vragenlijsten werden per profiel drie representatieve leraren geselecteerd, met vergelijkbare leraarloopbaan (ervaring, satisfactie). Ze werden geïnterviewd over hun praktijk. De verschillen tussen de profielen blijken in deze interviews minder scherp als op grond van de enquête mocht worden verwacht. Het profiel Culturele vorming behoudt de duidelijkste contouren. De andere profielen zijn te zien als afwijkingen daarvan in de richting van literair-esthetische of maatschappelijke vorming, dan wel individuele ontplooiing. Tegelijkertijd constateert Janssen dat de gerapporteerde praktijk zeer divers is. Binnen de profielen blijken de individuele docenten zeer verschillende praktijken te rapporteren.

Deelonderzoek 3 brengt de geobserveerde praktijk van vier docenten (een uit elk drielik geïnterviewden) in beeld. Centraal stonden de taakeisen die docenten hun leerlingen stelden in vragen en opdrachten. Gegevens over deze taakeisen werden geïnventariseerd met behulp van observatieschema's, geconstrueerd op basis van een classificatie van tekstgerichte (genrespecifieke) en niet-tekstgerichte leerstofdomeinen (culturele achtergrondkennis bijv.). Janssen baseert zich in dezen op literair-theoretische en literatuurdidactische publicaties. Daarnaast werden taakeisen geïnclassificeerd in termen van cognitieve taak en fasen in het leerproces. Referentiepunten hier zijn onder meer Guilford en De Corte, naast opnieuw literatuurdidactische gegevens. Janssen slaagt erin een bruikbaar instrumentarium te ontwikkelen, al brengt ze het met de nodige reserve: "Wat lezers met teksten doen, is eerder een zeer ingewikkeld proces waar allerlei activiteiten aan te pas komen, zoals waarnemen, associëren interpreteren en waarderen, en waarbij ook emoties - opgeroepen door nieuwsgierigheid, spanning, inleven, en meeleven - een belangrijke rol spelen" (112). De statistische verwerking van de gegevens laat zien dat de profieldocenten verschillen. Overigens zijn de verschillen in inhoud (tekstkeuze) groter dan die in leerlingengedrag (vereiste cognitieve activiteiten).

Deelonderzoek 4 gaat na in hoeverre de leerresultaten van de vier prototypische docenten verschillen. Per docent zijn van twee klassen learner reports verzameld. Janssen gaat uitvoerig in op de validiteit van het instrument. De learner reports zijn geanalyseerd op de gecombineerde dimensie Inhoud & Gedrag. Na statistische verwerking (chikwadraattoetsen) blijken er enkele grote verschillen in de leerresultaten. Binnen de groepen zijn er overigens weer grote verschillen in het aantal gerapporteerde leerervaringen.

Deelonderzoek 5 be vraagt oudleerlingen (examen zes jaar terug) van de vier leraren naar leerresultaat op langere termijn. Ook in dit deelonderzoek gebruikte learner reports. De resultaten tonen weinig verschillen tussen de vier groepen leerlingen. Ook nu is er sprake van een grote diversiteit aan leerervaringen binnen de groepen. De enige overeenkomst lijkt dat bijna alle leerlingen melding maken van een attitude (positief dan wel negatief) ten opzichte van literatuur als resultaat van het onderwijs. Janssen concludeert, wijzend op mogelijke feilen van haar onderzoek - zo geeft ze aan dat bijvoorbeeld de respons gekleurd kan zijn doordat juist enthousiaste oudleerlingen reageerden - dat de door de docenten gehanteerde vorm van literatuuronderwijs er op den lange duur niet zoveel toe doet.

Vier profielen

Een belangrijk gegeven is dat er, ondanks alle diversiteit in opvatting en praktijk vier profielen van literatuuronderwijs te construeren zijn. Deze verschillen van elkaar in een aantal opzichten.

Enquête en interview wijzen uit dat docenten binnen het profiel **Culturele vorming** meer tijd (dan andere docenten) besteden aan literatuurgeschiedenis, literair-historische achtergrondinformatie en aan poëzie. Ze leggen een sterker accent op doceren (dictaat geven, vertellen). De observatie bevestigt dit beeld. Opvallend is de aandacht voor reproducerende leeractiviteiten. De leerlingen rapporteren significant meer resultaten in termen van declaratieve kennis van literair-historische en niet-literaire achtergrondkennis. Ze noemen meer metacognitieve vaardigheden. Ze tonen meer dan andere leerlingen een positieve attitude ten opzichte van literatuur en literatuuronderwijs, ze hebben niet zoveel problemen met verplicht lezen. Ze rapporteren gemiddeld ook meer metacognitieve leerervaringen betreffende studievaardigheden. De oudleerlingen onderscheiden zich door een groter effect in termen van een positieve attitude.

Het profiel **Literair-esthetische vorming** onderscheidt zich vooral in leerstofinhoud en in mindere mate in didactiek. Docenten in dit profiel besteden meer tijd aan literair-theoretische onderwerpen en aan formele aspecten van teksten. De lesinhoud bestaat voornamelijk uit de structuuranalyse en het bijbehorende begrippenapparaat. Docenten in dit profiel toetsen vaker op een productieve wijze: boek- en/of leesverslag. De interviewgegevens bevestigen dit beeld. De lespraktijk is gevarieerd: zelfstandig werken, doceren, klasgesprekken. Specifieke taakeisen bleken uit de observatie moeilijk te destilleren. Wat opvalt is dat onder meer het reconstrueren van de inhoud van een tekst minder vaak wordt gevraagd. De leerlingen tonen vaker een negatieve attitude en rapporteren minder procedurele en metacognitieve kennis. De oudleerlingen rapporteren significant minder effect in termen van studievaardigheid.

Het profiel **Maatschappelijke vorming** onderscheidt zich vooral in didactisch opzicht. De docenten peilen vaker leesvoorkeuren en leesgewoonten van leerlingen en ze toetsen vaker in het voorindexamenjaar. De leerstof verschilt in die zin van andere profielen dat er meer aandacht is voor niet-literaire, sociaalpolitieke achtergronden.

In de interviews blijken de docenten een tamelijk diffuse groep te vormen. Ze komen overeen in hun aandacht voor niet-literaire achtergronden en het belang dat ze hechten aan kritisch leren oordelen. In de observatie blijkt de tekstkeuze te verschillen van andere profielen: meer teksten van vóór 1945. Voorts blijkt de docent minder vaak dan die uit de andere profielen verbanden te (laten) leggen tussen tekst en context, in tegenstelling tot wat er op basis van enquête en interview mocht worden verwacht.

Leerlingen uit dit profiel onderscheiden zich in geen enkel opzicht significant van andere leerlingen. Opvallend is wel dat er veel positieve én veel negatieve attitudes worden gerapporteerd. De scores ontlopen elkaar nauwelijks (64 resp. 62% van de leerlingen). De oudleerlingen verwoorden minder positieve attitudes.

Docenten met het profiel **Individuele ontplooiing** onderscheiden zich met name in doelstelling. Het profiel lijkt op dat van de Maatschappelijke vorming door de geringere aandacht voor literatuurgeschiedenis en -theorie. Er is meer aandacht voor de belevingswereld van de leerlingen en voor lezersgerichte begrippen. In vergelijking met Maatschappelijke vormers wordt minder belang gehecht aan niet-literaire achtergronden. In de interviews toont de groep zich homogener dan op grond van de enquête mocht worden verwacht. De docenten delen het belang van leesplezier en de persoonlijke leeservaring; ze vertonen wat Janssen noemt 'een zekere experimenteerlust of vernieuwingsdrang'; ze hechten minder aan literatuurgeschiedenis en structuuranalyse. Literatuuronderwijs is voor hen vooral lezen van en praten over teksten. In de observatie blijkt de docent meer tekstgericht dan zijn collega's uit de andere profielen. Vooral korte verhalen worden gebruikt, en leerlingen krijgen evaluatieve en elaboratieve taken. Opvallend is dat evaluatie van de tekst vrijwel steeds het startpunt vormt voor het onderwijs, terwijl dat bij de andere geobserveerde docenten meestal het sluitstuk is. Het lezen geschiedt buiten de les, dit in tegenstelling tot de andere profielen. De leerlingen rapporteren absoluut en gemiddeld minder leerervaringen dan die uit andere profielen. Ze noemen leerresultaten in termen van interpretatievermogen, en vermelden zowel declaratieve als metacognitieve kennis en vaardigheden. Het verschil met andere groepen leerlingen in dezen is

significant. Ze geven minder literair-historische kennis aan. Minder leerlingen rapporteren een negatieve attitude, tegelijkertijd verwoorden ook minder leerlingen een positieve attitude. Wel is het zo dat het percentage leerlingen dat uitsluitend een positieve attitude verwoordt, relatief groot is: 38%. Oudleerlingen rapporteren minder leerresultaten in literatuuranalyse en op het gebied van studievaardigheid.

Conclusie

Janssen concludeert dat 'het' literatuuronderwijs in Nederland niet bestaat. Het enquête-onderzoek toont grote standaarddeviaties die een grote variatie tussen docenten op vrijwel alle onderzochte kenmerken indiceren. Die variatie komt terug in de interviews, de observaties en verslagen van leerlingen en oudleerlingen.

Er zijn op basis van de vragenlijst vier benaderingen te onderscheiden. Die zijn ook nog te herkennen in de gerapporteerde praktijk (interviews), maar al weer veel minder in de geobserveerde praktijk. Janssen constateert dus een zekere spanning tussen retoriek en praktijk. Die spanning verschilt per profiel, en lijkt het kleinst in het profiel Culturele vorming, daarna in dat van de Individuele ontplooiing. Die beide profielen lijken dan ook in de praktijk het duidelijkst vormgegeven en het meest elkaars tegenpolen: veel respectievelijk weinig literatuurgeschiedenis, oude poëzie versus modern proza, doceren versus interactie, niet-tekstgerichte versus tekstgerichte taakeisen, weinig versus veel evaluatieve taken.

De leeropbrengst verschilt in de mate van een positieve dan wel negatieve attitude, en lijkt samen te hangen met het profiel van de docent. Opvallend is dat de literair-historische benadering zich in dezen positief onderscheidt. Janssen suggereert dat de duidelijke herkenbaarheid van dit profiel, ook in de lessituatie, in dezen een belangrijke factor vormt. Opvallend is ook dat de docenten, hoe hun opvatting ook moge zijn, erin slagen hun leerlingen dat bij te brengen dat ze nastreven.

De relatie tussen leeropbrengst en docentprofiel is bij oud-leerlingen moeilijk nog te vinden. Oud-leerlingen rapporteren, evenals de leerlingen, veel en gevarieerde resultaten.

Discussie

Tanja Janssen's vijf deelonderzoeken vormen een trechter. Een breed overzicht van de retoriek mondt uit in een nauwkeurige blik op praktijken van literatuuronderwijs. Een grootschalig enquête-onderzoek mondt uit in een kleinschalig interviewonderzoek, in n=1 observaties en learner reports. Het onderzoek voert de lezer van leraardoelen naar leerlingresultaat. Janssen gaat bij dit alles uiterst zorgvuldig te werk. Ze besteedt veel aandacht aan de instrumentontwikkeling (pretest- en validiteitsstudies), aan de dataverzameling (rekening houden met effecten van observatieparadox), in de statistische verwerking ervan. De transcriptie van de lessen werd zorgvuldig verricht, de lesprotocollen werden geautoriseerd door de leraren, de scores gecontroleerd op inter- en intrabeoordelaarsbetrouwbaarheid. Mutatis mutandis geldt dat voor de learner reports. Toch heb ik bij alle zorgvuldigheid wel enkele methodologische vragen.

Haar design karakteriseert ze zelf als een "empirisch, sociaal-wetenschappelijk onderzoek" (20). Dit type onderzoek plaatst ze dan tegenover "meer interpretatie onderzoeksvormen" (20). Ze verantwoordt haar keuze vanuit een toenemend belang dat literatuurwetenschappers hechten aan empirisch onderzoek. Ik heb hier wat moeite mee. De tegenstelling 'empirisch' versus 'interpretatief' lijkt me onjuist. 'Empirisch' geeft aan wat voor data verzameld worden, is een aanduiding van objectbepaling. 'Interpretatief' duidt een epistemologische en methodologische keuze aan. De juistere tegenstelling lijkt me die tussen een empirisch-objectiverende en een empirisch-interpretatieve traditie (vgl. Hetebrij & Wardekker 1986). Janssen kiest voor onderzoek binnen de objectiverende traditie. Dat levert in haar dissertatie een goed controleerbaar en repliceerbaar onderzoek op, waarin zorgvuldig wordt omgegaan met criteria van validiteit en betrouwbaarheid. Tegelijkertijd levert die benadering ook problemen op. De vraag lijkt gewettigd of het observatie-onderzoek en dat naar leerresultaten niet beter vanuit de interpretatieve traditie hadden kunnen worden opgezet. Toegegeven, de generali-

satie van zulk onderzoek is problematisch, maar vermoedelijk niet meer dan nu al het geval is. Ik werk enkele van mijn gedachten kort uit.

Het observatie-onderzoek beperkt zich tot de taakeisen die docenten stellen aan leerlingen. Andere aspecten van het gegeven onderwijs blijven buiten beschouwing. De praktijk van drie van de vier prototypische docenten blijkt met deze methode te karakteriseren. Niet echter die van de 'Literair-esthetische vormer'. Janssen constateert dan ook dat blijkbaar "de proportie van docentvragen en -opdrachten in dit geval geen goede graadmeter (is) voor de taakeisen die de docent in het literatuuronderwijs stelt" (144). Met een meer empirisch-interpretatieve benadering was dit probleem op te lossen geweest. Een onderzoeker had dan andere aspecten van onderwijsgedrag in beeld kunnen brengen, aspecten die naar de mening van de docent zijn taakeisen aan leerlingen wel representeren. Dat zou dan wel, gezien vanuit Janssens optiek, ten koste gegaan zijn van de onderlinge vergelijkbaarheid van de docenten. Zij wil die vergelijking langs dezelfde parameter (taakeisen). Verschil in docentenpraktijk kan echter ook blijken uit het gegeven, dat die praktijken zich eenvoudigweg niet lenen voor eenzelfde instrument. Bovendien benadert Janssen de taakeisen hier volledig vanuit het perspectief van de docent. Interessant zou het zijn na te gaan of de leerlingperceptie overeenkomst met wat de docent beoogt. Wellicht zien de verschillen er dan wat anders uit.

Eenzelfde redenering kan gelden voor de 'learner reports'. Janssen constateert terecht dat "het invullen van een open leerverslag voor leerlingen een moeilijke taak is" (158). Daarom structureert ze de opdracht. Daarmee worden de leesverslagen vergelijkbaarder, maar tegelijkertijd worden de (oud)leerlingen gestuurd. Spontane herinneringen, meer persoonlijke herinneringen, dat wat naar de mening van de (oud)leerlingen het meest typerend was, krijgt nu minder reliëf doordat het instrument de herinneringen genereert en er wellicht meer en/of andere oproept dan die waar leerlingen zelf mee zouden komen. De vraag hierachter is natuurlijk wat je beschouwt als leerresultaat: dat wat leerlingen uit zichzelf rapporteren of dat wat door de vragen van de onderzoeker wordt opgeroepen.

De analytische lijn van Janssen speelt haar ook parten bij de observatie van de docent Maatschappelijke vorming. De observatie leidt niet tot duidelijke karakteristieken. Maar hier is sprake van een 'ik zou wel willen maar mijn sectie': de docent in kwestie geeft aan dat haar eigen praktijk in de geobserveerde lessen niet representatief is voor haar opvattingen. De rol van de sectie, als factor achter de retoriek-praktijkproblematiek wordt voor de andere docenten niet benadrukt. Misschien omdat de observaties daar aardig met de veronderstelde verschillen sporen. In het algemeen kan worden opgemerkt dat Janssen in haar onderzoek vooral op zoek is naar verschillen en overeenkomsten, en die koppelt aan de hypothese van vier profielen. Waarom docenten verschillen, waarom leerlingen verschillen komt nauwelijks aan de orde. Haar empirisch-analytische benadering verhindert door te vragen naar 'waarom's en stelt zich tevreden met 'dats'.

Tanja Janssen is bescheiden ten aanzien van haar resultaten. Ze vraagt zich terecht af in hoeverre ze generaliseerbaar zijn. Ze kiest voor haar observaties en learner reports prototypische docenten, maar laat zich niet verleiden tot directe generalisatie. Ze constateert namelijk ook uit haar interviews dat er binnen de profielen grote verschillen bestaan. Daar komt bij dat kernbegrippen als 'literatuur', 'leesplezier', 'tekstervaring', 'cultureel besef', 'poëzie-analyse' niet eenduidig gedefinieerd zijn, maar verschillende inhouden dekken bij verschillende leraren. Sterker nog, niet voor alle leraren vormt literatuur een apart onderdeel van het schoolvak. Dat brengt me bij een kernvraag voor wat betreft de onderzoeksmethode. Janssen hanteert een gestandaardiseerd onderzoeksinstrument waarmee ze een gevarieerde praktijk in beeld wil brengen. Dat kan heel goed, als die praktijk voornamelijk verschilt in de mate waarin leraren verschillen op helder en eenduidig gedefinieerde variabelen. Het is de vraag of het analytische instrument toereikend is waar leraren kwalitatief verschillen in de kernbegrippen. Wat voor de een 'literatuur' is, is het blijkbaar niet voor de ander. Wat weten we dan als we vernemen dat de ene leraar er meer aandacht aan besteedt dan de andere? Genereert een gestandaardiseerd instrument die kenmerken van literatuuronderwijs die de geobserveerde docenten zelf wezenlijk vinden?

Twee kanttekeningen tot slot. Janssen constateert dat het profiel 'Culturele vorming' het meest gemarkeerd is. Ze veronderstelt dat de positieve rapportage door de (oud)leerlingen zeker ten dele aan de herkenbaarheid ervan kan worden toegeschreven. Ik voel er wel voor die veronderstelling te steunen. Vermoedelijk is de lespraktijk van dat profiel het meest 'schools': een duidelijk afgebakende leerstof, een duidelijke rolverhouding leraar-leerling, reproductieve taken voor leerlingen. Daar komt bij dat dit profiel een lange onderwijs traditie achter zich heeft: in de geschiedenis van het literatuuronderwijs in Nederland wordt voortdurend geklaagd over de praktijk van het literatuuronderwijs die te zeer historisch gericht zou zijn.

Het tweede enigermate herkenbare profiel is dat van de individuele ontplooiing. Ik vraag me af of ook dat profiel zijn herkenbaarheid te danken heeft aan de onderwijs traditie. Waar de culturele vorming kan bogen op een eeuw historisch-biografisch literatuuronderwijs, kan het profiel Individuele ontplooiing worden herleid tot de reformpedagogiek, die ook in het literatuuronderwijs van het begin van deze eeuw opduikt. Lezersgericht literatuuronderwijs is niet pas van de jaren tachtig, zoals Janssen opmerkt (De Moor 1984, Van de Ven 1990). Als dat waar is, dan blijkt opnieuw hoe sterk onderwijs tradities de praktijk van nu vormgeven. Dan blijkt dat ook de beide andere profielen op hun plaats vallen. Het Literair-esthetische profiel zou dan in het onderwijs als alternatief gelden voor het profiel van de Culturele vorming, parallel aan de vakwetenschappelijke strijd tussen een historisch-biografisch en een literair-structureel paradigma. De Maatschappelijke vorming presenteert dan een jaren zestig alternatief voor de individueel gerichte reformpedagogische traditie. Persoonlijk vind ik dit een perspectief dat de moeite waard zou zijn nader te onderzoeken. Janssen daagt me uit tot hernieuwd historisch onderzoek in dezen. En dat brengt me naar de vraag naar de bruikbaarheid van haar onderzoek. Janssen zegt daar zelf niets over. Mij lijkt dat, ondanks mijn kanttekeningen, Tanja Janssen ons een ongekende hoeveelheid gegevens verschaft over literatuuronderwijs. Ik denk dat die op verschillende terreinen bruikbaar zijn. In het kort:

Op **beleidsniveau** geeft Janssen aan dat de invoering van het Studiehuis en van de nieuwe eindexameneisen complexe en gevarieerde innovatiesituaties oproept. Het loskoppelen van het literatuuronderwijs van de andere onderdelen van het schoolvak Nederlands en de verplichte samenwerking met leraren andere talen is ongetwijfeld inhoudelijk verdedigbaar, maar betekent wel dat je leraren Nederlands raakt in wat ze het belangrijkste onderdeel van hun vak vinden. Dat levert ongetwijfeld acceptatie- en dus implementatieproblemen op. Het nieuwe examenprogramma, dat onder het concept van 'literaire competentie' de doelen van Culturele en literair-esthetische vorming en Individuele ontplooiing aan elkaar verbindt, impliceert dat veel docenten hun doelen en dus hun werkwijzen zullen moeten veranderen. Sterker nog, de inhoud van hun onderwijs wordt verruimd. En dat gaat gepaard met minder lessen leesuren. De commotie waar ik in mijn inleiding aan refereerde, is met behulp van Janssen werk erg verklaarbaar geworden.

Janssen bevindingen lijken me uiterst belangrijk voor **opleiding en nascholing**. Janssen constateert onder meer dat de doelstellingen van literatuuronderwijs "over het algemeen vaag (zijn) geformuleerd. Het gaat om complexe, veelomvattende zaken als 'culturele vorming', 'literair-esthetische vorming', 'maatschappelijke vorming', 'individuele ontplooiing' of met een moderne verzamelnaam 'literaire competentie'. Maar wat bedoelt men hier nu eigenlijk mee? Uit de interviews kwam naar voren dat docenten niet zo'n helder zicht hebben op hun doelstellingen. Literatuurlessen worden over het algemeen ook niet aan de hand van welomschreven doelen gepland. Docenten onderschrijven algemene doelen in meer of mindere mate, maar in de praktijk behandelen zij nar eigen zeggen vooral wat zij 'van belang', 'leuk' en/of 'interessant' vinden" (145). Daar kan aan worden toegevoegd dat, al zouden die doelen welomschreven zijn, ze hoogstens op een persoonlijk niveau eenduidig zouden zijn gezien de nogal persoonlijke invulling door leraren van kernbegrippen als 'literatuur' en 'cultuur'. Dat leidt tot aanzienlijke problemen in sectie-overleg en tot persoonlijke interpretaties én lespraktijken van wat in zulk overleg als gemeenschappelijk plan ontstaat (vgl. Rensman 1990). Elke onder-

wijs-leersituatie - opleiding dan wel nascholing - vergt dan ook dat (aanstaande) leraren hun subjectief concept pogen te verhelderen. Janssen indeling in vier profielen en haar operationalisatie van taakeisen zijn belangrijke instrumenten in een dergelijk leerproces.

Literatuur

- Hetebrij, M. & Wardekker, W.L. (1986). De kwaliteit van discussies over kwalitatief onderzoek. *Pedagogische Studiën* 63, 485-497.
- Hoogveen, M. & Bonset, H. (1998) *Het schoolvak Nederlands onderzocht*. Een inventarisatie van onderzoek naar onderwijs Nederlands als eerste en tweede taal in Nederland en Vlaanderen. Leuven/Apeldoorn: Garant.
- Moor, W.A.M. de (1984). Alwat vanzelve wast, behoeft men niet te zaaien. Over de ontwikkeling van tekst-ervaringsmethoden. *Forum der Letteren* 25, 4, 259-274.
- Rensman, Th. (1990). *Seeds of innovation. Report of a case study in the field of reading texts in Dutch lessons (1982-1988)*. Occasional Papers in Mother Tongue Education 3. Enschede: VALO-M.
- Ven, P.H. van de (1990). *'Het nut van het nutteloze'. Een interpretatie van visies op literatuuronderwijs sinds het einde van de negentiende eeuw*. Literatuurdidactische Verkenningen 29. Nijmegen: KUN, AKW.
- Ven, P.H. van de (1996). *Moedertaalonderwijs. Interpretaties in retoriek en praktijk, heden en verleden, binnen- en buitenland*. Groningen: Wolters-Noordhoff (diss. UU).

Piet-Hein van de Ven
Universitair Instituut voor de LerarenOpleiding
Erasmusplein 1, 20.01A
Postbus 9103, 6500 HD Nijmegen

Jaarverslag Vereniging voor Onderwijs Research 1998 - 1999

1. Samenstelling bestuur 1998 - 1999

Prof. dr. N. Verloop	Voorzitter	Leiden	1996-1999
Prof. dr. M. Mulder	Secretaris	Wageningen/Twente	1998-1999
Dr. W. Kuiper	Penningmeester	Twente	1998-2001
Prof. dr. J. van den Akker	Divisie Curriculum	Twente	1998-2001
Dr. A.B. Dijkstra	Divisie Onderwijs & Samenleving	Groningen	1998-2001
Prof. dr. W. Jochems	Divisie Hoger Onderwijs	Heerlen	1998-2001
Prof. dr. J. Pieters	Divisie Leren & Instructie	Twente	1998-1999
Prof. dr. P. Sanders	Divisie Methodologie & Evaluatie	Amsterdam	1998-2001
Prof. dr. F. van Wieringen	Divisie Beleid & Organisatie	Amsterdam	1998-2001
Dr. B. Versloot	Divisie Beroepsonderwijs, bedrijfsopleidingen en volwasseneneducatie	Utrecht	1996-1999
Prof. dr. Th. Wubbels	Divisie Lerarenopleiding en Leraarsgedrag	Utrecht	1997-2000
Mw. Drs. A. Veerman	Promovendi	Utrecht	1998-2001

2. Aantal leden

Divisies	jan 99	jan 98	1997	1996
BBV	151	161	162	170
LI	142	132	128	139
HO	107	103	102	110
LL	96	95	86	95
CU	80	76	72	74
BO	80	83	88	91
ME	77	83	77	85
OS	61	53	56	60
Totaal aantal VOR-leden	550	554	549	596

Deze ledenaantallen wijken iets af van die in het jaarverslag over het Verenigingsjaar 1997-1998, omdat voor dat verslag is uitgegaan van de ledenaantallen per mei 1998.

Het aantal leden van de VOR blijft gedurende de afgelopen jaren stabiel, nadat het in 1997 behoorlijk is gedaald. Die ledenaantallen van de voorgaande jaren waren echter geflatteerd. In feite wordt het ledenbestand steeds verder opgeschoond en komen er nieuwe leden bij. Daardoor blijven de cijfers redelijk stabiel, maar is er toch sprake van groei van het aantal leden. Het ledenaantal per april 1999 is gegroeid tot 576.

3. Ledenwerving

Er zijn dit jaar enkele acties ondernomen in het kader van de ledenwerving. Geconstateerd is dat leden van het VOR Promovendi Overleg (VPO) niet automatisch lid zijn van de VOR. Voor de VOR als geheel is het beter als de VPO-leden ook VOR-leden zijn. De ledenlijsten

van het (VPO) is gescreend en de VPO-leden-niet-VOR-leden zijn benaderd met het verzoek lid te worden van de VOR. Binnen het VPO bestaat de indruk dat het percentage VPO-leden dat VOR-lid is toeneemt.

De deelnemerslijst van de ORD '97 te Leuven is eveneens gescreend en de deelnemers die geen VOR-lid waren hebben een VOR-brochure toegestuurd gekregen.

Verder zijn de ledenbestanden van de VOR en de NVO naast elkaar gelegd om te bezien hoe groot de overlap tussen de leden van deze verenigingen is. Het doel daarvan is de NVO-leden een gereduceerd lidmaatschap van de VOR aan te bieden, zoals dat ook mogelijk is voor VELON- en NVvO-leden. Uit de vergelijking van de bestanden is gebleken dat er slechts elf dubbelleden zijn. De NVO wordt een mededeling gestuurd die wordt geplaatst in het Verenigingsorgaan van de NVO waarin de mogelijkheid om tegen een gereduceerd tarief lid te worden van de VOR wordt aangeboden. Verder zal daarin enige informatie over de VOR als vereniging worden opgenomen.

Het VOR-bestuur wijst op de mogelijkheid het VOR-Jaarboek ook als wervingsmiddel te gebruiken.

De procedure met betrekking tot de verwerking van aanmeldingen van nieuwe leden is als volgt. De secretaris ontvangt de aanmeldingsformulieren. Vervolgens stuurt deze:

- de nieuwe leden een welkomstpakket van de VOR (een welkomstbrief, het VOR-Jaarboek, een recent nummer van het TOR, een recente ORD-brochure, de statuten en het huishoudelijk reglement van de VOR en de VOR-gedragscode);
- de Divisiesecretaris(sen) van de Divisie(s) waarvan het nieuwe VOR-lid lid wil worden een kopie van het aanmeldingsformulier.

Van de Divisie(s) wordt verwacht dat ze het nieuwe lid ook welkom heten en hen de laatste Divisie-informatie toesturen (nieuwsbrief, activiteitenoverzicht e.d.).

Daarnaast krijgen Divisies (secretarissen) periodiek uitdraaien van hun ledenbestanden.

4. TOR

De redactie van het TOR heeft het VOR-bestuur verzocht de heren Glas, Mulder en Westershof te (her)benoemen als leden van de redactie. Het VOR-bestuur is hiermee akkoord gegaan.

Het VOR-Bestuur heeft een notitie ontvangen van de TOR-redactie waarin gewijzigd beleid wordt voorgesteld met betrekking tot het uitgeven van het TOR. Deze voorgestelde beleidswijziging komt voort uit het opdrogen van de kopij. De stand van zaken per mei 1999 is dat jaargang 1998 nummer 4 vol is maar nog moet verschijnen en dat voor jaargang 1999 nummer 1 een themanummer is ingediend. Daarna is de kopij praktisch op. Jaargang 2000 (de 25^e Jaargang van het TOR) wordt geheel gewijd aan het TOR-jubileum en zal bestaan uit review-artikelen vanuit de VOR-divisies. De TOR-redactie heeft voorgesteld het TOR met ingang van Jaargang 26 (2001) in het Engels te gaan uitgeven, of het aantal nummers per jaar terug te brengen naar twee. Een alternatief is de redactieformule en de identiteit van het TOR te wijzigen. De redactie staat op het standpunt dat er niet moet worden getornd aan de kwaliteitseisen. In de loop van het jaar zal over de toekomst van het TOR nadere besluitvorming plaatsvinden.

5. Ledenadministratie

De ledenadministratie is verder geoptimaliseerd en het huidige bestand is redelijk actueel. Het aantal klachten naar aanleiding van de gebrekkige ledenadministratie van voorgaande jaren is inmiddels sterk verminderd. Er komen de laatste maanden nauwelijks meer klachten binnen.

6. Contributie

De tariefstructuur voor het jaar 1999 is als volgt:

Basislidmaatschap gewone leden	f 110,-
Basislidmaatschap studenten	30,-
Basislidmaatschap promovendi (AIO-OIO- bursaal)	60,-
Basislidmaatschap combinatieleden (NVvO/VELON)	75,-
Lidmaatschap van een divisie is gratis	
Voor elk lidmaatschap van een <u>extra</u> divisie betalen de leden	10,-
Prijzen van tijdschriftabbonnementen:	
Tijdschrift voor Onderwijs Research	25,-
EERA Bulletin	25,-
International Journal of Educational Research	55,-
Learning and Instruction	45,-
Educational Research and Evaluation	55,-

7. Elektronische communicatie binnen de VOR

Vanuit de Divisie Hoger Onderwijs is een Website ontwikkeld voor de gehele VOR. De Divisies en de VOR zijn gevraagd contactpersonen aan te wijzen voor het aanleveren van kopij. Dat is gebeurd en de betrokkenen hebben dit gedaan. De Website is inmiddels operationeel en kan worden bezocht. De VOR bedankt de betrokkenen die zich voor het ontwikkelen van de VOR-website hebben ingespannen.

8. VOR-jaarboek

Tijdens de ORD 1998 is de 'nulde' versie van het VOR-Jaarboek uitgekomen. Het voornemen is ieder jaar een geactualiseerd Jaarboek uit te brengen. De reacties op het verschijnen van het VOR-Jaarboek zijn positief. Bij de voorbereiding van het Jaarboek 1999 is een buitengewoon grote hoeveelheid mutaties binnengekomen van de leden. Deze zijn verwerkt en het VOR Jaarboek 1999 bevat een actueel overzicht van het ledenbestand. Besloten is alle bezoekers van de ORD ook een exemplaar ter beschikking te stellen. Bij de voorbereiding van het VOR-Jaarboek 2000 iemand speciaal advertenties moeten gaan acquireren om de productiekosten te drukken.

Veel leden hebben trefwoorden met betrekking tot hun expertise ingevuld. Deze zullen worden gebruikt om een trefwoordenlijst te maken voor het volgende jaarboek. De leden zullen dan worden gevraagd drie trefwoorden aan te kruisen. Het trefwoordenregister zal worden opgenomen in het Jaarboek 2000.

9 ORD

De 25^e ORD heeft in 1998 plaats gehad in Enschede. Voorzitter van de organisatiecommissie was Prof. dr. J.M. Pieters.

De 26^e ORD in 1999 heeft plaats in Nijmegen. Voorzitter van de organisatiecommissie is Prof. dr. P.R.J. Simons. Naar aanleiding van de organisatie van deze ORD is vastgesteld dat er in de toekomst overleg moet zijn tussen het DB van de VOR en de lokale organisatiecommissie over de plaats en de inhoud van de VOR-bijeenkomsten (de plenaire opening, de Algemene Ledervergadering en de Divisievergaderingen) in het programma.

De 27^e ORD in 2000 wordt georganiseerd door de Universiteit Leiden.

De 28^e ORD in 2001 wordt georganiseerd door de Universiteit van Amsterdam.

De 29^e ORD in 2002 wordt georganiseerd door de Universiteit van Antwerpen.

10. EERA

Voorzitter en secretaris van de VOR hebben gedurende het afgelopen jaar vergaderingen van de EERA bijgewoond. Gesproken is over de verdere uitbouw van de EERA. De ECER conferentie in 1999 is in Lahti (Finland) en in 2000 in Edinburgh (Schotland).

Het EERA Bulletin gaat qua signatuur en naam veranderen. Er wordt onderscheid gemaakt in een peer reviewed gedeelte voor artikelen over onderwijsresearch die interessant zijn voor onderwijsonderzoekers in Europa, en in een verenigingsgedeelte voor nieuws vanuit de netwerken, en ledenverenigingen. De VOR is een lid-vereniging en wordt opgeroepen meer gebruik te maken van de mogelijkheid nieuws te publiceren in het EERA Bulletin. De nieuwe naam van het EERA Bulletin wordt: European Educational Researcher.

11. AERA

De Divisie Lerarenopleiding & Leraarsgedrag heeft in 1999 de VOR-presentatie verzorgd tijdens de AERA-conferentie. De Divisie Curriculum had ook belangstelling getoond voor dit jaar, maar heeft te kennen gegeven ook geïnteresseerd te zijn in 2000. De Divisie O&S heeft belangstelling voor de hierop volgende jaren. Er zal een meerjarenschema worden opgesteld voor Divisiepresentaties tijdens de AERA om ad hoc beslissingen in de toekomst te voorkomen.

12. VOR-prijs

De VOR-prijs 1998 is toegekend aan Dr. H. Biemans van de Universiteit Wageningen. Promotor van de laureaat was Prof. dr. P.R.J. Simons.

Predictoren van de keuze van natuurkunde in de tweede fase VO

Karel M. Stokking
Capaciteitsgroep Onderwijskunde, Universiteit Utrecht

ABSTRACT

The popularity of exact and technical subjects and studies is declining. A short review is given of research into this topic. On request of the Dutch Ministry of Education research has been done into the choice of physics in MAVO and HAVO (1986-1988) and VWO (1992-1995). The VWO-project replicated the MAVO-HAVO study with an additional focus on the impact of the curriculum, the stability of the choice behavior, and the consistency between analysis techniques. No significant differences between schoolbooks were found, but the perception by the students of several characteristics of the curriculum turned out to be of some relevance. The choice behavior appeared quite unstable. Different analysis techniques partly led to different outcomes, in particular with regard to the differences between the sexes and between the schools.

1. INLEIDING EN VRAAGSTELLING

De afgelopen 15 jaar is de belangstelling voor exacte en technische vakken, studies en beroepen aan het afnemen. In een aantal landen is beleid ontwikkeld om deze trend te keren. In Nederland worden de meeste exacte vakken door meisjes minder gekozen dan door jongens. De landelijke overheid voert vanaf medio 80-er jaren beleid om dit te veranderen. Vooral natuurkunde is bij meisjes weinig populair. Dit vak baart internationaal de meeste zorgen. In Nederland zijn hiervoor curriculaire vernieuwingen ontwikkeld en worden op scholen methoden gebruikt die duidelijk van elkaar verschillen.

In de periode 1986-1988 is op verzoek van het Ministerie van Onderwijs een onderzoek uitgevoerd in MAVO en HAVO naar de verschillen tussen jongens en meisjes en tussen kiezers en niet kiezers en de oorzaken van de beperkte populariteit van het vak (Jörg, Man in 't Veld, Wubbels & Verwey, 1990). Dit onderzoek is, zoals veel onderzoeken op dit terrein, gebaseerd op het zgn. Parsons/Eccles model van het kiezen van schoolvakken en de factoren die daarbij een rol spelen. De meeste in dit type onderzoek betrokken factoren zijn ascriptie-kenmerken.

Uit het onderzoek van Jörg et al. bleek dat jongens en meisjes al in het begin van de 2e klas duidelijk verschillen in interesse voor natuurkunde. Andere variabelen met relatief duidelijke verschillen waren de ervaren moeilijkheid, het gevoel van competentie, en de aan het vak toegekende relevantie voor de eigen toekomst. De voornaamste predictoren voor de uiteindelijke keuze van natuurkunde waren: relevantie, competentie, interesse, prestatie, moeilijkheid, en waardering (bij meisjes alle zes, bij jongens de eerste vier).

In de periode 1992-1995 is, op verzoek van hetzelfde Ministerie, een onderzoek uitgevoerd in het VWO, bedoeld om na te gaan in hoeverre de bevindingen van Jörg et al. daar ook opgaan, en deze bevindingen zo mogelijk aan te vullen (Stokking, 1995). De vraagstellingen, voorzover analoog aan die van Jörg et al., waren:

- (A) Welke verschillen doen zich voor tussen jongens en meisjes en tussen kiezers en niet kiezers van natuurkunde als examenvak?
- (B1) Welke factoren hangen met het kiezen van natuurkunde samen?
- (C1) Hoe verloopt de ontwikkeling in de klassen 2 t/m 4 van de belangstelling voor dit vak en de daarmee samenhangende factoren?

Bij vraag B1 was het onderzoek van Jörg et al. behalve op ascriptieve kenmerken gericht op een vergelijking tussen scholen die het experimentele PLON-curriculum gebruikten met scholen met een reguliere natuurkundemethode. In het VWO-onderzoek is de factor curriculum uitgebreider onderzocht, op drie manieren: door vergelijking tussen acht verschillende natuurkundemethoden; door oordelen van deskundigen over deze methoden als factoren mee te nemen; door oordelen van leerlingen als factoren mee te nemen. Vandaar vraag (B2): In hoeverre hangen de methode resp. de kenmerken van het natuurkunde-onderwijs met het kiezen van natuurkunde samen?

Een tweede aanvulling op het MAVO/HAVO-onderzoek betreft de gebruikte analyse-technieken. Waar wenselijk zijn aanvullende technieken ingezet, waaronder met name multiverniveau-analyse.

Een derde aanvulling betreft een uitgebreidere analyse van de variatie over de tijd. Dit leidde tot vraag (C2): Hoe stabiel zijn de scores op en de relaties tussen de variabelen over de tijd?

In dit artikel vatten we eerst kort samen wat de literatuur ons leert over keuzepercentages, verschillen tussen jongens en meisjes, en beleid om die percentages resp. die verschillen te beïnvloeden (par. 2). Daarna gaan we in op verklarende modellen en hun empirische status (par. 3), en op de mogelijke rol van het curriculum (par. 4). Dan beschrijven we de opzet, methoden en het meetinstrumentarium in het VWO-onderzoek (5.1 en 5.2). Vervolgens rapporteren we over de uitkomsten in relatie tot de eerste drie genoemde onderzoeksvragen, werkend met dezelfde variabelen en analysetechnieken als in het MAVO/HAVO-onderzoek (5.3.1). Tenslotte gaan we in op de uitkomsten van de bovengenoemde aanvullingen (5.3.2-5.3.4). We sluiten af met conclusies en discussie (par. 6).

2. ACHTERGRONDEN: KEUZEPERCENTAGES, VERSCHILLEN JONGENS-MEISJES, EN BELEID

Het belang van exacte en technische kennis wordt steeds meer onderkend. Zulke kennis is nodig voor steeds meer beroepen en is in een samenleving die drijft op wetenschap en techniek te zien als basiskennis. Tegen deze achtergrond is een (eventueel) afnemende interesse voor exacte en technische vakken en studies bij veel beleidmakers reden tot zorg. In hoeverre is deze zorg terecht?

De interesse van leerlingen voor natuurwetenschap neemt tijdens hun schoolloopbaan af (zie bijv. Walberg, 1991). Dat wil echter nog niet zeggen dat opeenvolgende cohorten zulke vakken minder kiezen. Zo concluderen bijv. Sobolewski en Doran (1996) voor de staat New York dat het percentage dat natuurkunde kiest al enkele decennia constant is. Ook in Nederland zijn de percentages die als examenvak natuurkunde kiezen de afgelopen decennia constant gebleven: afgerond 30% onder de HAVO-leerlingen en 45% onder de VWO-leerlingen (Dekkers, 1993; Spiering, 1994; Jörg, 1994). De dalende interesse betreft vooral enkele andere vakken in het VO (scheikunde in het VWO (van 50 naar 35%), en biologie in HAVO en VWO (van 50% naar 30%), alsmede de exacte en technische studies en opleidingen na het VO.

Het min of meer stabiel blijven van de keuzepercentages van natuurkunde in het VO geldt ook voor jongens en meisjes apart (afgerond in HAVO 50 resp. 15, in VWO 60 resp. 30). Wat is verder bekend over verschillen tussen jongens en meisjes in keuzegedrag?

Veel studies tonen aan dat meisjes als groep andere keuzes maken dan jongens (Rennie en Parker, 1993; Kelly, 1988; Oakes, 1990).

Diverse studies laten zien dat jongens en meisjes verschillen in interesse voor onderwerpen (Taber, 1991, Groot-Brittannië: jongens meer in mechanische en meisjes meer in lichamelijke onderwerpen; Oakes, 1990, VS: jongens meer in dingen, meisjes meer in mensen). Voor oorzaken wordt verwezen naar verschillen in attitudes en ervaringen (Jones en Wheatley, 1990; Oakes, 1990; Marion en Colodarci, 1993). Veel studies laten zien dat de meeste keuzemotieven en keuzepredictoren voor jongens en meisjes dezelfde zijn (Marion e.a., 1993; Kelly, 1988; McCarthy en Moss, 1994; Dekkers, 1993). Veel studies tonen aan dat geslacht op zich, indien andere predictoren zijn uitgepartialiseerd, geen belangrijke voorspeller meer is (Crawley & Black, 1992; Gallagher, 1994; Kelly, 1988).

Internationaal zijn diverse typen activiteiten ondernomen om de interesse te stimuleren en (veronderstelde) keuzetrends te keren. Een eerste type is de wervingsactie. In diverse landen zijn door de overheid campagnes gevoerd. Het grootste 'zorgenkind' is het vak natuurkunde. In de VS doen ook de leraren in dat vak veel moeite om leerlingen te blijven trekken (Schaeffer, 1991; Hellman, 1992; Boone & Roth, 1992; Grote, 1994; Gallagher, 1994; Lapp, 1996).

Een tweede type aanpak zoekt oplossingen in veranderingen in het curriculum, om dat voor leerlingen aantrekkelijker te maken.

Beide hiervoor genoemde aanpakken gaan vaak gepaard met een streven naar meer gelijkheid tussen de geslachten door te stimuleren dat meer meisjes de genoemde vakken volgen en verwante studies kiezen.

Een derde aanpak betreft het terugdringen van de keuzevrijheid. In Groot-Brittannië is inmiddels een nationaal curriculum ingevoerd, met daarin een breed vak natuurwetenschap (Harding en Parker, 1995). Een soortgelijke beweging in de richting van een nationaal curriculum is ook in de VS ingezet (Walberg, 1991). Ook in andere landen is men meer bepaalde vakken verplicht gaan stellen, zoals in Australië, Zweden en Denemarken (Harding en Parker, 1995).

De drie beschreven typen activiteiten zijn ook in Nederland beproefd. Er is getracht via landelijke campagnes de interesse te bevorderen door leerlingen, en vooral meisjes, te wijzen op het belang van exacte en technische vakken voor de eigen toekomstmogelijkheden. Daarnaast is door vakdidactici gewerkt aan veranderingen in het curriculum, qua inhoud en aanpak. En recent is besloten de keuzevrijheid terug te dringen en een groter deel van het curriculum als gemeenschappelijk verplicht te stellen, waarnaast nog slechts vier mogelijke profielen resteren, waarmee een betere aansluiting met het hoger onderwijs wordt beoogd.

3. VERKLARENDE MODELLEN EN UITKOMSTEN VAN ONDERZOEK

Bijna alle modellen voor keuzeprocessen gaan uit van de aanname dat individuen rationele afwegingen maken tussen alternatieven op basis van (informatie over) de waarden die deze voor hen vertegenwoordigen en de ingeschatte kans op succes (realisering). Bij de waarden wordt meestal onderscheid gemaakt tussen intrinsieke waarde (interesse en plezier) en utilitaire waarde (nut).

De relaties van interesse, toegekend nut en succesverwachting (ingeschatte haalbaarheid) met keuzevoornemens en feitelijke keuzes zijn vaak aangetoond (Meece, Parsons, Kaczala, Goff & Futterman, 1982; Eccles, 1987; Kelly, 1988; Altling en Pelgrum, 1990; Van Heugten, 1993; Mellink, Voncken, Bosma & Karstanje, 1993; Eccles, Barber, Updegraff & O'Brien, 1995). Dit wil niet zeggen dat niet ook andere factoren een rol kunnen spelen. Bij het zoeken naar aanvullende verklaringen is door veel onderzoekers gebruik gemaakt van het model van Parsons/Eccles (zie bijv. Meece et al. 1982). In dit model worden cognities gepostuleerd die tot genoemde waarden en verwachtingen leiden: percepties, interpretaties en attributies van ervaringen, resp. socialiserende invloeden. Met name de factoren die leiden tot de uiteindelijke succesverwachting zijn in deze traditie verder uitgewerkt, rond begrippen als waargenomen taakmoeilijkheid, beeld van eigen kunnen en zelfvertrouwen. De ervaringen die tot

deze cognities leiden betreffen onder meer succeservaringen (de ervaring iets goed te kunnen, in het onderwijs vaak in relatie tot behaalde cijfers) en verwachtingen van anderen. De exogene determinanten in het model zijn met name milieu, aanleg en geslacht.

Het model is (nog) niet in zijn geheel getoetst. Onderzoekers hebben steeds slechts onderdelen ervan in hun onderzoek opgenomen. Men kan zich afvragen in hoeverre het model geschikt is als theoretisch verklarend model, dan wel (slechts) is bedoeld als een ordening van variabelen voor instrumentconstructie en data-analyse. Ook is de vraag of de in het model aangegeven relaties alleen (verwachte) analyse-uitkomsten op geaggregeerd niveau beschrijven, of ook processen op individueel niveau. Verdere vragen zijn in hoeverre het model in feite zowel de ontwikkeling van leerlingen (een proces van jaren) als het maken van keuzes (vaak meer gebonden aan momenten of korte perioden) wil beschrijven, of het daarin slaagt, en of zo'n combinatie in het algemeen wel zo gemakkelijk is te maken.

In het model, tenslotte, ontbreken diverse potentieel relevante variabelen, zoals invloeden die op een school rond een keuzeprocess kunnen optreden (het curriculum; sociale normen in de klas; de beschikbare keuzeruimte; keuzebegeleiding), en de wijze waarop de keuze wordt gemaakt (de informatie die erbij wordt betrokken en de verwerking daarvan). In het hier te beschrijven onderzoek zijn, aanvullend op de Parsons/Eccles factoren, ook aspecten van het curriculum en verschillen tussen klassen onderzocht.

Onderzoeken laten het volgende zien. Het geslacht heeft zoals gezegd temidden van andere predictoren weinig voorspellende waarde. Het milieu hangt volgens sommige onderzoeken positief samen met het kiezen van natuurwetenschappelijke vakken (Gallagher, 1994 (de VS)), volgens andere niet (Dekkers et al., 1993, 1997 (Nederland)).

Blijkens sommige onderzoekers zijn interesse, toegekend nut en beeld van eigen kunnen de dominerende factoren (zie bijv. Kelly, 1988). Anderen menen te kunnen nuanceren dat jongens meer nadruk leggen op utilitaire waarden en meisjes meer op intrinsieke waarden (Bügel, 1993; Willems en De Grip, 1993), dat het beeld van eigen kunnen bij meisjes een grotere rol speelt dan bij jongens (Man in 't Veld e.a., 1990), en dat intrinsieke waarden bij leerlingen op een lager opleidingsniveau (bijv. VBO) een minder grote rol spelen dan op een hoger niveau (bijv. VWO) (Van Paridon en Vlaanderen, 1992). Men kan zich overigens afvragen of binnen het construct 'intrinsieke waarde' interesse en plezier niet meer moeten worden onderscheiden. Interesse verwijst immers vooral naar kennisinhouden, plezier meer naar de omgang daarmee, waarbij ook de sociale context een belangrijke rol zal spelen (verg. Erickson en Schultz, 1992).

Een aantal auteurs beschrijft dat leerlingen worden 'gekanaliseerd' naar bepaalde leerwegen (Kelly, 1988; Oakes, 1990; Ainley, 1993; Rennie e.a., 1993; Willems e.a., 1993). Keuzes van leerlingen zouden zo reeds sterk worden bepaald door het beleid dat de school voert. Sobolewski en Doran (1996) vonden inderdaad verschillen tussen scholen. Bosker en Dekkers (1994) vonden geen verschillen tussen scholen; Smits (1994) vond wel verschillen tussen klassen.

In recente Nederlandse onderzoeken is geconstateerd dat een (waarschijnlijk toenemend) aantal leerlingen probeert keuzes uit te stellen, resp. keuzes te maken die zoveel mogelijk opties open laten (Van Paridon et al., 1992; Mellink et al., 1993; Den Boer en Guldemond, 1996). Het motief 'daar kun je nog alle kanten mee op' wordt in deze onderzoeken ongeveer net zo vaak genoemd als de bekende reeks motieven 'interessant, leuk, nuttig, ben er goed in'.

De literatuur over de keuze van schoolvakken (academic choice) is tot dusver vrijwel gescheiden gebleven van die over de keuze van een beroep of beroepsloopbaan (career choice). Rondom het maken van loopbaankeuzes bestaan diverse theoretische stromingen: (1) Kiezen als beslissen, en als afstemming tussen persoon en beroep (zie voor een recent overzicht: Holland, 1997); (2) Kiezen als uitvloeisel van een gefaseerd proces van ontwikkeling en socialisatie (zie bijv. Super, 1980; Gottfredson, 1981); (3) Kiezen op basis van sociaal-cognitieve leerprocessen (zie Krumboltz, 1994; Lent, Brown en Hackett, 1994; Brown en Lent, 1996);

(4) Kiezen in relatie tot de ontwikkeling van een eigen identiteit (zie Spijkerman, 1992; Meijers en Piggott, 1995; Wijers en Meijers, 1996).

Het is wenselijk de literatuur over academic choice en die over career choice nauwer op elkaar te betrekken (zie Lent et al., 1994; Stokking, 1997). De hiervoor beschreven modellen over het kiezen van schoolvakken sluiten het duidelijkste aan bij de derde zojuist genoemde stroming. Volgens de vierde stroming hebben steeds meer jongeren steeds meer moeite met kiezen; zij stellen besluiten uit of zijn zelfs besluiteloos (zie Jones, 1989; Heppner en Hendricks, 1995). Leerlingen wisselen veel van voornemen, en een substantieel deel verandert achteraf alsnog van keuze of betreurt de gemaakte keuze en zou, als ze opnieuw kon kiezen, een andere keuze maken. In Nederland betreft dit momenteel zo'n 30% van de leerlingen (verg. De Boer en Guldemond, 1996; Dekkers en Smeets, 1997).

4. DE POTENTIËLE ROL VAN HET CURRICULUM

Zoals gezegd heeft het model van Parsons/Eccles vooral betrekking op door socialisatie beïnvloede cognities, en ontbreken potentieel belangrijke factoren die met het onderwijs zelf te maken hebben. Het onderwijs kan op diverse manieren invloed uitoefenen op keuzeprocessen en -beslissingen: via het aanbod van opleidingswegen en vakken; via werving, selectie en determinatie van leerlingen; via keuzebegeleiding en advisering; via het curriculum van de schoolvakken. Op de laatste mogelijkheid gaan we nu nader in.

Volgens een aantal onderzoekers zijn voor het bevorderen van de keuze van exacte vakken bepaalde kenmerken van het curriculum van belang. Deze betreffen deels de inhoud: aandacht voor activiteiten van mensen (natuurwetenschap als activiteit *van* mensen, en als belangrijk *voor* mensen); het belang voor opleiding en arbeidsmarkt; het ontbreken van *sexebias*. Deels betreft het de aanpak: actieve werkvormen; interactie en samenwerking; zelfstandig onderzoek; probleemgestuurde instructie in onderzoeksvaardigheden (Walberg, 1991; Schaeffer, 1991; Mason, Kahle & Gardner, 1991; Fensham, 1992; Gallagher, 1994; Young, 1995; Vazquez-Abad, Winer & Derome, 1996).

Volgens diverse auteurs is het geïnvolveerd raken in onderzoek van een probleem van belang voor de ontwikkeling van motivatie en interesse voor het vak (Gallagher, 1994; Vazquez-Abad et al., 1996). Dit komt overeen met inzichten in de cognitieve psychologie (Pratt, 1989, 1991). Dat zelfstandig onderzoek door leerlingen leidt tot een goede kennisstructuur is echter niet onomstreden. Walberg (1991) beschrijft hoe de beweging tot vernieuwing in de VS vanaf 1960 in de natuurwetenschappelijke vakken, gericht op zelfstandig onderzoek door leerlingen en het stellen van hogere orde vragen, en later ook op de ontwikkeling van het vak en de relevantie ervan voor de samenleving, leidde tot kritiek op de onduidelijke resultaten en tot de conclusie dat zo slechts een aftreksel van het vak overblijft.

In Nederland is, vooral bij natuurkunde, deze problematiek expliciet onderwerp geweest van theorievorming, ontwikkelwerk en onderzoek. Jörg et al. (1990) beschrijven het (traditionele) vak natuurkunde als formeel-abstract, met veel leerstof en weinig samenhangende inhoud, over onderwerpen die weinig aansluiten bij de interesses van meisjes en zich niet goed lenen voor toepassing in het dagelijks leven. Niet-vakgebonden contexten zouden natuurkunde aantrekkelijker kunnen maken door te letten op de herkenbaarheid in het dagelijks leven, de nuttigheid, en de toepasbaarheid (bijv. in een beroep).

Wierstra (1990) rapporteert over evaluatie-onderzoek in het Nederlandse PLON-project. Het PLON-materiaal is gemaakt naar aanleiding van het Harvard Project Physics en de ook in Nederland bij veel natuurkundeleraars gevoelde onvrede over het functioneren van hun vak. Veel leraren hadden het gevoel dat de leerstof en de lesvormen te weinig herkenbaar en uitdagend waren. Volgens Wierstra is natuurkunde-onderwijs leefwereldgericht in de mate waarin de kennis in verband wordt gebracht met de buitenschoolse ervaringen van de leerlingen. Het participatiegerichte karakter koppelt hij aan werkvormen waarin leerlingen zelf-

standig taken kunnen uitvoeren, en over hun onderzoek aan klasgenoten kunnen rapporteren. Ook Kuiper (1993) beschrijft deze dubbele vernieuwingsgerichtheid: de inhoud moet zich meer oriënteren op de leefwereld, en er moet sprake zijn van meer gevarieerde handelingsgerichte leerlingactiviteiten, ter bevordering van interesse, motivatie en vaardigheidsontwikkeling.

Al met al is er consensus over kenmerken waaraan curricula moeten voldoen om leerlingen te motiveren, nl. leefwereld- en participatiegerichtheid. Het eerste wijst naar het belang van verbindingen met buitenschoolse contexten. Het tweede komt vooral tot uiting in het zelf actief zijn van leerlingen, bij natuurkunde vooral via het uitvoeren van proeven en experimenten, en het discussiëren en rapporteren over bevindingen. In het hierna te beschrijven onderzoek is nagegaan in hoeverre deze kenmerken van natuurkunde-onderwijs samenhangen met de keuze voor het vak als examenvak.

5. HET ONDERZOEK IN HET VWO

5.1 Design

De keuze van het vakkenpakket voor het VWO-eindexamen (dat wordt afgenomen aan het einde van het zesde leerjaar) diende tot voor kort uiterlijk aan het einde van het vierde leerjaar te worden gemaakt. Op veel scholen was het beleid dat de leerlingen reeds aan het einde van het derde leerjaar een voorselectie van vakken kiezen, waarvan er dan een jaar later alsnog een deel moet afvallen. Het vak wordt gegeven vanaf het tweede leerjaar. De vraagstelling van het onderzoek besloeg daarmee het 2e t/m 4e leerjaar. Er werd budget beschikbaar gesteld voor twee dataverzamelingjaren. Besloten werd het onderzoek in twee cohorten uit te voeren: 2e-3e leerjaar en 3e-4e leerjaar (op dezelfde scholen), met twee metingen per jaar.

5.2 Selectie van scholen, schaalconstructie, respons, analyses

Selectie van scholen

Begin 1992 waren er in Nederland 500 scholen met een VWO-opleiding. Alle secties natuurkunde daarvan zijn aangeschreven met een korte enquête. Van 159 werd een ingevulde vragenlijst ontvangen, waarvan 100 bereid waren aan het onderzoek mee te doen en 59 niet. Beide groepen bleken op een aantal relevante variabelen niet significant te verschillen. Ook bleken de respondenten ($n=159$) niet significant te verschillen van een aselechte steekproef uit de nonrespondenten ($n=50$), die telefonisch op een aantal variabelen zijn bevraagd.

Procedure bij de constructie van de vragenlijst voor de leerlingen

Er is een schriftelijke enquête ontwikkeld voor leerlingen. Een concept is voorgelegd aan vijf inhoudsspecialisten ter beoordeling van de inhoudsvaliditeit. Ook is een try-out van de enquête gehouden bij leerlingen (132 leerlingen in zes 3e klassen op drie scholen). Hen is gevraagd de enquête in te vullen; op basis daarvan zijn de schalen op betrouwbaarheid gecheckt. Ook zijn gerichte vragen gesteld over de begripelijkheid en eenduidigheid van de formuleringen (verg. Belson, 1982). Op basis van de beoordelingen door de inhoudsspecialisten, de itemanalyses van de try-out en de beoordeling door de leerlingen is de enquête bijgesteld.

Gegevensverzameling en respons

De 100 vaksecties natuurkunde zijn viermaal benaderd: september 1992, mei 1993, oktober 1993, mei 1994. Steeds werd tweemaal gerappelleerd. De sectieleiders mochten in september 1992 zelf bepalen welke 2e resp. 3e klas in het onderzoek zou worden betrokken; dezelfde klassen dienden ook de volgende drie meetrondes mee te doen. In de eerste meetronde werden ingevulde vragenlijsten verkregen van 4564 leerlingen van 91 scholen, in de volgende rondes waren de aantallen als volgt: 3921 van 84, 2493 van 69, 2756 van 70. In totaal hebben 96 scholen meegedaan. Van 1371 leerlingen van 60 scholen zijn (identificeerbaar) gegevens

verkregen over alle vier meetrondes. Uitval kwam voort uit drie factoren. Ondanks de steeds bijgevoegde toelichting hebben de docenten in volgende rondes niet steeds dezelfde klassen benaderd. Een deel van de leerlingen is tijdens het onderzoek overstapt naar een andere opleiding dan VWO of verliet de school na een verhuizing. En een aantal leerlingen heeft niet (steeds) identificeerbaar deelgenomen (door niet de eigen naam op te geven).

De in het onderzoek betrokken variabelen en de schaalanalyses

Figuur 1 toont de variabelen die in de analyses zijn meegenomen.

Figuur 1: Overzicht van variabelen

variabelen op methodeniveau [t.b.v. onderzoeksvraag (B2)]

- gebruikte natuurkundemethode
- kenmerken van de gebruikte natuurkundemethode (schaalscores, gebaseerd op oordelen van onafhankelijke deskundigen)

variabelen op klasniveau

- het percentage meisjes in de klas
- het percentage leerlingen in de klas dat zich voorneemt te kiezen voor natuurkunde, resp. (laatste meetmoment) dat kiest voor natuurkunde

variabelen op leerlingniveau

losse variabelen [t.b.v. de onderzoeksvragen (A) en (B1)]

- ouderlijk milieu (beroep en opleiding van vader en moeder)
- geslacht
- cijfer voor natuurkunde (laatste rapportcijfer per meetmoment)
- keuze-advies van docent en/of decaan
- keuzevoornemen (5-puntsschaal) resp. keuze (nee/ja) voor natuurkunde
- motief of motieven voor de keuze (open vraag)

schaalvariabelen [t.b.v. onderzoeksvraag (B1)]

- de interesse in natuurkundige onderwerpen
- de ingeschatte relevantie van het vak voor de (eigen) toekomst
- de ervaren moeilijkheid van het vak natuurkunde
- de waardering voor de natuurkundelessen
- het zelfvertrouwen ten aanzien van het vak natuurkunde
- de mate waarin jongens en meisjes (in het algemeen) verschillen in interesse voor, affiniteit met, resp. inzet voor het vak natuurkunde

schaalvariabelen voor 'leerlingvriendelijkheid' [t.b.v. onderzoeksvraag (B2)]

- de ervaren duidelijkheid (van de leerstof, de leraar, de lessen, het gebruikte leerboek, de verkregen opdrachten)
- 'beroepen': de mate waarin de inhoud van de lessen verwijst naar (de relevantie van) natuurkunde in beroepen,
- en de mate waarin de leerling zou willen dat dit het geval was
- 'nut': de mate waarin door de docent resp. in de lessen aandacht wordt gegeven aan het gebruik en nut van natuurkunde in het dagelijks leven,
- en de mate waarin de leerling zou willen dat dit het geval was
- 'proeven': de mate waarin de leerlingen tijdens de natuurkundelessen proeven (kunnen of moeten) doen,
- en de mate waarin de leerling zou willen dat dit het geval was
- 'verwoorden': de mate waarin de leerlingen tijdens de natuurkundelessen bevindingen en meningen onder woorden (kunnen of moeten) brengen,
- en de mate waarin de leerling zou willen dat dit het geval was

De vier variabelen beroeps- en opleidingsniveau van vader resp. moeder vormden samen een schaal voor 'milieu' (alpha .84).

Op elk meetmoment is aan de leerlingen gevraagd of, als ze op dat moment al een vakkenpakket moesten kiezen, ze daarin ook natuurkunde zouden opnemen, met antwoordmogelijkheden lopend van "zeker niet" (1) en "waarschijnlijk niet" (2) via "misschien, weet ik nog niet" (3) tot "waarschijnlijk wel" (4) en "zeker wel" (5). Vanaf het derde meetmoment werd eerst gevraagd of ze al een vakkenpakket hadden moeten kiezen, en zo nee, om de hiervoor genoemde vraag te beantwoorden, resp. zo ja, of ze natuurkunde hadden gekozen.

Tabel 1: Overzicht van de schalen op leerlingniveau

schaal voorbeelditem	aantal items	type antwoord- modaliteit	range van de alpha's
interesse <i>ik lees graag boeken of tijdschriften die iets met natuurkunde te maken hebben</i>	5	'eens'	.71-.79
relevantie <i>met natuurkunde als examenvak kan ik later een heleboel kanten op</i>	3	'eens'	.82-.87
moeilijkheid <i>ik vind natuurkunde-proefwerken meestal moeilijk</i>	5	'eens'	.74-.81
waardering <i>meestal vind ik de natuurkundelessen leuk</i>	5	'eens'	.87-.91
zelfvertrouwen <i>als ik goed meedoe in de les dan haal ik wel goede cijfers voor natuurkunde</i>	3	'past'	.61-.81
verschillen jongens-meisjes <i>natuurkunde gaat over onderwerpen die jongens meer aanspreken dan meisjes</i>	5	'eens'	.76-.81
duidelijkheid <i>de leerstof wordt in het natuurkundeboek lang niet altijd duidelijk uitgelegd (-)</i>	9	'eens'	.68-.76
beroepen - feitelijk <i>in de natuurkundeles wordt aandacht besteed aan beroepen waarvoor kennis van natuurkunde nodig is</i>	7	'vaak'	.72-.78
beroepen - zou willen			.77-.82
nut - feitelijk <i>de leraar/lerares laat zien wat alledaagse dingen om je heen met natuurkunde te maken hebben</i>	9	'vaak'	.75-.78
nut - zou willen			.78-.82
proeven - feitelijk <i>leerlingen doen proeven in de natuurkundeles</i>	3	'vaak'	.69-.75
proeven - zou willen			.67-.79
verwoorden - feitelijk <i>in de natuurkundeles kun je vertellen hoe je een natuurkunde opdracht hebt opgelost</i>	5	'vaak'	.63-.66
verwoorden - zou willen			.67-.73

De schalen in de vragenlijst besloegen in totaal 75 items. Gewerkt is met een 4-puntsschaal en drie typen antwoordmodaliteiten: van 'helemaal niet mee eens' tot 'helemaal mee eens' (in tabel 1 aangeduid met 'eens'), van 'nooit' tot 'vaak' ('vaak'), en van 'past helemaal niet bij mij' tot 'past helemaal bij mij' ('past').

Op elk meetmoment is van alle schalen de betrouwbaarheid nagegaan, voor alle leerlingen tezamen en voor jongens en meisjes apart. Als criteria voor een goede schaal zijn gehanteerd: alle r_{ir} 's tenminste .35 en de Cronbach's alpha tenminste .70. Schalen met r_{ir} 's tussen .20 en .35 en/of een alpha tussen .60 en .70 zijn geaccepteerd, maar opgevat als zwakke schalen. Items met r_{ir} 's lager dan .20 en schalen met alpha's lager dan .60 zijn niet geaccepteerd.

Tabel 1 geeft per schaal een overzicht van het aantal items, een voorbeelditem, het type antwoordmodaliteiten, en de range van de (per meetronde en per geslacht) bepaalde Cronbach's alpha. Zoals is te zien, zijn de alpha's bij de meeste schalen opmerkelijk constant.

Ter controle op de validiteit van de conform de onderhavige onderzoekstraditie opgenomen factoren is elk meetmoment aanvullend een open vraag gesteld naar de motieven om wel of niet natuurkunde te kiezen. Deze bleken inderdaad grotendeels te passen binnen de drie bekende dimensies: of men het vak interessant vindt, of men denkt het nodig te hebben, en of men er goed in is.

De schalen voor de 'leerlingvriendelijkheid' van het natuurkunde-onderwijs zijn geconstrueerd op aanwijzingen van Jörg e.a. (1990), bedoeld om natuurkunde-onderwijs aantrekkelijk te maken voor meisjes, verdeeld over een aantal componenten: leerstofkeuze, contextkeuze, werkvormen, ondersteuning van het leerproces, en beeld van natuurkunde resp. natuurwetenschap. Bij de operationalisering in ons onderzoek is getracht duidelijk onderscheid te maken tussen: *het vak* natuurkunde, *de leerboeken* en lesmaterialen, de werkwijze en uitstraling van *de docent*, het verloop en de sfeer van *de lessen*, en de beoordeling van *'het onderwijs' als geheel*.

Naast een schaal over 'duidelijkheid' zijn vier schalen meegenomen die twee aan twee corresponderen met de genoemde curriculumkenmerken leefwereldgerichtheid en participatiegerichtheid (zie par. 5):

- er is aandacht voor natuurkunde in relatie tot beroepen;
- er is aandacht voor het praktisch nut van natuurkunde;
- er worden proeven gedaan;
- leerlingen kunnen (moeten) resultaten zelf verwoorden.

Per item is op twee manieren bevraagd: de mate waarin het gestelde volgens de leerling voorkomt, en de mate waarin de leerling zou willen dat het voorkomt. Voor beide manieren van bevragen en op alle vier meetmomenten geldt dat de eerste, tweede en vierde schaal ('beroepen', 'nut', 'verwoorden') kunnen worden samengevat in één schaal (alpha's tussen .75 en .82). De aparte positie van 'proeven' kan mogelijk worden verklaard doordat het daarbij (ook) gaat om (fysieke) handelingen, en niet alleen of primair om kennisinhouden.

Beoordeling van schoolboeken voor natuurkunde

De acht methodes die op de aan het onderzoek deelnemende scholen het meest werden gebruikt (dekking: 85%), zijn beoordeeld door twee onafhankelijke deskundigen, op twee dimensies: cursorisch-thematisch en sturend-geleidend. Deze zijn vakdidactisch als volgt ingevuld:

* *Cursorisch versus thematisch*: een indeling volgens de natuurwetenschappelijke discipline versus een opbouw op basis van thema's die uitgaan van onderwerpen en situaties die in de leefwereld van de leerling voorkomen. Beoordelingsaspecten: indeling van de leerstof, toepassingen in de eigen leefwereld, verwijzingen naar de factor mens, realistische situaties in proeven, realistische illustraties.

* *Sturend versus geleidend*: leerstof die de leerervaringen direct stuurt versus leerstof die voorwaarden biedt om zelfontdekkend te leren. Beoordelingsaspecten (van de opdrachten): productief-reproductief, open-gesloten, instruerend-verwerkend.

Alle methodes zijn gescoord op alle aspecten, op 5-puntsschalen. De interbeoordelaars-

betrouwbaarheid is vastgesteld met een jury-alpha. Daartoe is het analyse-instrument door beide deskundigen toegepast op het eerste deel van één van de methodes. Waar de scores lager correleerden dan .80, is de omschrijving aangescherpt. Na scoring van alle methoden is per dimensie over de (vijf resp. drie) aspecten de Cronbach's alpha berekend. Voor 'cursorisch-the-matisch' was deze .74; voor 'sturend-geleidend' .84. Op grond daarvan is per methode per dimensie een gemiddelde berekend. De methodes bleken duidelijk te verschillen op zowel de aspecten als de dimensies (de schalen). De schaalcores zijn toegevoegd aan de leerlinggegevens.

Beslissingen over de data-analyse

De 1371 leerlingen die (identificeerbaar) aan alle vier de meetmomenten hebben meegedaan bleken, op de onderscheiden variabelen, niet duidelijk af te wijken van de leerlingen die wel aan het eerste meetmoment hebben meegedaan maar niet aan alle drie volgende ($n=3193$), noch van de leerlingen die niet aan het eerste meetmoment meededen, maar wel aan één of meer volgende ($n=1252$). Het werd hierom verantwoord geacht de analyses te concentreren op de leerlingen waarvan de gegevens van alle meetmomenten beschikbaar zijn. De analyses zijn uitgevoerd per cohort per meetmoment, overall en apart voor jongens en meisjes en voor (uiteindelijk) kiezers en niet-kiezers van natuurkunde. Voor het analyseren van verbanden is, ter vergelijking met het MAVO/HAVO-onderzoek (zie par. 1), gebruik gemaakt van correlaties, multiple regressie-analyses, en discriminantanalyses. Aanvullend zijn logistische regressie-analyses, LISREL-analyses, variantie-analyses en multivariate-analyses uitgevoerd. (Deze worden verderop ter plekke toegelicht).

5.3 Uitkomsten

De beschrijvende uitkomsten (zoals de gemiddelden) zijn in dit onderzoek, gelet op de vraagstelling en de gebruikte 4-puntsschalen, op zich minder interessant, en vooral van belang bij de vraag naar het verloop in de tijd. Er wordt dan ook in dat verband over gerapporteerd (zie 5.3.4). Eerst rapporteren we over de verschillen tussen jongens en meisjes resp. kiezers en niet-kiezers, en de factoren die met de keuze van natuurkunde samenhangen.

5.3.1 De analyses ter vergelijking met het MAVO/HAVO-onderzoek

Verschillen jongens - meisjes en kiezers - niet-kiezers

Meisjes scoren gemiddeld significant lager op interesse, relevantie, waardering, duidelijkheid, zelfvertrouwen en cijfer natuurkunde, en hoger op moeilijkheid. De maten waarin de onderscheiden leerstof- en aanpak-kenmerken feitelijk voorkomen scoren bij meisjes gemiddeld lager (zij nemen deze kennelijk minder waar). Een advies om natuurkunde te kiezen kregen meisjes minder vaak. De maten waarin meisjes zich voornamen natuurkunde te kiezen en dit uiteindelijk daadwerkelijk doen, liggen gemiddeld lager. Duidelijke verschillen tussen jongens en meisjes in het verloop van de variabelen, die consistent zijn voor beide cohorten, zijn er niet.

Kiezers scoren gemiddeld significant hoger op interesse, relevantie, waardering, duidelijkheid, zelfvertrouwen en cijfers wiskunde en natuurkunde, en lager op moeilijkheid. De mate waarin in de lesstof het nut van natuurkunde naar voren komt is volgens kiezers hoger dan volgens niet-kiezers (de laatsten nemen dit kennelijk minder waar). Kiezers kregen vaker advies om natuurkunde te kiezen en namen zich dit ook vaker voor. Duidelijke verschillen tussen kiezers en niet-kiezers in verloop, consistent voor beide cohorten, zijn, dat interesse, waardering en zelfvertrouwen bij de niet-kiezers afnemen en bij de kiezers niet; en dat duidelijkheid en relevantie bij de niet-kiezers afnemen en bij de kiezers toenemen.

Voor beide geslachten en beide keuzegroepen geldt in beide cohorten dat de mate waarin in het natuurkundeonderwijs proeven een rol spelen, afneemt; de aan natuurkunde toegekende moeilijkheid toeneemt; de behaalde cijfers afnemen; en de waargenomen verschillen tussen beide geslachten in affiniteit voor natuurkunde op het vierde meetmoment (als echt moet worden gekozen) toenemen.

Bijna alle bevindingen bij de vergelijking tussen kiezers en niet-kiezers over alle leerlingen, blijken ook te gelden indien per keuzegroep voor jongens en meisjes apart bekeken. Er zijn dus geen aanwijzingen voor verschillen in *niveau* of *verloop* tussen kiezers en niet-kiezers die specifiek zouden zijn voor jongens of meisjes.

Aanvullend is met LISREL getoetst of de correlatiematrix over de onderzoeksvariabelen tussen de geslachten en tussen de keuzegroepen gelijk mogen worden geacht. Deze analyse is voor beide vergelijkingen achtmaal uitgevoerd: per cohort per meetmoment. De uitkomsten laten zien dat ook geen sprake van consistent verschillende *samenhangen* tussen de variabelen.

Predictoren voor de keuze van natuurkunde

Voor een eerste indruk zijn enkelvoudige correlatieberekeningen uitgevoerd. Als voornaamste predictoren van de keuzevariabele (correlaties tenminste .20 en consistent over meetmomenten) komen naar voren: interesse, relevantie en waardering, gevolgd door cijfer voor natuurkunde, zelfvertrouwen, moeilijkheid en duidelijkheid.

Voor een vergelijkende analyse van de samenhang tussen de diverse predictoren en het keuzevoornemen zoals bevraagd op het 1e, 2e en 3e meetmoment (een numerieke variabele) is multiple regressieanalyse een geschikte techniek. Dergelijke analyses zijn uitgevoerd per meetmoment, met en zonder geslacht en deskundigenoordelen (over de methodes) als predictoren. Als voornaamste voorspellers (bèta-coëfficiënten in absolute waarde consistent tenminste .15) komen dan naar voren: relevantie, waardering en zelfvertrouwen. Milieu, geslacht, de deskundigenoordelen, en het verkregen advies, blijken geen rol van betekenis te spelen. De variabele interesse, die een hoge enkelvoudige correlatie met keuzevoornemen liet zien, blijkt vrij sterk in andere predictoren te zijn verweven.

De uiteindelijke keuze, op het 4e meetmoment, is een dichotome variabele, en dus niet geschikt als criteriumvariabele in regressie-analyse. Als alternatief is, conform Jörg e.a. (1990), discriminant-analyse uitgevoerd, en wel per cohort en per geslacht. De uitkomsten bevestigen dat relevantie, waardering en zelfvertrouwen een rol spelen, nu naast behaalde cijfers, advies en geslacht.

Discriminantanalyse kent assumpties waaraan vaak niet is voldaan, zoals multivariate normaliteit van de predictoren en gelijke covariantiematrices per keuzegroep. Ter vergelijking zijn daarom aanvullend logistische regressieanalyses uitgevoerd. De enige duidelijke en significante predictor is dan relevantie (partiële regressiecoëfficiënt .23 in het eerste cohort, .29 in het tweede).

5.3.2 Het curriculum apart bekeken

Op de deelnemende scholen werd in de onderzochte leerjaren voornamelijk een achttal methodes gebruikt (zie eindnoot). De relatie daarvan met de overige onderzoeksvariabelen is nagegaan met enkelvoudige variantieanalyses, per cohort per meetmoment.

De groepen scholen met dezelfde methode, verschillen significant op de variabelen waardering, duidelijkheid, en het voorkomen en de wens tot voorkomen van de leskenmerken 'beroep', 'nut' en 'verwoorden'. De ervaren moeilijkheid en de behaalde cijfers verschillen meestal significant, maar de rangorde tussen de methodes varieert. We nemen aan dat hier vooral de persoon van de docent een rol speelt. Het percentage leerlingen dat het advies kreeg natuurkunde in het pakket te houden varieerde tussen methodes van 21% tot 43%. Keuzevoornemen en feitelijke keuze verschilden echter niet tussen methodes.

Al met al zijn vooral verschillen tussen methodes zichtbaar op de variabelen waardering, duidelijkheid en het (volgens de leerlingen) feitelijk voorkomen van en de wenselijkheid van het voorkomen van de lesstof- en aanpak-kenmerken. We leiden daaruit af dat de waardering voor de lessen niet alleen samenhangt met de interesse in natuurkundige onderwerpen van de leerling zelf of met de wijze van lesgeven van de docent, maar ook met de methode die wordt gebruikt.

Drie methodes behoren op de meeste indicatoren tot de hoogst scorende. Dit gaat min of meer op voor het deskundigenkenmerk 'sturend-geleidend', voor drie van de vier door de leerlingen waargenomen kenmerken (niet voor 'proeven') en voor de variabele waardering. Dit illustreert dat de gebruikte methode kennelijk enigszins samenhangt met wat er in het natuurkunde-onderwijs gebeurt, dat leerlingen in staat zijn daarin verschillen op te merken, en dat hun waardering daarmee samenhangt.

De oordelen van de deskundigen discrimineren sterker tussen methodes dan de door de leerlingen beoordeelde kenmerken, en die weer sterker dan de variabele waardering. Kennelijk resulteren verschillen in het formele curriculum slechts ten dele in verschillen in het operationele resp. het ervaren curriculum.

Dit neemt niet weg dat leerlingen (jongens zowel als meisjes) die graag hadden gewild dat in de lessen meer aandacht was geschonken aan het belang van natuurkunde in arbeid en beroep, aan het nut van natuurkunde en praktische toepassingen, en aan het zelf onder woorden brengen en samen bespreken van natuurkundige onderwerpen (problemen, oplossingen, conclusies), gemiddeld minder vaak voor natuurkunde kozen, dan zij die hieraan minder behoefte hadden.

Aanvullend op de schalen die het natuurkunde-onderwijs beschrijven zoals waargenomen resp. gewenst door de leerlingen, zijn per schaal de verschillen berekend tussen 'waargenomen' en 'gewenst'. Op alle meetmomenten wil bij alle kenmerken tenminste 70% van de leerlingen graag dat ze *meer* voorkomen dan feitelijk het geval is.

5.3.3 De uitkomsten van multiniveau-analyses

De analyses in 5.3.1 betroffen uitsluitend variabelen op het niveau van de individuele leerlingen. Het feit dat groepen leerlingen op dezelfde school zitten en dat groepen scholen dezelfde methode gebruiken, werd hierbij genegeerd. (De deskundigenoordelen over de methodes waren gedesaggregeerd). In 5.3.2 is nagegaan in hoeverre sprake is van verschillen tussen groepen scholen die dezelfde methode gebruiken. Die waren er ten aanzien van de beoordeling door de leerlingen van kenmerken van het natuurkunde-onderwijs, de behaalde cijfers, en de keuze-adviezen, maar niet ten aanzien van het al dan niet kiezen van natuurkunde door de leerlingen.

Aanvullend zijn, ter controle op de in 5.3.1 en 5.3.2. beschreven uitkomsten, multiniveau-analyses uitgevoerd, per cohort per meetmoment, met als niveaus natuurkundemethode, school, en leerling. Daarbij is gebruik gemaakt van de programma's ML3 en MLwiN.

Eerst zijn, ter vergelijking met de uitkomsten van de multiple regressie-analyses, de gegevens van de eerste drie meetmomenten geanalyseerd, per cohort per meetmoment, met keuzevoornemen als kwantitatieve criteriumvariabele.

Deze analyses bevestigen het belang van de predictoren interesse (in natuurkundige onderwerpen), toegeschreven relevantie (van het vak), waardering (voor de lessen), en daarna behaald cijfer, ervaren moeilijkheid, zelfvertrouwen, en (enigszins) ervaren duidelijkheid. De specifieke oordelen van de leerlingen over de onderscheiden lesstof- en aanpakkenmerken speelden echter geen rol (meer).

Begin 2e klas, bij de aanvang van het natuurkunde-onderwijs, konden de verschillen tussen leerlingen in keuzevoornemen voor ongeveer 40% worden voorspeld door de in dit onderzoek betrokken predictoren; op de volgende twee meetmomenten was dit steeds rond 60%.

De multiniveau-analyses bevestigden tevens het feit dat er geen substantiële en over cohorten consistente significante verschillen zijn in het gemiddelde keuzevoornemen tussen groepen leerlingen die te maken hadden met eenzelfde natuurkundemethode.

De analyses lieten tenslotte amper of geen variantie zien op de keuzevariabele tussen scholen. De relaties van geslacht en milieu met de mate van voornemen om natuurkunde te kiezen, die er gemiddeld amper waren, bleken evenmin te variëren over scholen.

Vervolgens zijn, ter vergelijking met de uitkomsten van de discriminantanalyses en logistische regressie-analyses, de gegevens van het vierde meetmoment geanalyseerd, met de uiteindelijke keuze als dichotome criteriumvariabele. Deze analyse is, gezien het ontbreken van significante variantie op het niveau methode, toegespitst op de niveaus school en leerling, en op de drie belangrijkste predictoren interesse, relevantie, waardering, aangevuld met de variabele geslacht. Bij deze analyse is voor de criteriumvariabele een binomiale verdeling aangenomen.

In deze analyse blijken relevantie en waardering significante predictoren te zijn, maar interesse niet (meer), evenmin als de variabele geslacht. Wel is sprake van significante variantie (in de uiteindelijke keuze) op het niveau van de scholen.

Deze laatste uitkomst wijkt af van de bevinding op de eerste drie meetmomenten. Om op dit punt nader uitsluitsel te krijgen, mede gezien de wisselende conclusies in de literatuur (zie par. 3), is de navolgende aanvullende analyse uitgevoerd.

Aangezien per cohort per school slechts één klas aan het onderzoek deelnam, hebben analyses per cohort van verschillen tussen scholen feitelijk betrekking op verschillen tussen klassen. Nu kunnen klassen hun eigen 'klimaat' ontwikkelen, ook wat betreft de populariteit van bepaalde vakken en studiekeuzen (verg. Erickson en Shultz, 1992). Ook kan het percentage jongens resp. meisjes in de klas van belang zijn (een variabele die in de voorgaande analyses nog niet voorkwam).

De verschillen tussen klassen (in dit onderzoek overeenkomend met: per cohort tussen scholen) in de verhouding jongens-meisjes kunnen, door het verschil in percentage kiezers van natuurkunde tussen beide geslachten, van invloed zijn op de te constateren verschillen tussen scholen in keuzepercentage. Zulke verschillen kunnen worden verwacht als scholen de leerlingen duidelijk kanaliseren (zie par. 3).

Om dit na te gaan is een analyse uitgevoerd van verschillen tussen scholen in percentage kiezers van natuurkunde, rekening houdend met de verschillen in keuzepercentage tussen de geslachten en tussen de beide op de scholen in het onderzoek betrokken cohorten. Daartoe is het bestand van 58 scholen met complete gegevens op klasniveau per school verdeeld in (maximaal) vier cellen: cohort 1 - jongens, cohort 1 - meisjes, cohort 2 - jongens, cohort 2 - meisjes. In totaal leverde dit 192 niet lege cellen op. De beoogde analyse corrigeert voor verschillen in samenstelling per cel.

Een tweenniveau random intercept model voor de ware proportie kiezers, gelet op de variabelen cohort en geslacht, kan als volgt worden geschreven (verg. Prosser, Rasbash & Goldstein, 1991, p.25):

$$\Pi_{ij} = \frac{e^{(\gamma_0 X_0 + \gamma_1 X_{1ij} + \gamma_2 X_{2ij} + u_{0j})}}{1 + e^{(\gamma_0 X_0 + \gamma_1 X_{1ij} + \gamma_2 X_{2ij} + u_{0j})}}$$

met X_0 als een vector met 1-en, X_1 als cohort en X_2 als geslacht. Bovenstaande uitdrukking kan ook worden geschreven als:

$$\text{logit}(\Pi_{ij}) = \gamma_0 X_0 + \gamma_1 X_{1ij} + \gamma_2 X_{2ij} + u_{0j}$$

De u_{0j} specificeren de toevallige afwijkingen per school van het overall percentage kiezers, en $\sigma_{u_{0j}}^2$ is een maat voor de variantie tussen scholen. De analyse is uitgevoerd, gebruikmakend van de aanpak van Prosser et al. (1991, p. 107), met behulp van de volgende getransformeerde criteriumvariabele:

$$Y_{ij} = \ln \frac{m_{ij}}{n_{ij} - m_{ij}}$$

waarin de m_{ij} staan voor de aantallen kiezers per cel per school, en de n_{ij} voor de aantallen leerlingen in die cellen.

De analyse gaf een schatting van de tussen-scholen-variantie van .20 bij een standard error van .12, zodat de conclusie luidt dat er geen significante verschillen zijn tussen scholen in percentage kiezers, rekening houdend met de verschillen in percentage kiezers tussen de klassen (resp. cohorten) en tussen de geslachten.

De hiervoor beschreven analyses vullen elkaar aan. Ze zijn onderling meer of minder verschillend, maar betroffen alle de afzonderlijke meetmomenten. Afsluitend is een analyse verricht die gebruik maakt van het herhaalde metingen design. Uitgevoerd is een vierniveau analyse: methode, school, leerling, meetmoment.

De keuzevariabele op de eerste drie meetmomenten (keuzevoornemen, vijfpuntsschaal) wijkt af van die op het vierde (keuze, dichotoom). Daarom is de analyse toegespitst op de eerste drie meetmomenten.

Op een aantal scholen werden tijdens het onderzoek in verschillende leerjaren verschillende methoden gebruikt (vooral in het 2e cohort). De analyse is toegespitst op die situaties waarin in beide opeenvolgende leerjaren dezelfde methode werd gebruikt. Aanvullende eisen waren: tenminste vier leerlingen per cohort per school met volledige gegevens over de predictoren interesse, relevantie, waardering en geslacht. De analyse is uitgevoerd op 451 leerlingen in het 1e cohort op 40 scholen, verdeeld over 7 methoden.

De variantiecomponentenanalyse leverde de volgende schattingen voor de varianties van de criteriumvariabele keuzevoornemen op de vier niveaus: methode: 0, school: 0,048 (SE 0,026), leerling: 0,491 (SE 0,051) en moment: 0,694 (SE 0,033). Kortom: geen significante verschillen tussen scholen (4% van alle variantie) en grotere verschillen tussen meetmomenten (56%) dan tussen leerlingen (40%). Inbreng van de genoemde predictoren op leerlingniveau (die ook nu weer significant bleken) deed deze verdeling verder verschuiven, tot: 3% (school), 14% (leerling) en 83% (meetmoment).

Kennelijk is de keuzevariabele zelf allerminst stabiel. Dit brengt ons bij de volgende, laatste paragraaf met uitkomsten.

5.3.4 De stabiliteit van variabelen en relaties over de tijd

Tabel 2 bevat de gemiddelden per variabele per meetmoment, zowel overall als per cohort. Behalve bij de cijfers, waarvoor de in Nederland gebruikelijke schaal van 1 tot 10 geldt, is steeds een vierpuntsschaal van toepassing.

De meeste gemiddelde schaa scores blijven over de vier meetmomenten opvallend constant. De toegeschreven relevantie en moeilijkheid nemen in het tweede onderzoeksjaar licht toe (in beide cohorten, dus in het 3e leerjaar vergeleken met het 2e leerjaar, en in het 4e leerjaar vergeleken met het 3e leerjaar). De mate waarin met proeven wordt gewerkt neemt in het tweede onderzoeksjaar in het tweede cohort af (4e leerjaar; hier speelt een rol dat in een aantal scholen dan een andere methode wordt gebruikt). De waardering voor het natuurkunde-onderwijs neemt in het tweede leerjaar gemiddeld vrij duidelijk af (eerste cohort, 1e en 2e meting). In beide cohorten nemen in beide onderzoeksjaren de gemiddelde rapportcijfers voor wiskunde, Nederlands en natuurkunde af (het cijfer voor natuurkunde ontbreekt bij de 1e meting in het eerste cohort, omdat de leerlingen in de 1e klas nog geen natuurkunde hebben gehad).

Per subgroep bekeken, jongens-kiezers, jongens-niet-kiezers, meisjes-kiezers, en meisjes-niet-kiezers, treden meer verschuivingen aan het licht. Bij enkele variabelen is er vanaf het begin enig onderscheid tussen de vier groepen, en blijft dit min of meer gelijk. Dit patroon

doet zich vooral voor bij interesse en moeilijkheid. Bij enkele andere variabelen echter liggen de gemiddelden van de vier groepen op het eerste meetmoment nog vrij dicht bij elkaar, maar gaan deze in de loop van de tijd meer uiteen lopen. Dit betreft vooral relevantie en zelfvertrouwen, en ook waardering.

Tabel 2: Gemiddelden totaal en per cohort per meetmoment

meetmoment:	totaal				cohort 2e-3e klas				cohort 3e-4e klas			
	1	2	3	4	1	2	3	4	1	2	3	4
interesse	2,2	2,1	2,1	2,1	2,2	2,0	2,0	2,0	2,2	2,2	2,2	2,2
relevantie	2,9	3,0	3,1	3,1	2,7	2,7	2,9	3,0	3,2	3,4	3,3	3,2
waardering	2,8	2,6	2,6	2,6	2,9	2,5	2,5	2,5	2,8	2,8	2,8	2,7
moeilijkheid	2,1	2,2	2,4	2,5	2,1	2,2	2,4	2,4	2,2	2,2	2,6	2,6
zelfvertrouwen	3,1	3,1	3,0	2,9	3,1	3,0	3,0	2,9	3,2	3,2	3,0	3,0
verschillen j-m	1,8	1,7	1,7	1,9	1,8	1,7	1,7	1,8	1,8	1,7	1,8	1,9
duidelijkheid	3,0	2,9	2,9	2,9	3,0	2,9	2,9	2,9	3,0	3,0	3,0	3,0
beroepen feitelijk	1,6	1,5	1,6	1,6	1,6	1,5	1,6	1,6	1,5	1,6	1,5	1,6
beroepen zou willen	2,3	2,3	2,4	2,3	2,3	2,2	2,4	2,3	2,4	2,4	2,4	2,3
nut feitelijk	2,4	2,4	2,3	2,3	2,4	2,4	2,3	2,3	2,4	2,4	2,3	2,3
nut zou willen	2,9	2,9	2,9	2,9	2,8	2,9	2,9	2,9	3,0	3,0	2,9	2,9
proeven feitelijk	2,6	2,6	2,4	2,4	2,6	2,6	2,5	2,5	2,5	2,6	2,1	2,2
proeven zou willen	3,4	3,3	3,2	3,1	3,4	3,4	3,3	3,2	3,3	3,2	3,0	3,0
woorden feitelijk	2,2	2,1	2,0	2,0	2,2	2,1	2,0	1,9	2,1	2,0	2,1	2,0
woorden zou willen	2,6	2,6	2,5	2,5	2,6	2,6	2,5	2,4	2,6	2,6	2,6	2,5
cijfer wiskunde	7,6	7,1	7,1	6,8	7,6	7,0	7,1	6,8	7,5	7,2	7,2	6,7
cijfer nederlands	7,4	7,1	7,0	6,9	7,5	7,1	7,1	6,9	7,3	7,1	7,0	6,8
cijfer natuurkunde	-	7,2	7,1	6,7	-	7,2	7,1	6,7	7,5	7,3	7,1	6,6

Naast het verloop in gemiddelde per variabele over de tijd zijn ook de intercorrelaties per variabele tussen de meetmomenten nagegaan. Deze blijken steeds een simplex-structuur te bezitten (hoe groter de afstand in tijd, hoe lager de correlatie). De interesse toont zich, zo bezien, het meest stabiel, en heeft zelfs van eerste op vierde meetmoment een vrij hoge correlatie (.59). Vrij lage correlaties van eerste op derde en/of vierde meetmoment zien we vooral bij variabelen die de aard van het onderwijs beschrijven, hetgeen inderdaad van jaar tot jaar kan verschillen (afhankelijk van schoolboek en docent), en bij het keuzevoornemen. De relevantie, per meetmoment de sterkste predictor van de keuzevariabele, blijkt over meerdere meetmomenten bezien niet uitgesproken stabiel (correlaties ongeveer .35); kennelijk gaan toegekende relevantie en keuzevoornemen samen op en neer.

Naar aanleiding van deze uitkomsten zijn de eerder gerapporteerde logistische regressie-analyses, met als criteriumvariabele de uiteindelijke keuze (zie 5.3.1), nog driemaal uitgevoerd, nu met de waarden van de predictoren op de eerste drie meetmomenten. Daaruit blijkt dat, zo geanalyseerd, de voorspelbaarheid van de keuze in de loop van de tijd niet duidelijk toeneemt.

Tenslotte is de stabiliteit van de keuzevariabele zelf nader geanalyseerd. Eerst is berekend hoeveel procent van de leerlingen constant blijft dan wel wisselt in de keuze voor of niet voor natuurkunde. Daartoe is de variabele keuzevoornemen, die is gemeten op een vijfpuntsschaal, teruggebracht tot een variabele met drie waarden: 'nee' (zeker of waarschijnlijk niet), '?' (misschien, weet ik nog niet) en 'ja' (waarschijnlijk of zeker wel). Per cohort per geslacht per meet-

moment zijn overgangsmatrices gemaakt. Daaruit blijkt (opnieuw, verg. 5.3.3) dat de keuzevariabele in de tijd niet stabiel is. Een 'nee' lijkt gemiddeld definitiever te zijn dan een 'ja', vooral bij de jongens. Meisjes switchen duidelijk vaker dan jongens van 'nee' naar 'ja' en minder vaak van 'ja' naar 'nee'.

Vervolgens is nagegaan hoeveel procent van de leerlingen (jongens, meisjes) steeds, vanaf 1e tot en met 4e meetmoment, de keuze 'nee' maakte, hoeveel steeds 'ja' koos, en hoeveel ooit (minimaal één keer) radicaal wisselde van 'nee' naar 'ja' of andersom. Als aparte categorie is onderscheiden de 'twijfelaars'; dit zijn de leerlingen die de categorie 'misschien wel, weet ik nog niet' hebben gebruikt, exclusief de leerlingen die hun keuze ooit radicaal veranderden van 'nee' in 'ja' of andersom. De uitkomsten staan in tabel 3.

Tabel 3: Percentages vaste keuzen, veranderingen van keuzen, en twijfelaars

	cohort 2e-3e klas			cohort 3e-4e klas		
	allen	jongens	meisjes	allen	jongens	meisjes
steeds nee	6	2	9	2	1	3
steeds ja	14	23	6	39	54	21
ooit nee-ja	10	9	10	5	3	8
ooit ja-nee	14	13	14	25	16	36
twijfelaars	58	54	61	29	27	32

Bij het verschil tussen de cohorten dat hier zichtbaar wordt, moeten we bedenken dat het eerste cohort de 2e en 3e klas bestrijkt en het tweede cohort de 3e en 4e klas.

We constateren dat de meerderheid van de leerlingen op z'n minst heeft lopen twijfelen, en dat zelfs ongeveer een kwart tenminste éénmaal radicaal van keuze is veranderd (nee-ja of ja-nee).

7. CONCLUSIES EN DISCUSSIE

(Vraag A) Voor beide geslachten en beide keuzegroepen geldt dat de mate waarin in het natuurkundeonderwijs proeven een rol spelen, afneemt; de aan natuurkunde toegekende moeilijkheid toeneemt; de behaalde cijfers afnemen; en de waargenomen verschillen tussen de geslachten in affiniteit voor natuurkunde op het laatste keuzemoment toenemen. Bijna alle bevindingen bij de vergelijking tussen kiezers en niet-kiezers over alle leerlingen, blijken ook te gelden indien per keuzegroep voor jongens en meisjes apart bekeken. Er zijn geen aanwijzingen voor verschillen in niveau of verloop tussen kiezers en niet-kiezers die specifiek zouden zijn voor jongens of meisjes. Ook zijn er geen consistent tussen de geslachten of de keuzegroepen verschillende *samenhangen* tussen de onderzoeksvariabelen.

(Vraag B1) De volgende variabelen zijn vrij duidelijke voorspellers gebleken van het al dan niet kiezen in het VWO van natuurkunde als eindexamenvak: relevantie, waardering, zelfvertrouwen en interesse. (Vraag C1) De variabele interesse toont al op het eerste meetmoment relatief duidelijke verschillen (tussen jongens en meisjes, tussen uiteindelijke kiezers en niet-kiezers, tussen individuele leerlingen) en is op dat moment de beste voorspeller; op volgende meetmomenten raakt deze variabele steeds meer met de andere variabelen verweven. Aan het begin van de tweede klas is de keuzevariabele voor 40% voorspelbaar uit de in dit onderzoek betrokken predictoren, op de volgende meetmomenten steeds voor 60%.

(Vraag B2) Er zijn geen significante verschillen geconstateerd tussen de methodes in de mate waarin jongens en meisjes natuurkunde kiezen. Ook de deskundigenoordelen over de methodes deden er niet veel toe. Wel zijn door de leerlingen gepercipieerde kenmerken van

belang gebleken: zij nemen verschillen tussen methodes waar in leefwereld- en participatiegerichtheid; hun waardering voor het natuurkunde-onderwijs hangt daarmee samen (en deze waardering is een voorspeller van de keuze voor het vak); een grote meerderheid van de leerlingen wil graag dat deze kenmerken meer voorkomen dan nu het geval is; en leerlingen die voldoende aandacht daarvoor hebben gemist, kozen gemiddeld minder vaak voor natuurkunde.

(Vraag C2) De keuzevariabele zelf is niet stabiel. De meeste leerlingen hebben op z'n minst lopen twijfelen, en een kwart is tenminste éénmaal radicaal van keuze veranderd.

Evenals in MAVO en HAVO (zie Jörg e.a., 1990) blijkt ook in het VWO de aan natuurkunde toegekende relevantie de variabele te zijn die het sterkste samenhangt met het kiezen van dat vak, blijken leerlingen nogal eens te twijfelen en van keuze te veranderen, en zijn geen significante verschillen tussen scholen geconstateerd. In het VWO lijken variabelen met betrekking tot het kunnen van de leerlingen, zoals moeilijkheid en cijfers, en de gegeven adviezen, een minder grote rol te spelen dan in MAVO en HAVO.

Het belang van het verwachte nut van het vak voor later, het plezier in of de waardering voor het vak, en het aankunnen van het vak, bleek ook uit eerdere onderzoeken (Kelly, 1988; Alting en Pelgrum, 1990). Daarbij is blijkens zowel eerder onderzoek (zie bijv. Van Heugten, 1993) als onderhavig onderzoek het waargenomen nut van natuurkunde (voor studie, beroep, werk) het meest bepalend.

Net als bij Willems en De Grip (1993) bleek ook in ons onderzoek (in het VWO) dat leerlingen in de 90-er jaren in Nederland gemiddeld weinig meer van mening zijn dat tussen de geslachten grote verschillen bestaan in hun affiniteit met een vak als natuurkunde. Toch bleek ook dat al bij aanvang van het natuurkunde-onderwijs (op gemiddeld 13-jarige leeftijd) duidelijk sprake is van een gemiddeld verschil tussen jongens en meisjes in interesse voor natuurkunde. Zulke verschillen in interesse zijn legitiem. Beleid dat individuele keuzes wil beïnvloeden zal zich grondig moeten kunnen legitimeren. Men kan zich afvragen in hoeverre het economische belang van exacte en technische kennis voldoende legitimatie biedt, waar het zo is dat de aan beroeps- en bedrijfssectoren toegekende relevantie in de tijd wisselt (Wijers, 1987), en plausibel is dat de waarde van kennis en vaardigheden ten aanzien van bijvoorbeeld talen en het omgaan met en organiseren van mensen niet zal afnemen, eerder toenemen.

Het actuele beleid in het voortgezet onderwijs in Nederland op het terrein van loopbaanoriëntatie en loopbaanbegeleiding vraagt om aandacht voor de leerling als actieve actor die zelf aan keuze-opties betekenis moet kunnen geven. Dit uitgangspunt weerspiegelt de in dit onderzoek betrokken kenmerken van leefwereld- en participatiegerichtheid, en past goed bij actuele opvattingen over leren als actief en constructief proces.

NOTEN

1. De acht methodes zijn: Proef op de som, Eureka, Exact natuurkunde, Natuurkunde DBK, Natuurkunde in thema's, Natuurkunde overall, Natuurkunde voor nu en straks, Systematische Natuurkunde.

LITERATUUR

- Ainley, J. (1993). Participation in science courses in senior secondary school. *Research in Science & Technology Education* 11 (2), 207-223.
- Alting, A. en H. Pelgrum. (1990). SISS-ers in science: een aanvullende studie naar sekseverschillen in prestaties in en houding tegenover natuurwetenschappen bij leerlingen in derde klassen HAVO-VWO en MAVO. In: E. van Eck en M. Volman (red.), *Determinanten van seksespecifieke keuzes en prestaties: Symposium-verslag*. Amsterdam: SCO, pp. 23-44.
- Belson, W.A. (1981). *The design and understanding of survey questions*. Aldershot: Gower.

- Boer, P. den, H. Guldmond. (1996). *De effecten van vakkenpakketkeuze en omwegen in het voortgezet onderwijs op het studiesucces in het hoger onderwijs*. Groningen: GION.
- Boone, W.J., M.K. Roth. (1992). Organizing school science shows. *The Physics Teacher* 30, 348-350.
- Bosker, R.J., H.P.J.M. Dekkers. (1994). School differences in producing gender-related subject choices. *School Effectiveness and School Improvement* 5 (2), 178-195.
- Brown, S.D., R.W. Lent. (1996). A social cognitive framework for career choice counseling. *The Career Development Quarterly* 44, 354-366.
- Bügel, K. (1993). *Sekseverschillen in tekstbegrip bij moderne vreemde talen*. Groningen: proefschrift.
- Crawley, F.E., C.B. Black. (1992). Causal modeling of secondary science students' intentions to enroll in physics. *Journal of Research in Science Teaching* 29 (6), 585-599.
- Dekkers, H. (1993). Determinanten van seksspecifieke vakkenkeuzes in HAVO en VWO. *Pedagogische Studiën* 70, 445-459.
- Dekkers, H., E. Smeets. (1997). Van vakkenkeuze naar eindexamen: sekseverschillen? *Pedagogische Studiën* 74, 108-116.
- Eccles, J.S. (1987). Gender Roles and Women's Achievement-related Decisions. *Psychology of Women Quarterly* 1987 II, pp. 135-172.
- Eccles, J.S., B. Barber, K. Updegraff & K.M. O'Brien. (1995). *An expectancy-value model of achievement choices: the role of self-concepts, perceived task utility and interest in predicting activity choice and course enrollment*. San Francisco: AERA.
- Erickson, F., J. Shultz. (1992). Students' experience of the curriculum. In: P.W. Jackson (Ed.). *Handbook of research on curriculum*. New York: Macmillan, 465-485.
- Fensham, P.J. (1992). Science and technology. In: P.W. Jackson (Ed.). *Handbook of research on curriculum*. New York: Macmillan, 789-829.
- Gallagher, S.A. (1994). Middle school classroom predictors of science persistence. *Journal of Research in Science Teaching* 31 (7), 721-734.
- Gottfredson, L.S. (1981). Circumscription and compromise: a developmental theory of occupation aspirations. *Journal of Counseling Psychology Monograph* 28 (6), 545-579.
- Grote, M. (1994). Recruiting students for high-school physics classes. *The Physics Teacher* 32, 350-351.
- Harding, J., L.H. Parker. (1995). Agents for change: policy and practice toward a more gender-inclusive science education. *International Journal for Science Education* 17 (4), 537-553.
- Hellman, W. (1992). From High-School to College Physics. *The Physics Teacher* 30, 24-25.
- Heppner, M.J., F. Hendricks. (1995). A process and outcome study examining career indecision and indecisiveness. *Journal of Counseling and Development* 73, 426-437.
- Heugten, J. van. (1993). *Vakkenpakketkeuze op het VWO; attitudes tegenover een B-pakket en een technische studie*. Eindhoven: TU-OCTO.
- Jones, M.G., J. Wheatley. (1990). Gender differences in teacher-student interactions in science classrooms. *Journal of Research in Science Teaching* 27 (9), 861-874.
- Jörg, T. (1994). *De keuze van het vak natuurkunde als examenpakket en de wijze waarop die tot stand komt bij leerlingen in het MAVO en HAVO*. Utrecht: RUU.
- Jörg, T., Man in 't Veld, M., Wubbels, Th. en Verwey, P. (1990). *Oorzaken van de geringe populariteit van het vak natuurkunde als examenpakket bij meisjes in het MAVO en HAVO*. Utrecht: Cdb.
- Kelly, A. (1988). Option Choice for Girls and Boys. *Research in Science & Technology Education* 6 (1), 5-23.
- Krumboltz, J.D. (1994). Improving career development theory from a social learning perspective. In: M.L. Savickas, R.W. Lent (eds.). *Convergence in career development theories*. Palo Alto, CA: CPP Books, 9-31.
- Kuiper, W.A.J.M. (1993). *Curriculumvernieuwing en lespraktijk; een beschrijvend onderzoek op het terrein van de natuurwetenschappelijke vakken in het perspectief van de basisvorming*. Enschede: proefschrift.
- Lapp, D.R. (1996). Multiplying physics enrollment - strategies that work. *The Physics Teacher* 34, 102-103.
- Lent, R.W., S.D. Brown, G. Hackett. (1994). Toward a unifying social cognitive theory of career and academic interest, choice and performance. *Journal of Vocational Behavior* 45, 79-122.
- Man in 't Veld, M., T. Jörg en Th. Wubbels. (1990). De afweging van meisjes op het MAVO en HAVO om natuurkunde al dan niet op te nemen in het vakkenpakket. In E. van Eck en M. Volman (red.), *Determinanten van seksspecifieke keuzes en prestaties: Symposiumverslag*. Amsterdam: SCO, pp. 45-74.
- Marion, S.F., T. Coladarsi. (1993). *Gender differences in science course-taking patterns among college undergraduates*. Atlanta: AERA.
- Mason, C.L., J.B. Kahle & A.L. Gardner. (1991). Draw-a-scientist test: future implications. *School Science and Mathematics* 91 (5), 193-198.
- McCarthy, A.C., G.D. Moss. (1994). A comparison of male and female pupil perceptions of technology in the curriculum. *Research in Science & Technology Education* 12 (1), 5-13.

- Meece, J.L., J.E(ccl)es Parsons, C.M. Kaczala, S.B. Goff & R. Futterman. (1982). Sex differences in math achievement: toward a model of academic choice. *Psychological Bulletin* 91 (2), 324-348.
- Mellink, E., E. Voncken, Y. Bosma, P. Karstanje. (1993). *Functioneren en rendement van het schooldecaanat*. Amsterdam: Swets & Zeitlinger.
- Meijers, F., C. Piggott. (1995). Career guidance and ethnic minorities in Holland and Britain: confronting fear and anger. *British Journal of Guidance and Counseling* 23 (1) 53-67.
- Oakes, J. (1990). Opportunities, achievement and choice: women and minority students in science and mathematics. *Review of Research in Education* 60, 153-222.
- Paridon, C. van, en A. Vlaanderen. (1992). De toekomst van lbo- en mavo-meisjes. Toekomstplannen van meisjes zijn nog steeds traditioneel. *Didactief januari 1992*, pp. 7-8.
- Prawat, R.S. (1989). Promoting Acces to Knowledge, Strategy, and Disposition in Students: A Research Synthesis. *Review of Educational Research* 59(1)1-41.
- Prawat, R.S. (1991). The value of ideas: The immersion approach to the development of thinking. *Educational Researcher* 20 (2) 3-10, 30.
- Rennie, L.J., L.H. Parker. (1993). Curriculum reform and choice of science: consequences for balanced and equitable participation and achievement. *Journal of Research in Science Teaching* 30 (9), 1017-1028.
- Schaeffer, R. (1991). Science credits, don't leave school without them! *The Science Teacher* 58 (9), 32-35.
- Smits, H. (1994). Leerling laat zich bij studiekeuze vooral door klas leiden. *NRC/Handelsblad 13 januari 1994*.
- Sobolewski, S.J., R.L. Doran. (1996). Replication of a path analysis model of secondary physics enrollments: 20 years later. *Journal of Research in Science Teaching* 33 (5), 501-512.
- Spiering, H. (1994). Talenknobbel wordt niet uitgedaagd. *NRC/Handelsblad 19 mei 1994*.
- Spijkerman, R. (1992). Studie- en beroepskeuzerijpheid. *Handboek voor School- en Beroepskeuzebegeleiding*. Alphen a/d Rijn: Samsom, 1300-1 t/m 30.
- Stokking, K.M. (1995). *Het kiezen van natuurkunde als examenvak in het VWO, verschillen tussen jongens en meisjes?* Utrecht: ISOR.
- Stokking, K.M. (1997). *Determinanten van school- en beroepskeuze in theorie en onderzoek*. Utrecht: ISOR.
- Super, D.E. (1980). A life-span, life-space approach to career development. *Journal of vocational behavior* 16, 282-298.
- Taber, K.S. (1991). Gender differences in science preferences on starting secondary school. *Research in Science & Technology Education* 9 (2), 245-251.
- Vasquez-Abad, J., L. Winer & J-R. Derome. (1996). *Why some stay: a study of factors contributing to persistence in undergraduate physics*. New York: AERA.
- Walberg, H.J. (1991). Improving school science in advanced and developing countries. *Review of Educational Research* 61 (1), 25-69.
- Wierstra, R.E.A. (1990). *Natuurkundeonderwijs tussen leefwereld en vakstructuur*. Utrecht: CdB.
- Willems, E.J.T.A., A. de Grip. (1993). *Young people and technology. Choice of course of study and profession, evaluation and image with regard to technology*. Den Haag: Ministerie van Economische Zaken.
- Wijers, G.A. (1987). Een ontwikkelingspsychologisch concept van beroepskeuze en arbeidsmarktprognoses. Den Haag: OSA.
- Wijers, G.A., F. Meijers. (1996). Career guidance in the knowledge society. *British Journal of Guidance and Counseling* 24 (2) 185-198.
- Young, D.J. (1995). Effect of the science learning environment on science achievement: a comparison of 12 countries from the IEA second international science study. *Educational Research and Evaluation* 1 (2), 129-158.

De score op een ja/nee-woordenschattoets: correctie voor raden en persoonlijke antwoordstijl¹

Ineke Huibregtse en Wilfried Admiraal

IVLOS, Universiteit Utrecht

(Correspondentie-adres: Ineke Huibregtse, IVLOS, Universiteit Utrecht, Postbus 80127, 3508 TC Utrecht, e-mail: i.huibregtse@ivlos.uu.nl)

ABSTRACT

The use of yes/no-tests seems to be a promising method in measuring the size of the receptive vocabulary knowledge of foreign language learners. In a yes/no-test items each consist of either a word or a pseudo word. Subjects are asked to indicate whether or not they know the meaning of these words. This article attempts to tackle the problem of determining a meaningful score for this type of test, specifically for the 'English as a Foreign Language' vocabulary test (Meara, 1992a). Such a score should contain correction for guessing as well as for subjects' response style. Three possible methods are discussed, but none of these measures appear to have an adequate theoretical foundation. Signal Detection Theory is applied and a new, more accurate measure is suggested. Based on theoretical as well as empirical considerations recommendations are made about the choice for the measure to be used in a yes/no-vocabulary test.

HET TOETSEN VAN WOORDENSCHAT

Opvattingen over tweede-taalverwerving en tweede-taalonderwijs veranderen in de tijd en met die veranderingen varieert het belang dat wordt gehecht aan woordenschat. Het ontwikkelen van methoden voor het meten van woordenschat is desondanks nooit afgebroken (voor een overzicht, zie Read, 1997). Een van de aandachtspunten in die ontwikkeling is het meten van de omvang van de woordenschat. Een mogelijkheid hiervoor, wat betreft de receptieve woordenschat, is het gebruik van een ja/nee-toets.

Een ja/nee-woordenschattoets heeft twee soorten items: echte woorden en pseudoworden. Pseudoworden zijn letterreeksen die voldoen aan de fonologische regels van een taal en dus wel woorden zouden kunnen zijn, maar die in die taal geen bestaande betekenisdragende woorden zijn. De toets, die bestaat uit een lijst samengesteld uit woorden en pseudoworden, wordt visueel aangeboden. Bij elk item moeten de proefpersonen aangeven of ze de betekenis van het betreffende woord kennen. Elk item wordt op deze manier beantwoord met "ja" of "nee". Het is de proefpersonen bekend dat de toets woorden bevat die niet bestaan, maar niet hoeveel woorden het betreft en waar ze staan.

Het maken van een ja/nee-toets kost weinig tijd per item, waardoor het mogelijk is relatief veel items te gebruiken en daarmee een betrouwbare meting van de omvang van de woordenschat te doen. Een nadeel van een ja/nee-woordenschattoets betreft een mogelijke geringe ecologische validiteit. In de toets worden woorden zonder context aangeboden, terwijl een taalgebruiker woorden zelden in isolatie leert of gebruikt. De toets zal daardoor vaak een onderschatting van de werkelijke woordkennis opleveren.

Ja/nee-toetsen lijken een bruikbare werkwijze te bieden voor het meten van de omvang van

receptieve woordenschat. De beste manier om op basis van het antwoordgedrag van proefpersonen een toetscore te bepalen is door de aard van de toets niet bij voorbaat duidelijk. Het optellen van het aantal correcte antwoorden lijkt niet voor de hand te liggen. Omdat er twee soorten items en twee mogelijke antwoorden ("ja" en "nee") zijn, zijn er immers vier mogelijke stimulus-responscombinaties. In Figuur 1 wordt de bijbehorende stimulus-responsmatrix weergegeven. De stimulus-alternatieven worden aangeduid met kleine letters, w voor 'woord' en p voor 'pseudowoord'. De respons-alternatieven zijn 'ja' (J) en 'nee' (N). Uit de figuur wordt duidelijk dat er twee soorten correcte antwoorden mogelijk zijn: een "ja"-antwoord bij een echt woord (*treffer*) en een "nee"-antwoord bij een pseudowoord (*terechte verwerping*). Incorrecte antwoorden zijn een "nee"-antwoord bij een echt woord (*misser*)² en een "ja"-antwoord bij een pseudowoord (*loos alarm*).

Figuur 1: Respons-alternatief

		J	N
Stimulus-alternatief	w	$P(J w)$ Treffer	$P(N w)$ Misser
	p	$P(J p)$ Loos alarm	$P(N p)$ Terechte verwerping

Een voorbeeld van een ja/nee-toets voor woordenschat is de 'English as a Foreign Language Vocabulary Test' (EFL-test), ontwikkeld door Meara (1992a) voor het meten van de receptieve woordenschat Engels van vreemde-taalleerders. De EFL-test bestaat uit een groot aantal toetsen op zes verschillende niveaus met oplopende moeilijkheidsgraad. De niveaus zijn gebaseerd op lijsten van woorden geordend naar de frequentie van voorkomen in Engelse teksten (Hindmarsh 1982; Nation, 1986). Het eerste niveau meet de kennis van de duizend meest frequent voorkomende woorden, het tweede niveau bestrijkt de duizend daarop volgende woorden, enzovoort. Voor ieder frequentiegebied bestaan tien parallele tests met 60 items: 40 echte woorden (een aselechte steekproef uit het betreffende frequentiegebied) en 20 pseudoworden. De pseudoworden zijn geconstrueerd door afzonderlijke lettergrepen van woorden uit het betreffende frequentiegebied willekeurig achter elkaar te zetten. De pseudoworden zijn vervolgens door native speakers gecontroleerd op hun fonologische kenmerken. Wanneer ze voldeden aan de fonologische regels werd hun woordsoort bepaald (met name 'zelfstandig naamwoord' of 'werkwoord'). De pseudoworden zijn vervolgens overeenkomstig de samenstelling van het betreffende frequentiegebied ingevoegd, waarbij ook rekening is gehouden met het aantal lettergrepen van de woorden.

Omdat de EFL-test in verschillende onderzoeken wordt gebruikt is het belangrijk een goede methode te hanteren voor het berekenen van een zinvolle score op een dergelijke toets. Meara (1992b) stelt een score (Δm) voor die is gebaseerd op de 'Theory of Signal Detection', een theorie die een veelbelovend kader biedt voor het vaststellen van een adequate score op een ja/nee-toets. De score Δm heeft twee opvallende kenmerken die discussie kunnen oproepen. Ten eerste nadert de waarde van de score bij minder goede prestaties al heel snel tot nul,

ook als de score nog ver boven kansniveau ligt. Ten tweede kunnen kleine verschillen in gedrag grote verschillen in scores veroorzaken, vooral wanneer het aantal "ja"-antwoorden op echte woorden klein is. Wanneer bijvoorbeeld in de helft van de gevallen "ja" wordt gezegd bij een echt woord en in geen enkel geval bij een pseudowoord, bedraagt de score 0,50. Wanneer één keer ook "ja" wordt gezegd bij een pseudowoord, daalt de score tot 0,37. Het verschil tussen deze twee scores bedraagt meer dan tien procent van het totale scorebereik en dat lijkt erg veel voor een verschil van één keer onterecht "ja" zeggen. Naar aanleiding van deze eigenschappen van de maat Δm , rijst de vraag of de door Meara voorgestelde scoreberekening accuraat is en of er een andere en betere manier mogelijk is voor het berekenen van de score op een ja/nee-woordenschattoets.

De berekening van de score op een ja/nee-toets wordt allereerst bepaald door de aard van de toets. Een ja/nee-toets kent twee soorten correcte antwoorden en twee soorten incorrecte antwoorden die verschillende betekenissen kunnen hebben voor het meten van de omvang van de woordenschat. Daarnaast, omdat een ja/nee-toets in feite een meerkeuzetoets met twee antwoord-alternatieven is (zoals een goed/fout- of waar/ onwaar-toets), moet rekening worden gehouden met factoren die invloed kunnen hebben op het antwoordgedrag van de proefpersoon op een dergelijke toets. Met name twee variabelen zijn hierbij van belang. De eerste betreft *raden*. De proefpersoon kan steeds kiezen uit twee antwoord-alternatieven en kan bij twijfel gokken. De kans op een correct antwoord is aanzienlijk, namelijk vijftig procent bij een volledig willekeurig antwoord. In de praktijk zal het in geval van een woordenschattoets vrijwel niet voorkomen dat een proefpersoon bij twijfel een willekeurig antwoord geeft. Woordkennis is zelden alles of niets. Het kan immers zo zijn dat de proefpersoon het woord wel denkt te kennen, maar niet helemaal zeker is van de betekenis. Er kan sprake zijn van een bepaalde gradatie van kennis en in dat geval zal de proefpersoon geen zuiver gokgedrag vertonen (zie Anderson & Freebody, 1983). Het gaat dus niet alleen om raden, maar met name om *beredeneerd raden*. De tweede intermedieërende variabele is de *persoonlijke antwoordstijl* of 'response style' van een proefpersoon (zie Nunnally & Bernstein, 1994). Bij twijfel zal de ene proefpersoon eerder geneigd zijn "ja" te zeggen dan "nee", terwijl een andere proefpersoon behoudender is en pas "ja" zal zeggen als er geen twijfel bestaat. Dat heeft gevolgen voor het antwoordgedrag bij zowel echte woorden als pseudoworden. Iemand met een behoudende antwoordstijl zal niet snel "ja" zeggen bij een pseudowoord maar ook niet bij een echt woord.

Dit betekent dat de berekening van een score op een ja/nee-toets moet voldoen aan verschillende voorwaarden. In de score moet rekening worden gehouden met het feit dat:

- er twee verschillende soorten correcte en twee verschillende soorten incorrecte stimulus-responscombinaties zijn;
- de proefpersoon de mogelijkheid heeft bij twijfel (beredeneerd) te raden, en
- de proefpersonen verschillen vertonen wat betreft persoonlijke antwoordstijl.

In het nu volgende worden vier mogelijke berekeningsmethoden voor de score op een ja/nee-toets besproken: 'het aantal correcte antwoorden', 'correction for guessing', de Δm van Meara (1992b) en een nieuwe maat die ook is afgeleid uit de Signaal Detectie Theorie. Het doel is om op grond van bovengenoemde criteria te komen tot een beslissing over de meest adequate berekeningsmethode voor de toetsscore.

BEREKENINGSMETHODEN VOOR DE SCORE OP EEN JA/NEE-TOETS

Het aantal correcte antwoorden

De meest eenvoudige methode is het sommeren van het aantal correcte antwoorden. Correcte antwoorden zijn de treffers en de terechte verwerpingen van pseudoworden (zie Figuur 1). In de score die hieruit volgt wordt geen rekening gehouden met het feit dat deze twee typen antwoorden niet op dezelfde manier correct zijn en derhalve voldoet deze methode niet aan

de hierboven gestelde eerste voorwaarde. Bij toepassing van deze methode is het mogelijk dat een proefpersoon die op bijna alle items "ja" zegt, en bijvoorbeeld 39 treffers en één terechte verwerping van een pseudowoord heeft, dezelfde score krijgt als iemand die 20 treffers heeft en op geen enkel pseudowoord "ja" zegt en dus ook 20 terechte verwerpingen heeft. Het tellen van het aantal correcte antwoorden biedt geen mogelijkheid onderscheid te maken tussen deze twee zeer verschillende soorten antwoordgedrag en is daarom geen adequate manier voor het berekenen van de score op een ja/nee-toets. Het lijkt zinvoller de proportie treffers te verminderen met de proportie loos-alarmresponsen. De proefpersoon krijgt dan als het ware punten voor treffers en strafpunten voor loos alarm. Deze berekeningsmethode voldoet echter niet aan de derde voorwaarde dat de score moet worden gecorrigeerd voor de persoonlijke antwoordstijl van de proefpersoon. Een proefpersoon die bijvoorbeeld acht treffers (een proportie treffers van 0,20) en nul loos-alarmresponsen heeft, krijgt een score van 0,20. Dergelijk antwoordgedrag wijst op een zeer behoudende antwoordstijl (de proefpersoon zegt zelden "ja") waarmee in deze bepaling van de score geen rekening wordt gehouden. In deze scorebepaling wordt alleen een correctie toegepast zodra iemand risico's neemt en een proportie loos-alarmresponsen heeft die groter is dan nul.

'Correction for guessing'

Deze methode is in het verleden door verschillende onderzoekers toegepast bij de berekening van scores op een ja/nee-woordenschattoets (zie Anderson & Freebody, 1983; en Meara & Buxton, 1987). 'Correction for guessing' is gebaseerd op het 'blind guessing model', een model voor het corrigeren van scores voor raden. Het uitgangspunt is dat er bij elk item sprake is van twee mogelijkheden. De eerste is dat de proefpersoon het correcte antwoord weet, zodat de kans dat het correcte antwoord wordt gegeven gelijk is aan 1. De tweede mogelijkheid is dat de proefpersoon het antwoord niet weet en een willekeurig antwoord geeft. In dat geval is de kans dat de proefpersoon het correcte antwoord geeft gelijk aan $1/K$, waarbij K staat voor het aantal antwoord-alternatieven per item (zie bijvoorbeeld Nunnally en Bernstein, 1994). Green en Swets (1966) maken aannemelijk dat het 'blind guessing model' inhoudt dat bij een ja/nee-toets elk "ja"-antwoord op een pseudowoord (loos alarm) het resultaat is van raden. De geobserveerde proportie loos-alarm-antwoorden is volgens het 'blind guessing model' gelijk aan de proportie van de geobserveerde treffers die correct is geraden. De formule van het 'blind guessing model' is in termen van een ja/nee-toets als volgt weer te geven:³

$$P(h) = P'(h) + P(f)[1 - P'(h)] \quad (1)$$

waarbij

$P(h)$ = de geobserveerde proportie treffers

$P'(h)$ = de ware proportie treffers

$P(f)$ = de proportie loos-alarmresponsen

Herschrijving van vergelijking (1) resulteert in de formule voor 'correction for guessing', waarmee de ware proportie treffers kan worden bepaald (zie Green & Swets, 1966; Anderson & Freebody, 1983):

$$P'(h) = \frac{P(h) - P(f)}{1 - P(f)} \quad (2)$$

Bij het 'blind guessing model' kan een aantal kanttekeningen worden geplaatst. Ten eerste voldoet deze maat niet aan de hierboven gestelde derde voorwaarde: er wordt bij het berekenen van de score geen rekening gehouden met mogelijke persoonlijke voorkeuren voor het

geven van bepaalde antwoorden ('response style'). Ten tweede wordt, als gevolg van de aanname dat wanneer proefpersonen "ja" zeggen, zij *of* het item echt kennen, *of* gokken, geen rekening gehouden met verschillende mogelijke maten van kennis naast perfecte kennis of afwezigheid van kennis en dus met beredeneerd raden (de tweede gestelde voorwaarde). Er wordt uitsluitend gecorrigeerd voor zuiver gokgedrag. Ten derde heeft de berekeningsformule een eigenaardigheid die niet zo gemakkelijk is op te lossen. Wanneer de geobserveerde proportie treffers gelijk is aan één, dan is de gecorrigeerde proportie treffers, ongeacht de proportie loos-alarm-antwoorden, ook gelijk aan één. Dit heeft tot gevolg dat iemand die op alle items correct antwoordt ("ja" bij alle echte woorden en "nee" op alle pseudowoorden) dezelfde score krijgt als iemand die "ja" zegt op alle echte woorden en ook op een groot aantal pseudowoorden.

Op grond van de problemen met (de aannames van) het 'blind guessing model', en met name omdat het model niet voldoet aan de voorwaarden dat rekening wordt gehouden met beredeneerd raden en de persoonlijke antwoordstijl van proefpersonen, kan worden geconcludeerd dat de methode 'correction for guessing' niet geschikt is voor het berekenen van de score op een ja/nee-toets voor woordenschat.

Signaal Detectie Theorie: Meara's Δm

De Signaal Detectie Theorie (SDT) levert een geavanceerd model voor het corrigeren voor gokgedrag (Nunnally & Bernstein, 1994). In dit model wordt rekening gehouden met beredeneerd raden door de aanname dat elk antwoord-alternatief van een item een bepaalde geloofwaardigheid heeft en dat de proefpersoon het antwoord-alternatief met de grootste geloofwaardigheid zal kiezen. Een dergelijk model biedt dus een alternatief voor het 'alles-of-niets'-principe van 'correction for guessing' en voldoet daarmee aan de tweede voorwaarde waaraan de berekening van de toetscore moet voldoen.

De SDT wordt niet vaak gebruikt in de toegepaste taalwetenschap en de bruikbaarheid ervan is weinig besproken. De theorie wordt met name toegepast in experimenten die met woordherkenning te maken hebben (Hoshino, 1991; MacLeod & Kampe, 1996) en er zijn ook voorbeelden te vinden van toepassingen op het gebied van woordkennis (Phillips & Grodsky, 1985; Behrend, 1988). Meara (1992b) stelt voor de SDT te gebruiken voor de berekening van de score op de EFL-test en heeft hiervoor een berekeningsformule afgeleid.

Centrale begrippen in de SDT. Het antwoordgedrag van een proefpersoon op een ja/nee-toets kan worden beschreven aan de hand van een stimulus-responsmatrix zoals in Figuur 1. Aangezien er per rij van deze matrix slechts één getal vrij kan worden ingevuld, heeft de matrix twee vrijheidsgraden. Alle informatie over het antwoordgedrag van een proefpersoon kan derhalve worden weergegeven door één punt in een tweedimensionale grafiek: een combinatie van $P(J|w)$ (de proportie treffers) en $P(J|p)$ (de proportie loos-alarmresponsen).

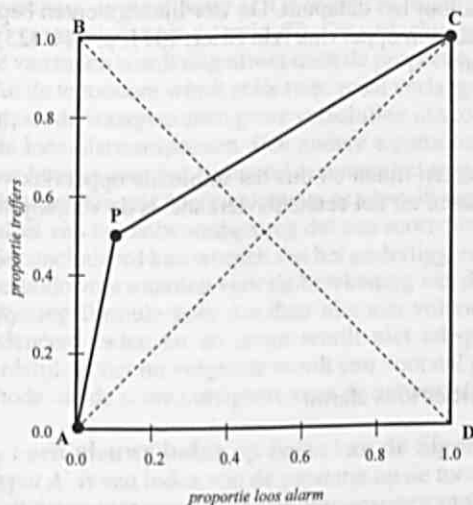
De proportie treffers en de proportie loos alarm en de verhouding tussen die proporties zijn niet alleen afhankelijk van de woordkennis en de persoonlijke antwoordstijl van de proefpersoon, maar ook van de kennis die de proefpersoon heeft over de gevolgen van een bepaald antwoord. Dat betekent dat het antwoordgedrag kan worden beïnvloed door de instructies bij de toets te veranderen, bijvoorbeeld door een beloning aan de respons "ja" te verbinden. Dit zal resulteren in een andere verhouding tussen de proporties treffers en loos alarm en dus in een nieuwe punt in de grafiek. Op die manier kan voor iedere denkbare set van instructies een punt worden gegenereerd. De op deze manier verkregen punten kunnen worden verbonden door een lijn, die de *Receiver-Operating-Characteristic (ROC) curve* wordt genoemd (zie bijvoorbeeld Green & Swets, 1966). Een ROC-curve begint altijd in (0,0) en eindigt altijd in (1,1). Deze twee punten vertegenwoordigen twee extremen van antwoordgedrag: in het geval van punt (0,0) wordt nooit "ja" gezegd en in het geval van punt (1,1) gebeurt dat altijd.

Green en Moses (1966, p. 229-230) en Green en Swets (1966, p. 45-50) hebben aangetoond dat het oppervlak onder de ROC-curve van een ja/nee-taak gelijk is aan het percenta-

ge correcte antwoorden in een bijbehorende geforceerde-keuzetaak met twee alternatieven. Dat laatste houdt in dat een proefpersoon twee stimuli krijgt aangeboden en dat hij of zij moet beslissen welk van de twee stimuli de 'juiste' is. Deze door Green en Swets theoretisch geformuleerde relatie is empirisch geverifieerd in een herkenningstaak door Green en Moses (1966) en ook in een detectietaak met auditieve signalen (Emmerich, 1968, beschreven in Pollack & Hsieh, 1969). Green en Swets (1966) concluderen dat het oppervlak onder de ROC-curve een index is van de onderscheidbaarheid van een signaal en daarmee van de vaardigheid van een proefpersoon signalen te onderscheiden.

Green en Swets (1966) en Pastore en Scheirer (1974) maken aannemelijk dat deze maat non-parametrisch is. Dat is wenselijk in de context van de EFL-test omdat er onvoldoende informatie beschikbaar is over de onderliggende verdelingen van de treffers en de loos-alarm-antwoorden. De maat is gebaseerd op de geometrie van het *eenheidsvierkant* (een vierkant waarvan het oppervlak één eenheid bedraagt). In dit vierkant kan elke combinatie van treffers en loos-alarmresponsen worden uitgezet aan de hand van twee assen, de *x*-as voor loos alarm en de *y*-as voor de treffers. Wanneer de toets één keer met één set van instructies is afgenomen en er derhalve slechts één datapunt beschikbaar is, wordt een curve geconstrueerd die de punten (0,0), (x,y) en (1,1) verbindt. In Figuur 2 wordt een voorbeeld gegeven.

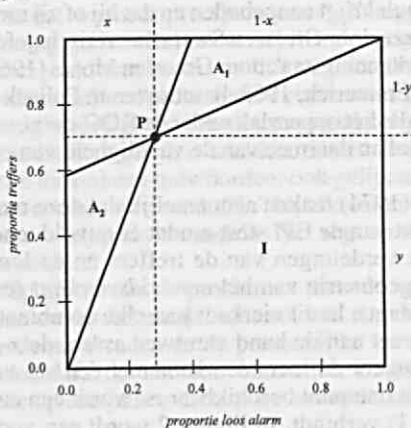
Figuur 2



Als de proportie treffers groter is dan de proportie loos-alarmresponsen zal het datapunt P in de driehoek ABC liggen. Uitsluitend gokken, dus met evenveel kans op treffers als op loos alarm, levert een punt op die op de diagonaal AC ligt. Punt B vertegenwoordigt een proportie treffers van 1,0 en een proportie loos-alarmresponsen van 0,0.

De index komt voort uit het oppervlak van de vierhoek APCD. Pollack en Norman (1964) stellen dat deze maat niet eenvoudigweg gelijk is aan het oppervlak onder de curve, omdat er slechts één datapunt beschikbaar is. De precieze vorm van de curve is derhalve niet bekend. Zij stellen een oppervlak A' voor dat het gemiddelde is van het grootste en het kleinste oppervlak dat met een bepaald punt kan worden geassocieerd.⁴ Grier (1971) beschrijft een manier waarop dit oppervlak kan worden berekend.

Figuur 3



Zoals in Figuur 3 is te zien, geven de doorgetrokken lijnen door (x,y) en $(0,0)$ en door (x,y) en $(1,1)$ twee elkaar niet overlappende driehoeken A_1 en A_2 . Deze twee driehoeken bepalen het bereik van alle mogelijke ROC-curves door het datapunt. De vier lijnsegmenten bepalen de boven- en de ondergrens van het te berekenen oppervlak (zie Grier, 1971, p. 424-425). De door Grier voorgestelde vergelijking is dan:

$$A' = I + \frac{1}{2}(A_1 + A_2) \quad (3)$$

waarin I het gebied is onder de doorgetrokken lijnen en dus het minimale oppervlak vertegenwoordigt. Aan de hand van de coördinaten uit het eenheidsvierkant⁵ is de vergelijking te herschrijven tot:

$$A' = \frac{1}{2} + \frac{(y-x)(1+y-x)}{4y(1-x)} \quad (4)$$

In termen van proporties treffers en proporties loos alarm:

$$A' = \frac{1}{2} + \frac{(h-f)(1+h-f)}{4h(1-f)} \quad (5)$$

Deze vergelijking is uitsluitend zinvol wanneer de proportie treffers (h) gelijk is aan of groter is dan de proportie loos alarm (f). Wanneer de twee proporties gelijk zijn, zijn de oppervlakken van A_1 en A_2 gelijk aan nul. Het betreffende datapunt ligt dan op de positieve diagonaal van het vierkant (lijnstuk AC in Figuur 2). De bijbehorende waarde van A' van 0,5 is de minimale waarde. Het oppervlak onder de ROC-curve heeft dus een bereik van 0,5 (waarbij de kans op een treffer gelijk is aan de kans op een loos alarm) tot 1,0 (waarbij de kans op een treffer gelijk is aan 1 en de kans op een loos alarm gelijk is aan 0).

Meara's Δm . De formule die wordt gehanteerd door Paul Meara bepaalt wat de proportie treffers van de betreffende proefpersoon zou zijn geweest wanneer hij of zij nooit "ja" zou zeggen op een pseudowoord. Meara (1992b) noemt deze maat Δm . De formule die door Meara hiervoor gehanteerd wordt, is:

$$\Delta m = \frac{(h-f)(1+h-f)}{h(1-f)} - 1 \quad (6)$$

Deze formule is een lineaire transformatie van de vergelijking voor A' (4A'-3). Een probleem van deze werkwijze is dat de score niet wordt gecorrigeerd voor de persoonlijke antwoordstijl van de proefpersoon. Het doel van Meara is namelijk de proportie treffers te bepalen die de proefpersoon zou hebben gehad als hij of zij geen loos-alarmresponsen had gegeven: wanneer in de bovenstaande vergelijking $f=0$ wordt ingevuld, dan is $\Delta m=h$. Dat betekent dat de proportie treffers wordt bepaald die de proefpersoon zou hebben gehad als hij of zij uitsluitend "ja" had gezegd als er geen twijfel was en dus extreem behoudend antwoordgedrag had vertoond. Omdat in dat geval bij twijfel altijd "nee" wordt gezegd, is de op deze manier gecorrigeerde proportie treffers waarschijnlijk een onderschatting van de ware proportie treffers. Dit wordt duidelijker aan de hand van een herschrijving van de formule van Meara tot:

$$\Delta m = \frac{(h-f)}{(1-f)} - \frac{f}{h} \quad (7)$$

De eerste term van de formule van Meara blijkt gelijk te zijn aan de formule voor 'correction for guessing' (vergelijking 2). De score Δm is gelijk aan de uitkomst van 'correction for guessing' verminderd met de ratio f/h . De gedachtegang van het 'blind guessing model' dat de mate van raden wordt uitgedrukt door de proportie loos-alarm-antwoorden wordt dus gehandhaafd: de toetsscore wordt zelfs twee maal verlaagd wanneer er loos-alarmresponsen zijn. Dit verklaart de waargenomen grote verschillen in scores bij relatief kleine verschillen in de proportie loos-alarmresponsen. Een andere eigenschap van de formule voor Δm is dat er alleen scores kunnen worden berekend wanneer het oppervlak onder de ROC-curve 0,75 of groter is. De maat Δm zelf heeft weliswaar een bereik van nul tot één, maar dit bereik omvat slechts een deel van het antwoordgedrag dat een score boven kansniveau zou moeten opleveren.

Geconcludeerd kan worden dat het onderliggende model (de SDT) weliswaar voldoet aan de gestelde voorwaarden voor de berekening van de toetsscore, maar dat de uitwerking tot de berekeningsformule voor Δm daar niet aan voldoet. Er wordt geen rekening gehouden met beredeneerd raden en de score wordt niet adequaat gecorrigeerd voor persoonlijke antwoordstijl. In het nu volgende wordt een voorstel gedaan voor een alternatieve berekeningsmethode die de score corrigeert voor de antwoordstijl.

I_{SDT} : een nieuwe index op basis van de Signaal Detectie Theorie

De maat A' is een index van de prestatie op de toets. In deze index is rekening gehouden met het feit dat er vier mogelijke stimulus-responscombinaties zijn door de verhouding tussen de proportie treffers en de proportie loos-alarmresponsen centraal te stellen. Bovendien is rekening gehouden met beredeneerd raden. Het model gaat ervan uit dat ieder antwoord-alternatief een bepaalde geloofwaardigheid heeft en dat de proefpersoon een afweging maakt indien hij of zij twijfelt. In de prestatie-index A' is nog geen rekening gehouden met de persoonlijke antwoordstijl van proefpersonen.

Hodos (1970) beschrijft hoe de antwoordstijl kan worden geëvalueerd. Ook deze analyse is gebaseerd op het eenheidsvierkant waarin de kans op een treffer is uitgezet als functie van de kans op een loos alarm. Een punt op de y-as aan de linkerkant van het vierkant vertegenwoordigt volgens Hodos een maximale neiging om "nee" te zeggen. De kans op een loos alarm is in dat geval gelijk aan nul omdat bij eventuele twijfel altijd "nee" wordt gezegd. Een punt op de x-as aan de bovenkant van het vierkant vertegenwoordigt een maximale neiging om "ja" te zeggen. Bij twijfel wordt "ja" gezegd. Dat betekent dat de kans op een treffer gelijk is aan 1 (de proefpersoon weet het antwoord of twijfelt en zegt "ja") en dat er een bepaalde

kans bestaat op een loos alarm (die dan gelijk is aan de mate waarin de proefpersoon twijfelt over de pseudowoorden). Precies in het midden tussen die assen loopt de negatieve diagonaal van het eenheidsvierkant (lijnstuk BD in Figuur 2). Hodos stelt dat deze een 'unbiased performance' vertegenwoordigt: een punt op deze diagonaal staat voor een verhouding tussen de proportie treffers en de proportie loos-alarmresponsen bij een neutrale antwoordstijl.

Om op grond van een datapunt het bijbehorende punt gecorrigeerd voor persoonlijke antwoordstijl te vinden moet het snijpunt van de ROC-curve van de betreffende proefpersoon met de negatieve diagonaal worden bepaald. Grier (1971, p. 425) geeft de coördinaten van dat punt:

$$(x, y) = \left[\frac{2(1 - A')}{3 - 2A'}, 1 - \frac{2(1 - A')}{3 - 2A'} \right] \quad (8)$$

Invulling van vergelijking (6), de vergelijking voor A' , in het tweede deel van vergelijking (8) levert een berekeningsformule op voor de y -coördinaat:

$$y = 1 - \frac{2h(1 - f) - (h - f)(1 + h - f)}{4h(1 - f) - (h - f)(1 + h - f)} \quad (9)$$

De y -coördinaat is in feite de ware proportie treffers, de proportie treffers die zou zijn behaald wanneer het antwoordgedrag van de proefpersoon niet zou zijn beïnvloed door een persoonlijke antwoordstijl. De waarde van deze y -coördinaat is derhalve bruikbaar als score op de EFL-test. De minimale waarde van deze coördinaat bedraagt 0,5 omdat de onderste helft van de negatieve diagonaal in het gebied van het eenheidsvierkant ligt dat de scores beneden kansniveau vertegenwoordigt. De hier voorgestelde formule voor de score op de EFL-test is een lineaire transformatie ($2y - 1$) van vergelijking (9) die een score oplevert met een bereik van nul tot één. De formule luidt als volgt:

$$I_{SDT} = 1 - \frac{4h(1 - f) - 2(h - f)(1 + h - f)}{4h(1 - f) - (h - f)(1 + h - f)} \quad (10)$$

EEN VERGELIJKING OP BASIS VAN REKENVOORBEELDEN

Het is interessant na te gaan of de nieuwe, hier voorgestelde berekening van de score I_{SDT} ook daadwerkelijk andere waarden oplevert dan de drie andere berekeningswijzen (de proportie treffers verminderd met de proportie loos-alarmresponsen, 'correction for guessing' en de door Meara voorgestelde score Δm). Teneinde een indruk te krijgen van de waarden van de verschillende toetscores bij verschillende verhoudingen tussen de proportie treffers (h) en de proportie loos alarm (f) worden vergelijkingen gemaakt voor een serie combinaties die voorkomen in de gegevens van een onderzoek naar de effecten van tweetalig onderwijs (Engels-Nederlands) in het voortgezet onderwijs (zie Huijbregtse, De Bot, Coleman & Westhoff, 1998). De waarden van de verschillende maten zijn weergegeven in Bijlage 2. Voor verschillende mogelijke proporties treffers (in stappen van 0,10) zijn voor een alle mogelijke proporties loos-alarmresponsen (met intervallen van 0,05) de uitkomsten van de verschillende berekeningsmethoden gegeven.

De maten komen overeen als het gaat om de score van iemand die op kansniveau presteert (wanneer beide proporties gelijk zijn). Bij alle maten is de score gelijk aan nul (of ongelijk

in het geval van Δm). Ook voor de hele goede presteerders (maximale proportie treffers in combinatie met afwezigheid van of een zeer kleine proportie loos-alarmsresponsen) komen de waarden van de verschillende maten sterk overeen. Een proefpersoon behaalt in dit geval een score van 1,0 of een score die de waarde van 1,0 benadert. De maten Δm , 'h-f', en I_{SDT} bereiken de waarde 1,0 uitsluitend wanneer de proefpersoon een volledig correcte score heeft ($h=1$ en $f=0$).

Een uitkomst die in het genoemde onderzoek vaak voorkomt, is een proportie treffers van ongeveer 0,80 en een proportie loos alarm van ongeveer 0,20. In dit specifieke geval geeft de maat 'h-f' dezelfde score als de I_{SDT} . Voor andere, vergelijkbare uitkomsten geeft 'h-f' een score die de waarde van de I_{SDT} dicht benadert. De score volgens Δm is lager dan die van de andere maten. Dit is te verklaren door het feit dat Δm een score oplevert die de betreffende proefpersoon zou hebben bij een extreem behoudende antwoordstijl. De maat Δm levert dus een onderschatting van de ware score op, terwijl de I_{SDT} bij lage proporties loos alarm juist de score verhoogt door te corrigeren voor een behoudende antwoordstijl. De waarde van 'cfg' is hoger dan die van de I_{SDT} . In de maat 'cfg' wordt de proportie treffers sterk gewaardeerd, iets wat het duidelijkst tot uiting komt wanneer de proportie treffers gelijk is aan 1. De score krijgt dan, ongeacht het aantal keren dat een proefpersoon "ja" heeft gezegd op een pseudo-woord, altijd de waarde 1. Een grote proportie treffers levert, met name in combinatie met een redelijk kleine proportie loos alarm bij toepassing van 'cfg' een relatief hoge score op.

Wanneer een proefpersoon op zowel echte woorden als pseudoworden heel vaak "ja" zegt en derhalve beide proporties heel groot zijn, geven de scores 'h-f' en Δm allebei een lagere waarde dan de I_{SDT} . Wanneer de grootten van de beide proporties dicht bij elkaar liggen en de proportie treffers niet echt groot en ook niet echt klein is, heeft Δm geen geldige waarde. Deze scoremaat bereikt bij minder goede prestaties, en met name bij kleine proporties treffers, snel de waarde nul, zelfs wanneer de proefpersoon nog boven kansniveau presteert. Daarentegen is in deze gevallen de waarde van 'cfg' relatief hoog. De maat corrigeert niet voor de vertoende risicovolle antwoordstijl. De waarde van 'h-f' komt in dergelijke gevallen precies of bij benadering overeen met die van de I_{SDT} . Een proefpersoon die slechts in weinig gevallen "ja" zegt, en voor wie dus beide proporties klein zijn, krijgt met de I_{SDT} een hogere score dan volgens de andere berekeningswijzen. Dat is ook te verwachten, omdat bij een extreem kleine proportie loos alarm wordt gecorrigeerd voor een behoudende antwoordstijl. De scores volgens 'cfg' en 'h-f' zijn allebei erg laag, de eerste vanwege de kleine proportie treffers, de tweede vanwege het kleine verschil tussen beide proporties. De maat Δm levert ook in dergelijke gevallen geen geldige score op.

CONCLUSIES EN DISCUSSIE

De score op een ja/nee-toets moet een indicatie zijn van de omvang van de receptieve woordenschat van de proefpersoon. In dit artikel zijn enkele mogelijke manieren besproken om een dergelijke score te bepalen, in het bijzonder voor de 'English as a Foreign Language Vocabulary Test' (EFL-test) van Meara (1992a). Het betreft de maat 'h-f' (de proportie treffers verminderd met de proportie loos alarm), 'correction for guessing' en de door Meara (1992b) voorgestelde maat Δm . Voor al deze maten kan worden gesteld dat ze niet voldoen aan de eisen die aan een score op een ja/nee-toets kunnen worden gesteld: rekening houden met verschillende typen correcte en incorrecte antwoorden, beredeneerd raden en de antwoordstijl van de proefpersoon. Op basis van de Signaal Detectie Theorie (SDT) wordt een nieuwe index, de I_{SDT} , voorgesteld. In de I_{SDT} wordt op grond van de proportie treffers en de proportie loos-alarmsresponsen een score berekend, waarbij rekening wordt gehouden met zowel beredeneerd raden als de antwoordstijl van de proefpersoon.

Uit een vergelijking tussen de verschillende maten op basis van rekenvoorbeelden blijkt dat de maten 'cfg' en Δm in veel gevallen waarden opleveren die verschillen van de waarde

van de I_{SDT} . De maat Δm geeft een onderschatting van de beoogde score, terwijl de maat 'cfg' een overschatting oplevert bij grote proporties treffers. De berekeningswijze 'h-f' levert, ondanks het feit dat de score niet wordt gecorrigeerd voor antwoordstijl, waarden op die vergelijkbaar zijn met en soms zelfs exact gelijk zijn aan de waarden van de I_{SDT} . Uitzonderingen worden gevormd door gevallen waarin de proportie treffers en de proportie loos-alarmsresponsen allebei heel groot of allebei heel klein zijn. In die gevallen is de waarde van 'h-f' kleiner dan die van de I_{SDT} .

Het voordeel van de 'h-f'-maat is dat deze eenvoudig te beschrijven en te berekenen is. Omdat de genoemde extreme vormen van antwoordgedrag in de praktijk niet vaak voorkomen, kan worden gesteld dat de scoremaat 'h-f' een hanteerbare scoremaat is voor kleinschalige toepassingen in de onderwijspraktijk. Een belangrijk nadeel van de maat 'h-f' is dat de score niet wordt gecorrigeerd voor de antwoordstijl van proefpersonen. De I_{SDT} levert een score op die wel is gecorrigeerd voor beredeneerd raden en antwoordstijl en bij elk mogelijk type antwoordgedrag accuraat is. In toepassingen in een onderzoekscontext raden wij dan ook aan deze berekeningswijze te gebruiken.

Enkele kanttekeningen bij de voorgestelde bepaling van de score op een ja/nee-toets aan de hand van de SDT zijn hier op hun plaats. Ze betreffen de aannames van de SDT en de geldigheid daarvan in het geval van een woordenschattoets.

Een van de uitgangspunten van de SDT is dat een reeks items bestaat uit een aantal gelijksoortige signalen waarbij de proefpersoon twee klassen moet onderscheiden. Het is niet vanzelfsprekend dat de woorden in de EFL-test als gelijksoortige items kunnen worden beschouwd en een eventueel verschil in moeilijkheid tussen de items buiten beschouwing kan worden gelaten. Een aanname die ten grondslag ligt aan de toets is dat de frequentie waarmee woorden voorkomen, bepalend is voor de moeilijkheid van een item. Omdat de echte woorden uit een individuele test tot dezelfde frequentieband behoren, kan worden beargumenteerd dat deze als gelijkwaardige items kunnen worden beschouwd. Voor de pseudoworden is dat iets minder duidelijk. Deze zijn echter wel zodanig geconstrueerd dat ze de kenmerken vertonen van de woorden uit de betreffende frequentieband (bijvoorbeeld wat betreft het aantal lettergrepen en de woordsoort).

Een andere, mogelijk problematische aanname van de SDT ligt in de vertaling van de antwoordstijl van proefpersonen naar antwoordgedrag. Zoals beschreven, stelt Hodos (1970) dat bij een maximale neiging om bij twijfel "nee" te zeggen de kans op een loos alarm gelijk is aan nul. Daaraan ten grondslag ligt de veronderstelling dat een proefpersoon bij een pseudoword twee mogelijkheden heeft: met zekerheid "nee" zeggen of twijfelen. De optie van met zekerheid "ja" zeggen op een pseudoword is hiermee uitgesloten. Het is de vraag of deze veronderstelling klopt. Het lijkt echter onwaarschijnlijk dat een proefpersoon met zekerheid een pseudoword zal kennen. Iets vergelijkbaars geldt voor een maximale neiging om "ja" te zeggen. Volgens Hodos is de kans op een treffer in dit geval gelijk aan één, omdat ook bij een echt woord een proefpersoon wordt verondersteld twee mogelijkheden te hebben (met zekerheid "ja" zeggen of twijfelen). Bij een maximale neiging om "ja" te zeggen zal twijfel resulteren in een "ja"-antwoord. Het is echter niet uit te sluiten dat een proefpersoon op een echt woord met zekerheid "nee" zegt, dat wil zeggen, zeker weet dat hij of zij het woord niet kent. Omdat er sprake is van een toetssituatie en de proefpersonen dus prestatiegericht zijn, is de verwachting dat proefpersonen ten minste enige twijfel zullen hebben.

Een derde aanname in de SDT die vragen op kan roepen, is dat scores die beneden kansniveau liggen het gevolg zijn van meetfouten, invulfouten of het met opzet geven van foute antwoorden. In termen van de EFL-test betekent dat, dat geen score wordt toegerekend wanneer een proefpersoon relatief vaker "ja" zegt op een pseudoword dan op een echt woord. Er zijn gevallen denkbaar waarin een dergelijk resultaat eigenlijk tot een geldige score zou moeten leiden. Bij een heel laag niveau van woordkennis in combinatie met een antwoordstijl waarin bij twijfel snel "ja" wordt gezegd, is het mogelijk dat de proportie loos alarm net iets groter is dan de proportie treffers. Een score beneden kansniveau als geldige score beschou-

wen, zou inhouden dat informatie over dergelijke, overigens niet vaak voorkomende gevallen, niet verloren gaat. Tegelijkertijd zijn deze dan niet meer te onderscheiden van situaties waarin invulfouten worden gemaakt of opzettelijk foute antwoorden worden gegeven. Om te voorkomen dat in dergelijke gevallen scores worden toegekend, is het verdedigbaar scores beneden kansniveau buiten beschouwing te laten. De aanname dat een score niet geldig is wanneer de proportie treffers kleiner is dan de proportie loos alarm, heeft tot gevolg dat de gecorrigeerde proportie treffers volgens de nieuwe correctie nooit kleiner is dan 0,5. Ten behoeve van de interpretatie is de I_{SDT} lineair getransformeerd naar een maat met een bereik van nul tot één.

Een vraag die resteert, is hoe de score op de toets precies moet worden geïnterpreteerd. Meara (1992b) stelt dat de score Δm staat voor de proportie van de woorden uit het betreffende frequentiegebied dat door de proefpersoon wordt gekend. Bij de I_{SDT} gaat het om een proportie treffers die is gecorrigeerd voor (beredeneerd) raden en antwoordstijl. De maat is bruikbaar voor het vergelijken van groepen proefpersonen en het uitdrukken van de ontwikkeling van de omvang van de woordenschat. Nader onderzoek is wenselijk om te bepalen wat de score in de praktijk betekent, bijvoorbeeld voor toelating tot of afsluiting van een onderwijsprogramma. Een mogelijkheid hiervoor ligt in onderzoek naar de relatie tussen de score op de ja/nee-toets en andere, reeds genormeerde taalvaardigheidsmaten.

LITERATUUR

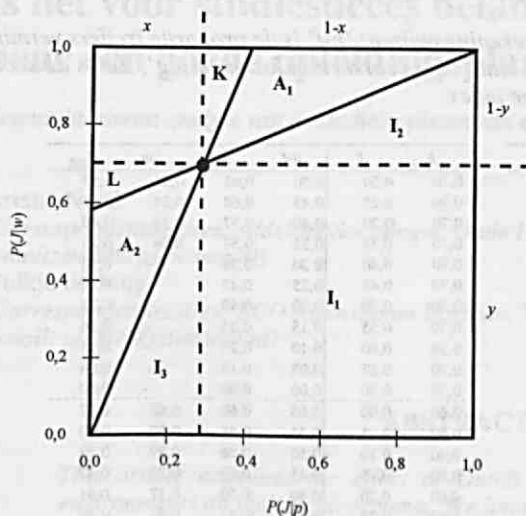
- Anderson, R.C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B.A. Hutson (Ed.), *Advances in reading/language research*. Greenwich, UK: JAI Press Inc.
- Behrend, D.A. (1988). Overextensions in early language comprehension: evidence from a signal detection approach. *Journal of Child Language*, 15, 63-75.
- Emmerich, D.S. (1968). Receiver-operating characteristics determined under several interaural conditions of listening. *Journal of the Acoustical Society of America*, 43, 298-307.
- Green, D.M., & Moses, F.L. (1966). On the equivalence of two recognition measures of short-term memory. *Psychological Bulletin*, 66, 228-234.
- Green, D.M., & Swets, J.A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley and Sons, Inc.
- Grier, J.B. (1971). Nonparametric indexes for sensitivity and bias: computing formulas. *Psychological Bulletin*, 75, 424-429.
- Hindmarsh, R. (1982). *Cambridge English lexicon*. Cambridge: Cambridge University Press.
- Hodos, W. (1970). Nonparametric index of response bias for use in detection and recognition experiments. *Psychological Bulletin*, 74, 351-354.
- Hoshino, Y. (1991). A bias in favor of the positive response to high-frequency words in recognition memory. *Memory and Cognition*, 19, 607-616.
- Huibregtse, I., De Bot, K., Coleman, L., & Westhoff, G. (1998). Een evaluatie van tweetalig voortgezet onderwijs in Nederland. *Toegepaste Taalwetenschap in Artikelen, jaargang 1998(1)*, 219-227.
- MacLeod, C.M., & Kampe, K.E. (1996). Word frequency effects on recall, recognition, and word fragment completion tests. *Journal of Experimental Psychology Learning, Memory, and Cognition*, 22, 132-142.
- Meara, P. (1992a). *EFL vocabulary test*. Swansea, UK: Centre for Applied Language Studies.
- Meara, P. (1992b). *New approaches to testing vocabulary knowledge*. (Draft paper). Swansea, UK: Centre for Applied Language Studies.
- Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing*, 4, 142-154.
- Nation, I.S.P. (1986). *Word lists (revised edition)*. Wellington: Victoria University English Language Centre.
- Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory*. (3rd ed.). New York: McGraw-Hill.
- Pastore, R.E., & Scheirer, C.J. (1974). Signal detection theory: considerations for general application. *Psychological Bulletin*, 81, 945-958.
- Phillips, G.W., & Grodsky, M. (1985, March-April). *Testing Piaget's theory of probability concept development: a Bayesian approach using the theory of signal detection*. Paper presented at the 69th Annual Meeting of the American Education Research Association, Chicago.

- Pollack, I., & Hsieh, R. (1969). Sampling variability of the area under the ROC-curve and of d' . *Psychological Bulletin*, 71, 161-173.
- Pollack, I., & Norman, D.A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, 1, 125-126.
- Read, J. (1997). Assessing vocabulary in a second language. In C. Clapham & D. Corson (Eds.), *Language testing and assessment*. Dordrecht, NL: Kluwer Academic Publishers.

NOTEN

1. De auteurs danken de overige leden van het onderzoeksteam, Kees de Bot, Loes Coleman en Gerard Westhoff voor hun bijdrage. In de periode waarin dit artikel is geschreven, is onverwachts onze collega Erik Schils overleden. Hij heeft bij de ontwikkeling van onze ideeën over de score op de ja/nee-woordenschattoets een belangrijke rol gespeeld en we zullen zijn kundige en enthousiaste inbreng node missen. Dit onderzoek wordt gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).
2. De benaming 'incorrect antwoord' is hier enigszins misleidend. Het betekent niet dat het antwoord fout is, want de proefpersoon kent het betreffende woord niet en zegt ook dat dat het geval is.
3. De symbolen h en f die worden gehanteerd voor respectievelijk treffers en loos alarm zijn afgeleid van de Engelse termen 'hits' en 'false alarms'.
4. De index A' wordt door Pastore en Scheirer (1974) A_g genoemd.
5. Zie Bijlage 1 voor een beschrijving van de afleiding van de vergelijking.

BIJLAGE 1: DE AFLEIDING VAN DE BEREKENINGSFORMULE VAN A'



$$A' = I + \frac{1}{2}(A_1 + A_2) = I_1 + I_2 + I_3 + \frac{1}{2}(A_1 + A_2)$$

$$I_1 + I_2 + I_3 = y(1-x) + \frac{(1-x)(1-y)}{2} + \frac{xy}{2}$$

$$A_1 = \frac{(1-x)(1-y)}{2} - K$$

$$A_2 = \frac{xy}{2} - L$$

De driehoeken K en I_3 zijn gelijkvormig, evenals de driehoeken L en I_2 . De verhoudingen tussen de zijden van twee gelijkvormige rechthoekige driehoeken zijn gelijk. Dat betekent dat de lange zijde van K gelijk is aan $(1-y)$ en de korte zijde van K gelijk is aan $x(1-y)/y$. De lange zijde van L is gelijk aan x en de korte zijde van L is gelijk aan $x(1-y)/(1-x)$.

Het oppervlak van K is dan: $\frac{1}{2}(1-y) \left[\frac{x(1-y)}{y} \right] = \frac{x(1-y)(1-y)}{2y}$

en het oppervlak van L is dan: $\frac{1}{2}x \left[\frac{x(1-y)}{(1-x)} \right] = \frac{x^2(1-y)}{2(1-x)}$

Invulling van alle elementen levert de vergelijking op voor A' :

$$A' = \frac{1}{2} + \frac{(y-x)(1+y-x)}{4y(1-x)}$$

BIJLAGE 2: EEN VERGELIJKING OP BASIS VAN REKENVOORBEELDEN

De uitkomsten van de verschillende berekeningswijzen: 'h-f' is de proportie treffers vermindert met de proportie loos alarm, 'cfg' staat voor 'correction for guessing', Δm is de score van Meara en I_{SDT} staat voor de nieuwe index.

<i>h</i>	<i>f</i>	<i>h-f</i>	<i>cfg</i>	Δm	I_{SDT}	<i>h</i>	<i>f</i>	<i>h-f</i>	<i>cfg</i>	Δm	I_{SDT}
1,00	0,00	1,00	1,00	1,00	1,00	0,70	0,20	0,50	0,63	0,34	0,50
1,00	0,05	0,95	1,00	0,95	0,95	0,70	0,25	0,45	0,60	0,24	0,45
1,00	0,10	0,90	1,00	0,90	0,90	0,70	0,30	0,40	0,57	0,14	0,40
1,00	0,15	0,85	1,00	0,85	0,86	0,70	0,35	0,35	0,54	0,04	0,35
1,00	0,20	0,80	1,00	0,80	0,82	0,70	0,40	0,30	0,50	--	0,30
1,00	0,25	0,75	1,00	0,75	0,78	0,70	0,45	0,25	0,45	--	0,25
1,00	0,30	0,70	1,00	0,70	0,74	0,70	0,50	0,20	0,40	--	0,21
1,00	0,35	0,65	1,00	0,65	0,70	0,70	0,55	0,15	0,33	--	0,16
1,00	0,40	0,60	1,00	0,60	0,67	0,70	0,60	0,10	0,25	--	0,11
1,00	0,45	0,55	1,00	0,55	0,63	0,70	0,65	0,05	0,14	--	0,06
1,00	0,50	0,50	1,00	0,50	0,60	0,70	0,70	0,00	0,00	--	0,00
1,00	0,55	0,45	1,00	0,45	0,57	0,60	0,00	0,60	0,60	0,60	0,67
1,00	0,60	0,40	1,00	0,40	0,54	0,60	0,05	0,55	0,58	0,50	0,60
1,00	0,65	0,35	1,00	0,35	0,51	0,60	0,10	0,50	0,56	0,39	0,53
1,00	0,70	0,30	1,00	0,30	0,48	0,60	0,15	0,45	0,53	0,28	0,47
1,00	0,75	0,25	1,00	0,25	0,45	0,60	0,20	0,40	0,50	0,17	0,41
1,00	0,80	0,20	1,00	0,20	0,43	0,60	0,25	0,35	0,47	0,05	0,36
1,00	0,85	0,15	1,00	0,15	0,40	0,60	0,30	0,30	0,43	--	0,30
1,00	0,90	0,10	1,00	0,10	0,38	0,60	0,35	0,25	0,38	--	0,25
1,00	0,95	0,05	1,00	0,05	0,36	0,60	0,40	0,20	0,33	--	0,20
1,00	1,00	--	--	--	--	0,60	0,45	0,15	0,27	--	0,15
0,90	0,00	0,90	0,90	0,90	0,90	0,60	0,50	0,10	0,20	--	0,10
0,90	0,05	0,85	0,89	0,84	0,85	0,60	0,55	0,05	0,11	--	0,05
0,90	0,10	0,80	0,89	0,78	0,80	0,60	0,60	0,00	0,00	--	0,00
0,90	0,15	0,75	0,88	0,72	0,75	0,50	0,00	0,50	0,50	0,50	0,60
0,90	0,20	0,70	0,88	0,65	0,70	0,50	0,05	0,45	0,47	0,37	0,52
0,90	0,25	0,65	0,87	0,59	0,66	0,50	0,10	0,40	0,44	0,24	0,45
0,90	0,30	0,60	0,86	0,52	0,62	0,50	0,15	0,35	0,41	0,11	0,38
0,90	0,35	0,55	0,85	0,46	0,57	0,50	0,20	0,30	0,38	--	0,32
0,90	0,40	0,50	0,83	0,39	0,53	0,50	0,25	0,25	0,33	--	0,26
0,90	0,45	0,45	0,82	0,32	0,49	0,50	0,30	0,20	0,29	--	0,21
0,90	0,50	0,40	0,80	0,24	0,45	0,50	0,35	0,15	0,23	--	0,15
0,90	0,55	0,35	0,78	0,17	0,41	0,50	0,40	0,10	0,17	--	0,10
0,90	0,60	0,30	0,75	0,08	0,37	0,50	0,45	0,05	0,09	--	0,05
0,90	0,65	0,25	0,71	--	0,33	0,50	0,50	0,00	0,00	--	0,00
0,90	0,70	0,20	0,67	--	0,29	0,40	0,00	0,40	0,40	0,40	0,54
0,90	0,75	0,15	0,60	--	0,24	0,40	0,05	0,35	0,37	0,24	0,45
0,90	0,80	0,10	0,50	--	0,18	0,40	0,10	0,30	0,33	0,08	0,37
0,90	0,85	0,05	0,33	--	0,11	0,40	0,15	0,25	0,29	--	0,30
0,90	0,90	0,00	0,00	--	0,00	0,40	0,20	0,20	0,25	--	0,23
0,80	0,00	0,80	0,80	0,80	0,82	0,40	0,25	0,15	0,20	--	0,17
0,80	0,05	0,75	0,79	0,73	0,76	0,40	0,30	0,10	0,14	--	0,11
0,80	0,10	0,70	0,78	0,65	0,70	0,40	0,35	0,05	0,08	--	0,05
0,80	0,15	0,65	0,76	0,58	0,65	0,40	0,40	0,00	0,00	--	0,00
0,80	0,20	0,60	0,75	0,50	0,60	0,30	0,00	0,30	0,30	0,30	0,48
0,80	0,25	0,55	0,73	0,42	0,55	0,30	0,05	0,25	0,26	0,10	0,38
0,80	0,30	0,50	0,71	0,34	0,50	0,30	0,10	0,20	0,22	--	0,29
0,80	0,35	0,45	0,69	0,25	0,46	0,30	0,15	0,15	0,18	--	0,20
0,80	0,40	0,40	0,67	0,17	0,41	0,30	0,20	0,10	0,13	--	0,13
0,80	0,45	0,35	0,64	0,07	0,37	0,30	0,25	0,05	0,07	--	0,06
0,80	0,50	0,30	0,60	--	0,32	0,30	0,30	0,00	0,00	--	0,00
0,80	0,55	0,25	0,56	--	0,28	0,20	0,00	0,20	0,20	0,20	0,43
0,80	0,60	0,20	0,50	--	0,23	0,20	0,05	0,15	0,16	--	0,29
0,80	0,65	0,15	0,43	--	0,18	0,20	0,10	0,10	0,11	--	0,18
0,80	0,70	0,10	0,33	--	0,13	0,20	0,15	0,05	0,06	--	0,08
0,80	0,75	0,05	0,20	--	0,07	0,20	0,20	0,00	0,00	--	0,00
0,80	0,80	0,00	0,00	--	0,00	0,10	0,00	0,10	0,10	0,10	0,38
0,70	0,00	0,70	0,70	0,70	0,74	0,10	0,05	0,05	0,05	--	0,16
0,70	0,05	0,65	0,68	0,61	0,68	0,10	0,10	0,00	0,00	--	0,00
0,70	0,10	0,60	0,67	0,52	0,62	0,00	0,00	--	--	--	--
0,70	0,15	0,55	0,65	0,43	0,56						

Is het voor studiesucces belangrijk om uit te kijken naar een goede opleidingsplaats?

Een multi-niveau analyse van de studieloopbaan van eerstejaarsstudenten van cohort 1995/96

Ariana Need

(Correspondentieadres: Afdeling Sociologie, Oude Hoogstraat 24, 1012 CE Amsterdam, e-mail: need@pscw.uva.nl)

Uulkje de Jong

(Correspondentieadres: SCO Kohnstamm Instituut, Wibautstraat 4, 1091 GM Amsterdam, e-mail: uulkje@educ.uva.nl)

ABSTRACT

This article examines the effect of Dutch College student's local study environments on their achievements. We found that local study environments only differ slightly in terms of grade averages, number of courses successfully completed, and dropout rates. Most of these differences are due to individual factors, but we find small effects of differences in student (self) selection between local study environments, and of differences in academic integration between local study environments. The results indicate that when it comes to higher education, the benefit of selecting the right local study environment is small. Academic achievement is shaped mainly by individual factors. The implications of our results are discussed.

1. INLEIDING

Uit een in 1994 door Elsevier gepubliceerd onderzoek bleek dat de meningen over de kwaliteit van hun opleiding behoorlijk verschilden tussen studenten van verschillende universiteiten, en ook tussen studenten van verschillende studierichtingen (Van Leeuwen, 1994). Niet alleen de beoordeling van de studenten over hun opleiding varieert sterk per studierichting, ook de behaalde resultaten van de studenten verschillen aanzienlijk. Zo liet Jansen (1996) voor verschillende studierichtingen aan de Rijksuniversiteit Groningen zien dat er zowel verschillen zijn in propedeuserendementen tussen studierichtingen als binnen cohorten. Deze onderzoeken suggereren dat er significante verschillen zijn tussen verschillende studieomgevingen in de resultaten van studenten. Daar staat tegenover dat uit een onderzoek naar de verschillen in oordelen van visitatiecommissies blijkt dat de oordelen slechts voor 4% gebonden worden door het kenmerk 'universiteit' (Hulshof, Van den Broek & Slijkerman, 1998).

Het doel van dit artikel is inzicht te krijgen in de omvang en de oorzaken van deze verschillen tussen studieomgevingen. Veel onderzoek naar studieloopbanen richt zich op individuele factoren die het studiesucces beïnvloeden (onder andere De Jong, Roeleveld & Webink, 1997). De theorie van Tinto (1987), die aan dit onderzoek ten grondslag ligt, veronderstelt echter dat niet alleen mechanismen op individueel niveau een rol spelen bij de studievoortgang, maar dat ook de studieomgeving van invloed is op de studieresultaten.

In dit artikel richten we ons op de studieresultaten en studieuitval van studenten in het hoger onderwijs na één jaar studeren. Hierbij boeken we vooruitgang ten opzichte van eerdere studies door niet alleen te kijken naar individuele factoren die het studiesucces beïnvloeden, maar

ook naar de invloed van de studieomgeving op de studieresultaten van studenten. Daarnaast volgen we *dezelfde* studenten in het hoger onderwijs gedurende één studiejaar, waardoor we de mechanismen die op individueel niveau werkzaam zijn adequaat kunnen bestuderen. Eerder onderzoek is veelal uitgevoerd binnen één studierichting of binnen één instelling in het hoger onderwijs. In dit onderzoek vergelijken we de studieresultaten van studenten aan *verschillende* studierichtingen en aan *verschillende* instellingen. We analyseren verschillen in studiesucces tussen *opleidingsplaatsen*, waarmee we de unieke combinatie van sector en instelling bedoelen. Binnen een sector zijn verschillende studierichtingen op basis van het gebruikte wetenschapparadigma samengevoegd. We beantwoorden een beschrijvende vraag naar de omvang van de verschillen in studieresultaten en studieuitval tussen opleidingsplaatsen, en een verklarende vraag die betrekking heeft op de factoren die deze verschillen verklaren. De beschrijvende vraag luidt als volgt:

Welk deel van de variantie in studieresultaten en studieuitval van studenten in het hoger onderwijs na één jaar studeren is te wijten aan verschillen tussen individuen en welk deel aan verschillen tussen opleidingsplaatsen?

Wanneer we de totale omvang van de variantie op beide niveaus (individueel en opleidingsplaats) hebben vastgesteld, willen we achterhalen welke individuele kenmerken en welke kenmerken van opleidingsplaatsen er nu precies verantwoordelijk zijn voor de gevonden verschillen. De Jong, Roeleveld, Webbink en Verbeek (1997) ontwikkelden een model waarin de nadruk ligt op het verklaren van studieresultaten via mechanismen op individueel niveau. In dit artikel vullen we deze individuele mechanismen aan met invloeden uit de directe studieomgeving. Dat er verschillen tussen instellingen zijn, lieten De Jong, Roeleveld en Webbink (1996) al zien. Hun onderzoek wijst uit dat er, zelfs na controle voor verschillen in individuele achtergronden van studenten, nog steeds significante verschillen in studieinzet tussen instellingen zijn. Jansen (1996) vond, zij het binnen een beperkt aantal studierichtingen aan de Rijksuniversiteit Groningen, ook significante verschillen in studievoortgang tussen studierichtingen. In dit artikel bekijken we of er ook verschillen gevonden worden wanneer we de combinatie van instelling en sector bekijken.

We kijken in dit artikel naar twee soorten invloeden van de directe studieomgeving op het studiesucces. Ten eerste kunnen de achtergrondkenmerken van studenten verschillen tussen opleidingsplaatsen. Deze verschillen in achtergrondkenmerken kunnen de gevonden verschillen in studieresultaten en studieuitval tussen opleidingsplaatsen verklaren. Wanneer bijvoorbeeld studenten die zelf geen hoog eindexamencijfer hebben betere studieresultaten halen in opleidingsplaatsen waarin veel studenten een hoog eindexamen hebben dan in andere opleidingsplaatsen (met een lager gemiddeld eindexamencijfer), kunnen we spreken van een invloed van de opleidingsplaats. Naast deze geaggregeerde invloed van de achtergrondkenmerken kijken we of er een geaggregeerde invloed is van het integratief vermogen van opleidingsplaatsen. We komen hier in de theoretische achtergrond van dit artikel op terug.

We beantwoorden in dit artikel de volgende verklarende vraag:

Welke individuele kenmerken en kenmerken van opleidingsplaatsen zijn verantwoordelijk voor verschillen in studieresultaten en studieuitval van studenten?

De opzet van dit artikel is als volgt. In de volgende paragraaf gaan we in op de theoretische achtergrond. We maken daarin een onderscheid tussen het loopbaanproces op individueel niveau en de invloed van de directe studieomgeving. In de derde paragraaf beschrijven we de gegevens die we gebruiken om de onderzoeksvragen te beantwoorden. In de vierde paragraaf gaan we in op de gebruikte methode. De beschrijvende onderzoeksvraag beantwoorden we in paragraaf 5, de verklarende vraag in paragraaf 6. In de laatste paragraaf bespreken we de conclusies.

2. THEORETISCHE ACHTERGROND

De vraagstellingen in dit artikel hebben betrekking op het keuzeproces van studenten in het hoger onderwijs in Nederland. De keuzes in dit proces bestaan uit beslissingen van individuele studenten over het tempo waarin de studie wordt afgelegd en over verandering van opleiding of doorgaan met de huidige opleiding. In deze paragraaf geven we aan welke mechanismen we veronderstellen werkzaam te zijn in dit keuzeproces, waarbij we een onderscheid maken tussen de mechanismen op individueel niveau en mechanismen op het niveau van opleidingsplaatsen.

Het loopbaanproces op individueel niveau

Voor de mechanismen op individueel niveau gebruiken we het onderzoek van De Jong, Roeleveld, Webbink en Verbeek (1997). In dit onderzoek is een model ontwikkeld voor het onderzoek naar studieloopbanen in het hoger onderwijs, waarin inzichten uit sociologische en economische theorieën zijn gecombineerd. Het model sluit nauw aan bij de theorie van Tinto (1987) en de human capital theorie en gaat uit van de veronderstelling dat de studieloopbaan moet worden opgevat als een longitudinale keten van: voorgeschiedenis - doelen - inschatting realiseerbaarheid doelen - commitments - ervaringen - integratie - bijstelling doelen - bijstelling inschatting realiseerbaarheid doelen, enz. In dit artikel onderzoeken we één volledige cyclus.

De doelen waarmee studenten aan de studie beginnen zijn afhankelijk van de ervaringen die ze in hun eerdere leven hebben opgedaan. Deze ervaringen meten we niet direct, we kijken naar achtergrondkenmerken als het *geslacht, het opleidingsniveau en het inkomensniveau van de ouders*, en de *eerdere schoolloopbaan* om deze eerdere ervaringen aan te duiden. De eerdere schoolloopbaan indiceert behalve de onderwijservaringen ook de initiële capaciteiten van de student in de vorm van *gemiddeld eindexamencijfer en leeftijd*, dat niet alleen naar zittenblijven verwijst maar ook naar de leerweg en aantal jaren dat men het voortgezet onderwijs al heeft verlaten.

De doelen zijn onder te verdelen in algemene doelen en operationele doelen (Lindenberg, 1990). Onder algemene doelen verstaan we de motieven die studenten hebben om aan de opleiding te beginnen. Deze doelen kunnen positief zijn, bijvoorbeeld een sterke *extrinsieke motivatie*, of veel meer negatief van aard zijn, bijvoorbeeld als men de studie vooral gekozen heeft vanwege de *haalbaarheid*. Los van de algemene doelen stellen studenten zich operationele doelen, waaronder we de intenties van studenten verstaan die richtinggevend zijn voor hun handelen. Voorbeelden van operationele doelen zijn de mate waarin studenten van plan zijn het diploma van de gekozen opleiding te halen, of *de intentie het propedeusediploma in één jaar te halen*. Operationele doelen zijn voor het keuzeproces van groot belang, omdat, zoals Esser (1997) stelt, de meeste mensen die zich in sterke mate hebben vastgelegd op een operationeel doel zich tijdens een loopbaan niet steeds afvragen of ze hun keuze moeten herzien, maar 'automatisch' doorgaan op de ingeslagen weg.

Studenten maken op basis van hun capaciteiten, motivatie, operationele doelen en hun mogelijkheden om de studie te financieren een *inschatting van hun kansen om het diploma te halen* en van de *studieduur* die ze nodig hebben om het diploma te halen. Oosterbeek (1992) toonde aan dat subjectieve inschattingen van succeskansen en kosten (studieduur), gecombineerd met subjectieve toekomstverwachtingen een betere weergave zijn van het keuzeproces dan de klassieke human capital theorie waarin alleen rekening wordt gehouden met het effect van de inschatting van de financiële bonus die een diploma oplevert.

Vervolgens komt het erop aan dat de intenties van studenten worden omgezet in handelen, in de eerste plaats af te lezen aan de studieinspanning, ofwel het aantal uren dat men bezig is met de studie. Behalve dit 'fysieke' commitment in de vorm van uren die men besteedt, is ook een 'mentaal' commitment van belang voor de loopbaan in de vorm van getoonde inzet voor de studie. De Jong, Koopman en Roeleveld (1991) toonden aan dat in een longitudinaal

model inzet een groot effect heeft op studietempo. Ferrari, Johnson en McCown (1995) en Schouwenburg (1994) onderzochten hetzelfde verschijnsel van de andere kant, en noemden het procrastinatie, resp. uitstelgedrag. Zij kwamen in hun analyses tot overeenkomstige conclusies.

De eerste ervaringen met medestudenten, staf en onderwijsinhoud en -vormen kunnen zozeer afwijken van de verwachtingen van de student dat hij gaat twijfelen of de keuze voor de opleiding geen fout is geweest en of hij zich om zijn algemene doelen te realiseren niet moet bezinnen op de toekomst. Ook kunnen de ervaringen in het onderwijs aanleiding zijn om de inschatting van de slaagkansen te heroverwegen, gezien de discrepantie tussen de vereiste capaciteiten en de inschatting van de eigen capaciteiten. Andere relevante ervaringen zijn het oordeel over de studeerbaarheid en de ervaren werkdruk. Tinto plaatste deze ervaringen in het licht van het concept van Van Gennep 'rites of passage', want of men zich goed heeft voorbereid of niet, het blijft een overgang naar een goeddeels nieuwe omgeving. In deze overgangperiode moet de student proberen aan de normen van de opleidingsplaats te voldoen en te integreren in de nieuwe omgeving. Dit wordt aangeduid als het proces van academische integratie.

Invloed van de studieomgeving

Iedere opleidingsplaats formuleert eindtermen, die bestaan uit de benoeming van de vereiste kennis, vaardigheden en attitudes van de afgestudeerden (Elshout-Mohr, de Jong & van den Bijtel, 1994). Per opleidingsplaats wordt zo een unieke set van eisen geformuleerd waaraan de student moet voldoen. Tijdens het proces waarin de student een opleidingsplaats kiest, probeert hij zich een beeld te vormen van de formele eisen en de minder geëxpliciteerde normen waaraan hij zal moeten voldoen. Dit keuzep proces voltrekt zich in Nederland als een vorm van zelfselectie, met als resultaat dat er grote verschillen in instroom bestaan tussen opleidingsplaatsen. Deze verschillen bestaan uit verschillen in eerdere schoolloopbaan en initiële capaciteiten, man -vrouw verhouding, en sociale achtergrond van studenten (Dronkers, 1982; Van Hout, 1996) en komen overeen met die in landen waar een strenge ingangselectie wordt toegepast (Tinto, 1987; Pascarella & Terenzini, 1991; De Jong, Roeleveld & Webbink, 1996). Daarnaast verschillen ook persoonlijkheidskenmerken van studenten sterk tussen opleidingsplaatsen (Smart & Elton, 1975; De Fruyt & Mervielde, 1996). Al deze verschillen duiden we aan als de verschillende aantrekkingskracht van opleidingsplaatsen.

Opleidingsplaatsen verschillen evenwel niet alleen in aantrekkingskracht, maar ook in integratief vermogen. Hiermee bedoelen we dat het individuele proces van academische integratie bij bepaalde opleidingsplaatsen een grotere kans op mislukking heeft dan bij andere. Dit kan komen doordat de eisen binnen sommige opleidingsplaatsen te hoog zijn, of doordat studenten de werkdruk niet aankunnen. Het proces van academische integratie kan ook mislukken doordat de opleidingsplaats heel anders was dan verwacht, door onvrede met de medestudenten, e.d..

We verwachten dat studievertraging en uitval uit het hoger onderwijs niet alleen individuele processen zijn, maar dat ook de sociale omgeving van de opleidingsplaats van invloed is op de studieloopbaan. Zo kunnen we bijvoorbeeld verwachten dat het studietempo van individuele studenten binnen opleidingsplaatsen met veel modale studenten trager is dan binnen opleidingsplaatsen met weinig modale studenten. Het studietempo van individuele studenten binnen opleidingsplaatsen waar studenten gemiddeld een hoge werkdruk ervaren zal ook trager zijn dan binnen opleidingsplaatsen waar studenten gemiddeld een lagere werkdruk ervaren.

3. DATA EN OPERATIONALISATIE

Om de onderzoeksvragen te beantwoorden, gebruiken we gegevens die verzameld zijn voor het onderzoek *Determinanten van de Deelname*. Dit bestand betreft een steekproef van 4412 eerstejaars studenten, getrokken uit studenten, die zich in 1995 voor het eerst inschre-

ven bij een opleiding van een bepaalde opleidingsplaats¹. Al deze studenten volgden een propedeuseprogramma aan een opleidingsplaats van het WO of HBO. Er is een gestratificeerde steekproef getrokken op basis van WO/HBO, sector, instelling en opleiding. Door deze opzet is het mogelijk om verschillen tussen opleidingsplaatsen te analyseren. Met opleidingsplaats bedoelen we zoals gezegd de unieke combinatie van sector en instelling. Na één jaar, in 1996, zijn dezelfde studenten nogmaals ondervraagd. Van de 4412 ondervraagden die deelnamen aan de eerste enquête hebben er 2701 ook de tweede enquête ingevuld. Voor 2071 respondenten hebben we volledige informatie voor alle relevante variabelen. Daarom worden in de analyses voor het WO 1235 studenten in 59 opleidingsplaatsen opgenomen en voor het HBO 836 studenten in 66 opleidingsplaatsen. In Appendix A staat een overzicht van de aantallen studenten per opleidingsplaats voor het WO en HBO apart.

Studieresultaten en studieuitval

Studieresultaten zijn met twee variabelen gemeten: het percentage behaalde vakken en het gemiddelde cijfer. Aan de studenten is gevraagd welk deel van het programma dat ze in het cursusjaar 1995/96 volgden op 1 september 1996 voldoende was afgerond. Bij deze vraag konden de studenten een getal tussen de 0 en de 100 invullen. Wanneer de studenten deze vraag niet konden beantwoorden, maar wel antwoord gaven op de vraag hoeveel studiepunten zij in het vorige studiejaar hadden behaald, dan is het aantal behaalde studiepunten uitgedrukt als een percentage van het totaal te behalen studiepunten. Wanneer de propedeuse na 1 jaar behaald was, is dit percentage gelijk aan 100. Het gemiddelde cijfer dat studenten over de studie-onderdelen van het afgelopen cursusjaar behaalden, is gemeten door de studenten te vragen hun gemiddelde cijfer te noemen. Wanneer studenten geen precieze cijfers kregen, is gevraagd of ze hun resultaten wilden vertalen naar 'rapportcijfers' van 1 tot 10.

Studieuitval is gemeten door aan de studenten te vragen of zij aan het begin van het tweede studiejaar gestopt waren met hun studie. De variabele 'doorgaan' heeft de waarde 0 als studenten aan een andere studie zijn begonnen, 0 als men het hoger onderwijs heeft verlaten, en 1 als studenten in 1996 nog steeds dezelfde studie volgden als in september 1995.

Operationalisatie van de verklarende variabelen

In de analyse bekijken we het effect van een aantal verklarende variabelen. Deze zijn als volgt geoperationaliseerd:

Voorgeschiedenis

Allochtoon

Of studenten al dan niet van allochtone afkomst zijn, is vastgesteld door aan de studenten naar hun eigen geboorteland te vragen en naar dat van hun ouders. Wanneer de studenten zelf of een van hun ouders niet in Nederland geboren zijn, beschouwen we ze als allochtoon, in alle andere gevallen als autochtoon. Binnen de groep allochtonen onderscheiden we ook een groep 'overig buitenland': studenten afkomstig uit Indonesië of een OESO-land worden ingedeeld in de groep 'overig buitenland', dit geldt niet voor studenten afkomstig uit Polen, Turkije, Tsjechië en Hongarije. Deze indeling komt overeen met de WBEAA-definitie van het Ministerie van Binnenlandse Zaken.

Inkomen ouders

Het inkomen van de ouders verwijst naar de som van het netto inkomen per maand van beide ouders.

Opleiding ouders

De opleiding van de ouders is in de volgende vijf categorieën ingedeeld:

(1) Maximaal lager onderwijs; (2) Lbo, Mavo, 3 jaar voortgezet onderwijs; (3) Mbo, Havo, Vwo; (4) HBO; (5) WO.

Modale Student

Of een student een modale student is, hangt af van de leerweg die hij voorafgaand aan het

hoger onderwijs heeft gevolgd. Studenten met maximaal een Havo-diploma zijn modale studenten in het HBO en studenten met maximaal een Vwo-diploma zijn modale studenten in het WO. De variabele modale student heeft de waarde 1 voor het HBO als een student een Havo-diploma heeft (en niet hoger) en voor het WO als een student een Vwo-diploma heeft. In alle andere gevallen neemt de variabele modale student de waarde 0 aan.

Gemiddeld eindexamencijfer

Aan de respondenten is gevraagd aan te geven welke cijfers zij behaald hebben bij hun Vwo-, Havo- of Mavo-examen. Studenten die meerdere diploma's bezitten, is gevraagd de cijfers van het hoogste diploma in te vullen. Het gemiddelde eindexamencijfer (minimaal 1, maximaal 10) is alleen berekend voor respondenten die voor minimaal vier vakken hun cijfers invulden.

Doelen

Extrinsieke motivatie

De extrinsieke motivatie van studenten is in 1995 gemeten. Extrinsieke motivatie is een schaal die van 1 tot 10 loopt en is geconstrueerd op basis van de volgende items:

- Door deze studie te volgen denk ik later zelfstandig werk te kunnen verrichten.
- Door deze studie te volgen denk ik later een leidinggevende functie te kunnen bekleden.
- Door deze studie te volgen denk ik later zeker een betaalde baan te kunnen krijgen.
- Door deze studie te volgen kan ik later een bepaald beroep gaan uitoefenen.

Bij elk van deze uitspraken konden de respondenten met een getal tussen 0 (deze overweging speelt totaal geen rol) en 10 (deze overweging speelt een uitzonderlijk sterke rol) aangeven hoe belangrijk die was bij de keuze van hun studie.

Intrinsieke motivatie

De intrinsieke motivatie van studenten is gemeten aan de hand van een schaal gebaseerd op twee vragen. Bij de eerste uitspraak kon de student aangeven in hoeverre de overweging belangrijk was bij de studiekeuze. De uitspraak is:

- Ik heb deze studie gekozen omdat ik het onderwerp interessant vind.

Op de tweede vraag konden de studenten met een getal tussen 0 en 100 aangeven in hoeverre de inhoud van de studie een overweging was bij de keuze ervan. Dit getal is door 10 gedeeld voordat we de intrinsieke motivatie schaal geconstrueerd hebben.

Plan propedeuse halen

De mate waarin studenten in 1995 het voornemen hadden om de propedeuse in één jaar te halen, is gemeten door de studenten te vragen met een getal tussen de 0 (dat ben ik beslist niet van plan) en de 10 (dat ben ik zeer zeker van plan) aan te geven hoe vast ze van plan zijn om de propedeuse ook daadwerkelijk binnen één jaar te halen.

Inschattingen

Verwachte studieduur

Hoe lang studenten verwachten over hun huidige studie te gaan doen, is gemeten door de studenten te vragen aan te geven hoe lang ze in totaal denken nodig te hebben voor het behalen van hun eindexamen.

Kans diploma

Hoe groot de kans is die studenten zichzelf geven om het eindexamen te behalen, is rechtsreeks aan de studenten gevraagd.

Commitments

Inzet

De inzet van de studenten is in 1995 gemeten. Inzet is een schaal die van 1 tot 10 loopt en is geconstrueerd op basis van de volgende items:

- Mijn bezigheden buiten mijn studie verhinderen mij volledig op te gaan in mijn studie *
- Mijn zelfdiscipline is goed

- Ik gebruik alle beschikbare tijd om zo snel mogelijk af te studeren
- Ik probeer steeds een zo hoog mogelijk cijfer te halen
- Ik ben tevreden over de studieprestaties die ik tot nu toe heb geleverd
- Ik vind het moeilijk om me in te spannen voor oninteressante onderdelen*
- Ik heb de neiging verplichtingen uit te stellen*
- Ik vind het moeilijk zelfstandig mijn studie te plannen*
- Waarom zou ik mijn studie sneller afronden dan nodig is, het is de mooiste tijd van mijn leven*

* deze items zijn omgecodeerd naar positief

Bij elk van deze uitspraken konden de studenten met een getal tussen 0 (geldt totaal niet voor mij) en 10 (geldt precies voor mij) aangeven in hoeverre de uitspraak op hen van toepassing is.

Inspanning

De inspanning van studenten is gemeten door in 1995 aan de respondenten te vragen hoeveel uren per week zij gemiddeld besteden aan hun opleiding. Voor studenten die aangaven meer dan 60 uur per week aan hun studie te besteden, is de inspanning vastgesteld op 60 uur per week. De inspanning van studenten varieert zodoende tussen de 0 en de 60.

Ervaringen

Problematische aansluiting studiekeuze bij capaciteiten

Deze variabele, gemeten in 1996, geeft aan of studenten in het afgelopen studiejaar problemen hebben ervaren in hun studie wat betreft de aansluiting van de studie bij hun capaciteiten en vaardigheden. Op grond van de volgende uitspraken is een schaal geconstrueerd:

- De studie was voor mij te moeilijk
- Het ontbrak mij aan studievaardigheden
- Ik miste essentiële voorkennis
- Leervakken geven mij problemen
- De studie was voor mij te gemakkelijk*

* dit item is omgecodeerd naar positief

Ook bij deze uitspraken konden de studenten met een getal tussen 0 en 10 aangeven in hoeverre deze uitspraken bij hun ervaringen of situatie in het afgelopen studiejaar passen.

Tevredenheid met het onderwijs

Aan de studenten is in 1996 gevraagd in hoeverre zij tevreden zijn over het onderwijs. Verschillende aspecten van het onderwijs kwamen hierbij aan de orde. Er is gevraagd naar de tevredenheid met de studiekeuze, de tevredenheid met de studiebegeleiding en de tevredenheid over de studeerbaarheid. Hieronder geven we voor elk van deze schalen weer op grond van welke uitspraken zij zijn geconstrueerd.

Tevredenheid met de studiekeuze

Ik denk er steeds vaker over om naar een andere studie(richting) over te stappen*.

- De studie is niet wat ik er van had verwacht; ze valt me tegen*.
- Als ik aan mijn studie denk word ik wel eens moedeloos*.
- Ik ben tevreden met mijn studiekeuze.
- Ik vind mijn studie over het algemeen boeiend.
- Ik heb over het algemeen plezier in mijn studie.
- Ik vind het leuk om steeds nieuwe dingen te leren in mijn studie.
- Meestal vind ik het wel prettig om aan de studiedag te beginnen.

* deze items zijn omgecodeerd naar positief

Tevredenheid met de studiebegeleiding

Ik heb de indruk dat docenten niet of nauwelijks geïnteresseerd zijn in hun studenten*.

- Ik ben over het algemeen ontevreden over de bereikbaarheid van docenten (telefonisch, op het instituut)*.
- Je krijgt onvoldoende gedetailleerd commentaar op je werk*.

- De onderlinge sfeer op de faculteit/school is over het algemeen goed.
- De meeste docenten stellen het op prijs wanneer je hen buiten colleges of werkgroepen benadert met vragen.
- Er zijn voldoende mogelijkheden om invloed uit te oefenen op wat er tijdens het onderwijs gebeurt.
- Het is gemakkelijk contact te leggen met docenten.
- De docenten geven je snel te horen of je goed of niet goed werkt.
- Docenten en studieadviseur zorgen voor voldoende sociaal-emotionele ondersteuning.
- Mijn studiebegeleider(s) (indien afwezig: mijn docenten) kennen me persoonlijk.
- Mijn studiebegeleider(s) (indien afwezig: mijn docenten) hebben altijd de goede gegevens bij de hand als je iets vraagt.

* deze items zijn omgecodeerd naar positief

Tevredenheid over de studeerbaarheid

- De studiebezigheden van verschillende onderdelen overlappen elkaar te veel*.
- De organisatie en onderwijsopzet omvat vele improductieve bezigheden en verliestijd*.
- De studiematerialen zijn over het algemeen te laat beschikbaar*.
- In het begin van een programma-onderdeel valt er meestal weinig te studeren (gedwongen onderbelasting)*.
- Aanvankelijk is het meestal onduidelijk wat je moet doen ter voorbereiding van een tentamen*.
- Ik kon moeilijk hoogte krijgen van de tentameneisen*.
- Bij de tentamenvoorbereiding kan ik over het algemeen niet goed taxeren of ik de stof voldoende beheers (te weinig feedback)*.
- Het rooster is onevenwichtig: op sommige dagen van de week moet je heel veel onderwijs volgen en op andere juist heel weinig*.
- 1 of meer tentamens waren zwaarder dan ik in redelijkheid had kunnen verwachten*.
- 1 of meer tentamens waren anders dan de tentameneisen die mij bekend waren*.
- Belangrijke delen van de stof zijn pas kort voor het tentamen behandeld*.
- Het volgen van onderwijs is niet nodig om het tentamen te halen*.
- De boeken en studiematerialen bieden mij onvoldoende steun voor doelmatige studie*.

* deze items zijn omgecodeerd naar positief

De *werkdruk* van studenten is in 1996 gemeten aan de hand van de volgende uitspraken. Bij elk van deze uitspraken moesten de studenten met een getal tussen 0 en 10 aangeven in hoeverre de uitspraak past bij hun situatie of ervaringen in het afgelopen studiejaar:

- Er blijft naast de studie voldoende tijd over voor andere zaken*.
- Ik heb voldoende tijd om me voor te bereiden op colleges en werkgroepen*.
- Ik heb voldoende tijd om me voor te bereiden op tentamens*.
- Ik vind de werkdruk niet erg groot*.
- Ik moet doorgaans zo hard studeren dat ik nauwelijks nog tijd heb om eens wat op adem te komen.
- Het is voor mij praktisch ondoenlijk om het studieprogramma in de gestelde termijn met succes te doorlopen.
- Ik heb grote moeite de stof onder de knie te krijgen.
- Het is voor mij over het algemeen praktisch ondoenlijk om tijdig met de tentamenvoorbereiding te beginnen.
- Ik kan het studietempo niet bijbenen.

* deze items zijn omgecodeerd naar positief

Om de resultaten gemakkelijker te kunnen interpreteren hebben we (binnen het HBO en WO apart) van alle variabelen die op interval niveau zijn gemeten de gemiddelde score over alle studenten afgetrokken. Dit heeft tot gevolg dat in de analyses de constante steeds verwijst naar de groep studenten die op alle variabelen een gemiddelde score heeft.

De kenmerken van de opleidingsplaatsen zijn als volgt gemeten. Binnen elke opleidings-

plaats wordt de gemiddelde score berekend van alle variabelen die de *voorgeschiedenis* van studenten meten en van alle variabelen die de *ervaringen* van de studenten meten. Alle studenten binnen dezelfde opleidingsplaats krijgen dus dezelfde score op deze geaggregeerde variabelen, die zo het percentage vrouwen, de gemiddelde leeftijd van studenten en de gemiddelde tevredenheid met de studie binnen de opleidingsplaats aangeven.

4. METHODE

Gezien de aard van de vraagstelling en de beschikbare data ligt het voor de hand om te kiezen voor een multi-niveau analyse. Studenten studeren immers binnen een bepaalde opleidingsplaats. Alleen door rekening te houden met het feit dat sommige studenten binnen dezelfde opleidingsplaats studeren is het mogelijk om tot een zuivere schatting van omgevingseffecten te komen.

De afhankelijke variabele, studiesucces, meten we aan de hand van drie variabelen: het percentage behaalde vakken, het gemiddelde cijfer en het al dan niet doorgaan aan de opleidingsplaats waar men zich in 1995 had ingeschreven. Het is te verwachten dat deze drie indicatoren voor studiesucces een onderlinge samenhang vertonen: studenten met een hoger gemiddeld cijfer zullen over het algemeen ook een hoger percentage van hun vakken halen, en studenten die weinig vakken halen, zullen eerder naar een andere studie overstappen dan studenten die veel vakken halen. In de analyse schatten we daarom twee modellen². Het eerste model schat de effecten op het percentage behaalde vakken en het gemiddelde cijfer simultaan. Dit betekent dat we het effect schatten van bijvoorbeeld de inzet van studenten op het percentage behaalde vakken onder controle voor het effect van de inzet op het gemiddelde cijfer. De effecten in dit model worden in een multivariaat model met twee afhankelijke variabelen geschat (zie Goldstein, 1995). In het tweede model bekijken we de keuze om al dan niet door te gaan met de opleiding: in dit model nemen we het percentage behaalde vakken en het gemiddelde cijfer als verklarende variabelen in de analyse op. De effecten worden in een logistisch model geschat (Goldstein & Rasbash, 1996). Om de effecten van beide modellen te schatten, gebruiken we het programma MLwiN.

In beide modellen nemen we de kenmerken van de opleidingsplaats stapsgewijs op. Op grond van de theoretische overwegingen zijn twee groepen kenmerken van de opleidingsplaats geselecteerd waarvan we verwachten dat ze het studiesucces en de studieuitval mede bepalen. De eerste groep betreft de (geaggregeerde) instroomverschillen tussen opleidingsplaatsen³, de tweede de (geaggregeerde) verschillen in academische integratie tussen opleidingsplaatsen⁴. Het gaat hier in totaal om 14 variabelen die kenmerken van de opleidingsplaats weergeven. In het HBO hebben we voor studenten in 66 opleidingsplaatsen gegevens en in het WO voor studenten aan 59 opleidingsplaatsen. Het aantal eenheden van de analyse op het niveau van de opleidingsplaatsen is dus in het HBO 66 en in het WO 59. Bij zo weinig cases willen we deze verklarende 14 variabelen niet allemaal tegelijk in de analyse opnemen. Vanuit een model met alle individuele kenmerken voegden we daarom de kenmerken van de opleidingsplaats beurtelings toe. Alle kenmerken van de opleidingsplaats die in deze eerste stap een significant effect hadden zijn vervolgens samen in een model opgenomen. Wanneer een kenmerk van de opleidingsplaats in het HBO een significant effect had, is het ook in de analyse van studiesucces in het WO opgenomen en andersom. Hierdoor kunnen we steeds alle effecten in het WO goed met die op het HBO vergelijken.

5. VERSCHILLEN TUSSEN OPLEIDINGSPLAATSEN

Tabel 1 beschrijft de indicatoren van studiesucces: gemiddeld cijfer, percentage behaalde vakken en doorgaan met studeren. In MLwiN is een regressiemodel geschat met daarin een con-

stante en twee variantietermen: één op individueel niveau en één op het niveau van de opleidingsplaats⁵. Tabel 1 geeft een antwoord op de beschrijvende vraag in hoeverre er verschillen in studieresultaten en studieuitval na één jaar studeren zijn tussen opleidingsplaatsen. We presenteren de resultaten voor HBO en WO apart.

Tabel 1: Gemiddelde en variantiecomponenten van gemiddeld cijfer, percentgae behaalde vakken en al of niet doorgaan aan de opleidingsplaats

	Gemiddelde	Variantiecomponenten		Variantiecomponenten (%)	
		Studenten	Opleidings- plaatsen	Studenten	Opleidings- plaatsen
HBO					
Gemiddeld cijfer	6.83 (0.04)	0.45 (0.02)	0.07 (0.02)	86.5	13.5
Percentage behaalde vakken	87.2 (0.96)	438 (22.3)	23.4 (10.4)	94.9	5.1
Doorgaan	0.83 (0.01)	0.14 (0.01)	0.00 (0.00)	100	0
WO					
Gemiddeld cijfer	6.75 (0.03)	0.79 (0.03)	0.02 (0.01)	97.5	2.5
Percentage behaalde vakken	81.7 (0.92)	665 (27.3)	13.7 (8.37)	98.0	2.0
Doorgaan	0.80 (0.02)	0.16 (0.01)	0.01 (0.00)	94.1	5.9

HBO studenten halen een gemiddeld cijfer van 6.83, in het WO is het gemiddelde cijfer iets lager (6.75). Gemiddeld halen studenten op het HBO 87 procent van de vakken in het eerste jaar, in het WO is dit percentage 82. Tot slot zien we dat 83 procent van de studenten op het HBO doorging met de opleiding waarmee ze in 1995 begonnen, in het WO gaat 80 procent van de studenten door met de opleiding.

De tweede en derde kolom van tabel 1 geven de variantiecomponenten weer, in de laatste twee kolommen staan deze variantiecomponenten omgerekend in percentages. Wanneer een variantiecomponent significant is ($\alpha \leq 0.05$) staat deze vetgedrukt in de tabel. Voor het HBO zien we dat alleen het gemiddelde cijfer en het percentage behaalde vakken significant verschillen tussen opleidingsplaatsen. In het HBO zijn er geen verschillen tussen opleidingsplaatsen in het percentage studenten dat doorgaat met de opleiding. In het WO zijn er alleen significante verschillen in 'doorgaan' tussen opleidingsplaatsen.

Tabel 1 laat zien dat niet alleen individuele factoren het studiesucces bepalen: in het HBO kan 13.5% ($100 * (0.07 / (0.07 + 0.45))$) van de variantie in het gemiddelde cijfer en 5.1% van de variantie in het percentage behaalde vakken worden toegeschreven aan verschillen tussen opleidingsplaatsen. Er zijn geen significante verschillen tussen opleidingsplaatsen binnen het HBO in het al of niet doorgaan met de opleiding (met een effect van 0.0019 en een standaardfout van 0.002 is het gevonden effect kleiner dan de bijbehorende standaardfout). In het WO is 5.9% van de variantie in het doorgaan met de opleiding te wijten aan verschillen tussen opleidingsplaatsen.

De vraag is hoe deze verschillen tussen opleidingsplaatsen tot stand komen. We kunnen verschillen in studiesucces en studieuitval tussen opleidingsplaatsen vinden door instroomverschillen, door verschillen tussen sectoren, of door kenmerken die direct gerelateerd zijn aan de academische integratie binnen opleidingsplaatsen. In de volgende paragraaf bekijken we daarom in hoeverre we de verschillen tussen opleidingsplaatsen kunnen verklaren met deze verschillende kenmerken.

6. EFFECTEN VAN OPLEIDINGSPLAATSEN OP STUDIEPRESTATIES EN STUDIEUITVAL

Om te kijken welke factoren van invloed zijn op de studieprestaties en uitval van studenten na een jaar studeren voerden we een multi-niveau analyse uit. De resultaten hiervan staan in tabel 2 (voor HBO) en tabel 3 (voor WO). Merk hierbij op dat de effecten voor de afhankelijke variabelen 'percentage behaalde vakken' en 'gemiddeld cijfer' verkregen zijn met een multivariate analyse; de effecten voor de afhankelijke variabele 'doorgaan' zijn vanwege het binaire karakter van de variabele 'doorgaan' geschat in een logistische regressieanalyse.

Het multivariate model houdt rekening met het feit dat mensen die gemiddeld hogere cijfers halen over het algemeen ook een hoger percentage van hun vakken halen. De analyse van 'doorgaan' aan de opleidingsplaats is apart uitgevoerd. Hier nemen we aan dat het percentage behaalde vakken en het gemiddelde cijfer voor de student een overweging zijn bij de beslissing om al dan niet aan de opleidingsplaats door te studeren. Zowel 'gemiddeld cijfer' als 'percentage behaalde vakken' worden daarom in de analyse van 'doorgaan' als onafhankelijke variabelen in de analyse opgenomen.

Vrouwen halen zowel in het HBO als in het WO over het algemeen hogere cijfers dan mannen. Daarnaast halen vrouwen in het WO ook een hoger percentage van de vakken en ze stoppen op minder vaak met hun studie dan mannen. In het WO halen allochtonen een significant lager percentage vakken dan autochtonen. In het HBO beëindigen allochtonen vaker hun studie aan de opleidingsplaats dan autochtonen. Oudere studenten in het WO behalen een lager percentage van hun vakken, in het HBO vinden we dit effect niet. In het HBO halen studenten met rijkere ouders een lager percentage van de vakken dan studenten met armere ouders. Modale studenten halen alleen in het HBO significant lagere cijfers dan studenten die niet via de gangbare weg naar het hoger onderwijs zijn gegaan. Daarnaast vinden we voor zowel HBO als WO dat studenten die in hun vooropleiding een hoger gemiddeld eindexamencijfer hadden in hun eerste jaar op het hoger onderwijs niet alleen hogere cijfers halen, maar ook een hoger percentage van de vakken.

Van de doelen die studenten zich stellen vinden we in het HBO en het WO een significant effect van zowel intrinsieke motivatie als extrinsieke motivatie op het al dan niet doorgaan aan de opleidingsplaats. Of studenten nu aan een studie beginnen omdat ze denken er een baan mee te vinden of omdat ze het onderwerp interessant vinden, beide motivaties maken dat studenten met meer waarschijnlijkheid doorgaan met hun studie. In het WO blijkt daarnaast dat zowel een relatief hoge extrinsieke motivatie als een relatief hoge intrinsieke motivatie samen gaat met een lager percentage behaalde vakken⁶. Ook is er een significant effect van het voornemen om de propedeuse binnen één jaar te halen: studenten die aangaven stelliger van plan te zijn de propedeuse in één jaar te halen, behalen gemiddeld hogere cijfers en meer vakken dan studenten die aangaven dat minder stellig van plan te zijn. Dit voornemen om de propedeuse in één jaar te halen heeft overigens geen invloed op de beslissing om al dan niet door te studeren aan de opleidingsplaats.

De inschattingen die studenten maken blijken in het HBO andere effecten te hebben dan op het WO. In het HBO stoppen studenten die verwachten langer over hun studie te doen eerder met hun studie. In het WO heeft de verwachte studieduur geen significant effect op studievoortgang of studieuitval. De inschatting van de kans om het diploma daadwerkelijk te halen, heeft in het HBO het effect dat studenten die zichzelf een hogere kans geven het diploma te behalen ook een hoger gemiddeld cijfer halen. In het WO zorgt een hogere inschatting van deze kans voor een hoger percentage behaalde vakken en voor een grotere kans door te gaan met de studie.

Een relatief hoge inzet van studenten aan het begin van de studie betaalt zich in het hoger onderwijs terug in gemiddeld hogere cijfers en meer behaalde vakken. Echter, gecontroleerd voor andere factoren zoals het gemiddelde cijfer en het percentage behaalde vakken, haken studenten die in het WO aan het begin van hun studie een grotere inzet vertonen juist eerder af⁷. Het aantal uren dat studenten besteden aan hun studie heeft, gecontroleerd voor de inzet, geen effect op het percentage behaalde vakken of het gemiddelde cijfer.

Tabel 2: Determinanten van studiesucces: HBO (N=836; standaardfouten tussen haakjes)

	Percentage behaalde vakken	Gemiddeld cijfer	Doorgaan vs niet doorgaan
Constante	68.1 (11.3)	5.67 (0.41)	-7.95 (3.75)
VOORGESCHIEDENIS			
Vrouwen	2.06 (1.56)	0.13 (0.05)	-0.09 (0.31)
Allochtoon (autochtoon=ref)			
Allochtoon	-3.67 (2.79)	-0.08 (0.08)	-1.01 (0.50)
Overig buitenland	-2.48 (3.95)	-0.04 (0.12)	-1.10 (0.67)
Leeftijd	-0.24 (0.23)	0.01 (0.01)	0.03 (0.05)
Inkomen ouders	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Opleiding ouders	-0.76 (0.71)	0.00 (0.02)	-0.16 (0.14)
Modale student	-2.52 (1.39)	-0.14 (0.04)	0.04 (0.27)
Gemiddeld eindexamencijfer	2.58 (1.19)	0.28 (0.03)	-0.39 (0.25)
DOELEN			
Extrinsieke motivatie	0.23 (0.38)	-0.02 (0.01)	0.23 (0.07)
Intrinsieke motivatie	0.57 (0.51)	0.03 (0.01)	0.29 (0.10)
Propedeuse behalen	1.48 (0.40)	0.03 (0.01)	-0.00 (0.07)
INSCHATTINGEN			
Verwachte studieduur	-1.56 (1.05)	0.04 (0.03)	-0.39 (0.19)
Kans diploma	0.01 (0.05)	0.00 (0.00)	0.01 (0.01)
COMMITMENTS			
Inzet begin studie	1.50 (0.48)	0.08 (0.01)	-0.11 (0.10)
Inspanning begin studie	0.02 (0.06)	0.00 (0.00)	-0.01 (0.01)
ERVARINGEN			
Tevredenheid studiebegeleiding	0.22 (0.54)	0.00 (0.02)	-0.38 (0.11)
Problematische aansluiting	-2.58 (0.53)	-0.09 (0.02)	0.12 (0.11)
Studeerbaarheid	-1.49 (0.54)	-0.01 (0.02)	-0.12 (0.12)
Tevredenheid studiekeuze	2.01 (0.44)	0.01 (0.01)	0.67 (0.08)
Werkdruk	-1.43 (0.55)	-0.04 (0.02)	-0.10 (0.11)
Gemiddeld cijfer			-0.22 (0.24)
Percentage behaalde vakken			0.05 (0.01)
STUDIEOMGEVING			
Percentage modale studenten	-0.90 (4.83)	0.51 (0.17)	
Gemiddelde inkomens ouders	0.00 (0.00)	-0.00 (0.00)	
Gemiddelde tevredenheid studiekeuze	1.35 (1.37)	0.12 (0.05)	
Gemiddeld cijfer			1.23 (0.61)
Sector (HBO economie=ref)			
HBO sociaal	5.50 (2.74)	0.34 (0.11)	-1.12 (0.59)
HBO gezondheid	2.25 (2.64)	0.15 (0.10)	-1.77 (0.52)
HBO landbouw	7.58 (3.32)	0.06 (0.12)	0.54 (0.75)
HBO laboratorium	-0.59 (2.88)	-0.12 (0.11)	1.07 (0.62)
HBO pedagogiek	-0.38 (3.00)	0.27 (0.11)	-1.32 (0.60)
HBO taal/cultuur	-0.46 (2.55)	0.08 (0.10)	-1.13 (0.48)
HBO techniek	0.40 (2.42)	0.24 (0.10)	-0.04 (0.49)
		Variantiecomponenten	
Random gedeelte			
Opleidingsplaatsen			
Gemiddeld cijfer	0.01 (0.01)	0.01 (0.01)	
Percentage behaald	0 (0)	0 (0)	
Gemiddeld cijfer/percentage behaald	0 (0)	0 (0)	
Doorgaan			0(0)
Studenten			
Gemiddeld cijfer	0.27 (0.01)	0.27 (0.01)	
Percentage behaald	319 (15.6)	319 (15.6)	
Gemiddeld cijfer/percentage behaald	3.25 (0.34)	3.25 (0.34)	
Doorgaan			1 (0)

Ook de ervaringen die studenten hebben opgedaan in hun eerste studiejaar zijn van invloed op hun studieresultaten en studieuitval. Van alle ervaringen die studenten in hun eerste jaar opdoen heeft alleen de tevredenheid met de studiebegeleiding (en alleen op het HBO) een effect op de studieuitval: studenten die meer tevreden zijn met de studiebegeleiding stoppen, *ceteris paribus*, eerder met de studie dan studenten die minder tevreden zijn⁸. Wanneer de capaciteiten en vaardigheden van de student niet aansluiten bij de studie halen studenten lagere cijfers en een lager percentage van hun vakken. Er is ook een effect van de tevredenheid met de studiekeuze: studenten die tevreden zijn met hun studie halen hogere cijfers, meer vakken en gaan eerder door met de studie. Studenten die aangaven minder problemen te hebben met de studeerbaarheid halen over het algemeen minder vakken dan studenten die aangaven daar meer problemen mee te hebben⁹. Studeerbaarheid is volgens de gangbare definitie geoperationaliseerd als het ontbreken van belemmeringen op de voortgang van de studie. Het studietempo van studenten die zeggen geen of weinig belemmeringen te ondervinden ligt dus lager dan dat van studenten die veel hinder ondervinden. Dit is tegengesteld aan de verwachtingen die er in Nederland bestonden over de effecten van studeerbaarheid, maar komt overeen met de uitkomsten uit de causale analyse op het cohort 1991 (De Jong, Roeleveld, Webbink & Verbeek, 1997).

Tot slot halen studenten die een hogere werkdruk ervaren over het algemeen minder vakken en lagere cijfers dan studenten die een lagere werkdruk ervaren. In het WO blijkt dat een hogere werkdruk ook samengaat met een grotere waarschijnlijkheid door te studeren aan de opleidingsplaats¹⁰. Het gemiddelde cijfer behaald in het eerste studiejaar blijkt niet van invloed op de beslissing om door te gaan met de studie, het percentage behaalde vakken is dit wel: hoe hoger het percentage behaalde vakken is, hoe groter de waarschijnlijkheid dat de student doorgaat met de studie.

Tot slot richten we ons op de effecten van de studieomgeving. Op individueel niveau zagen we dat op het HBO modale studenten lagere cijfers halen dan studenten die niet via de standaard leerweg op het hoger onderwijs komen. Studenten die echter meer modale studenten in hun studieomgeving hebben, halen hogere gemiddelde cijfers. In het WO is dit effect tegengesteld: een studieomgeving met meer modale studenten zorgt er in het WO voor dat studenten juist lagere cijfers halen. In het HBO maakt een studieomgeving met meer studenten met 'rijke' ouders dat individuele studenten een hoger percentage van hun vakken halen. Op individueel niveau is een hoger inkomen van de ouders juist nadelig voor het percentage behaalde vakken van de student. In het HBO maakt een studieomgeving waarin studenten over het algemeen meer tevreden zijn met hun studie ook dat individuele studenten hogere cijfers halen. Bij het al dan niet doorgaan met de studie is er maar één significant effect van de studieomgeving: studenten in het HBO gaan eerder door als het gemiddelde cijfer van de studenten in een opleidingsplaats hoger is.

Om te kijken of de gevonden verschillen toe te schrijven zijn aan verschillen tussen sectoren, is vervolgens voor sector gecontroleerd. De effecten op het niveau van de opleidingsplaats verwijzen in dit geval alleen naar verschillen tussen instellingen. We vonden in het WO geen significante verschillen tussen sectoren in de studieprestaties en de studieuitval. In het HBO zijn die verschillen er wel: vergeleken met HBO economie halen studenten binnen de sector HBO sociaal en HBO landbouw een groter deel van de vakken. Binnen de sectoren HBO sociaal, HBO pedagogiek en HBO techniek wordt door studenten hogere cijfers gehaald dan in de sector HBO economie. En in HBO gezondheid en HBO pedagogiek stoppen studenten eerder met hun studie dan in HBO economie.

Uit tabel 2 en 3 blijkt dat er na controle voor de omgevingskenmerken en sector geen verschillen meer zijn in het percentage behaalde vakken tussen opleidingsplaatsen: in het random gedeelte is de variatie tussen opleidingsplaatsen in het percentage behaalde vakken gereduceerd tot nul. Dat wil zeggen dat de verschillen in het percentage behaalde vakken geheel verklaard zijn met verschillen in individuele kenmerken en verschillen in de studieomgeving. Het gemiddeld cijfer van studenten varieert in het HBO echter nog wel significant tussen opleidingsplaatsen, zelfs na controle voor individuele kenmerken en omgevingskenmerken. In het WO is dat niet het geval.

Tabel 3: Determinanten van studiesucces: wo (N=1235; standaardfouten tussen haakjes)

	Percentage behaalde vakken	Gemiddeld cijfer	Doorgaan vs niet doorgaan
Constante	97.9 (16.6)	7.74 (0.64)	-1.37 (3.65)
VOORGESCHIEDENIS			
Vrouwen	3.91 (1.30)	0.09 (0.04)	-0.48 (0.21)
Allochtoon (autochtoon=ref)			
Allochtoon	-5.96 (2.52)	-0.14 (0.08)	0.61 (0.41)
Overig buitenland	-4.01 (2.43)	-0.01 (0.08)	0.00 (0.37)
Leeftijd	-0.59 (0.27)	0.01 (0.01)	0.05 (0.06)
Inkomen ouders	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Opleiding ouders	-0.39 (0.60)	-0.13 (0.09)	0.01 (0.09)
Modale student	1.80 (2.81)	0.06 (0.09)	-0.09 (0.43)
Gemiddeld eindexamencijfer	8.10 (0.92)	0.51 (0.03)	-0.12 (0.16)
DOELEN			
Extrinsieke motivatie	-0.67 (0.30)	-0.01 (0.01)	0.14 (0.05)
Intrinsieke motivatie	-1.08 (0.44)	-0.00 (0.01)	0.14 (0.07)
Propedeuse behalen	2.28 (0.36)	0.03 (0.01)	-0.00 (0.05)
INSCHATTINGEN			
Verwachte studieduur	0.77 (0.75)	-0.01 (0.03)	0.01 (0.12)
Kans diploma	0.09 (0.04)	-0.00 (0.00)	0.02 (0.01)
COMMITMENTS			
Inzet begin studie	1.77 (0.44)	0.09 (0.01)	-0.17 (0.07)
Inspanning begin studie	0.10 (0.06)	0.00 (0.00)	0.00 (0.01)
ERVARINGEN			
Tevredenheid studiebegeleiding	0.33 (0.47)	0.00 (0.02)	0.07 (0.07)
Problematische aansluiting	-3.38 (0.51)	-0.14 (0.02)	-0.02 (0.08)
Studeerbaarheid	-1.07 (0.49)	-0.01 (0.02)	-0.05 (0.08)
Tevredenheid studiekeuze	2.73 (0.41)	0.05 (0.01)	0.39 (0.06)
Werkdruk	-1.61 (0.47)	-0.03 (0.02)	0.19 (0.07)
Gemiddeld cijfer			0.25 (0.14)
Percentage behaalde vakken			0.03 (0.00)
STUDIEOMGEVING			
Percentage modale studenten	-9.02 (10.9)	-0.93 (0.43)	
Gemiddelde inkomen ouders	-0.00 (0.00)	-0.00 (0.00)	
Gemiddelde tevredenheid studiekeuze	1.57 (1.53)	0.07 (0.06)	
Gemiddeld cijfer			-0.07 (0.55)
Sector (wo economie=ref)			
wo sociaal	3.33 (2.68)	-0.01 (0.11)	-0.65 (0.47)
wo gezondheid	5.00 (3.53)	-0.02 (0.12)	0.06 (0.67)
wo landbouw	-3.87 (4.62)	0.02 (0.19)	-0.22 (0.96)
wo natuur	0.62 (2.34)	0.01 (0.10)	-0.86 (0.48)
wo recht	-1.11 (2.67)	-0.18 (0.10)	-0.14 (0.51)
wo taal/cultuur	-3.04 (2.63)	-0.08 (0.10)	-0.52 (0.50)
wo techniek	-4.74 (2.73)	-0.21 (0.11)	-0.63 (0.54)
		Variantiecomponenten	
Random gedeelte			
Opleidingsplaatsen			
Gemiddeld cijfer	0.01 (0.01)	0.01 (0.01)	
Percentage behaald	0.42 (2.89)	0.42 (2.89)	
Gemiddeld cijfer/percentage behaald	0.11 (0.10)	0.11 (0.10)	
Doorgaan			0.31 (0.14)
Studenten			
Gemiddeld cijfer	0.43 (0.02)	0.43 (0.02)	
Percentage behaald	384 (15.7)	384 (15.7)	
Gemiddeld cijfer/percentage behaald	5.10 (0.40)	5.10 (0.40)	
Doorgaan			1 (0)

7. SAMENVATTING EN CONCLUSIES

In bovenstaande analyse hebben we onderzocht of de studieomgeving een effect heeft op de studieresultaten van studenten. Het doel van de analyse was om uit te maken of een student er verstandig aan doet uit te kijken naar de opleidingsplaats die de beste kans biedt op goede studieresultaten.

Alleen in het HBO lijkt er voor het gemiddelde cijfer dat studenten halen over de studieonderdelen in het eerste jaar een redelijk groot verschil tussen opleidingsplaatsen te bestaan (13.5% van de variantie is toe te schrijven aan verschillen tussen opleidingsplaatsen, in het WO 2.5%). De relatief grote verschillen tussen opleidingsplaatsen in gemiddeld cijfer lijken erop te wijzen dat er in het HBO aanzienlijke verschillen bestaan in beoordelingsgewoonten. Van de variantie in het percentage behaalde studiepunten is in het HBO 5.1% toe te schrijven aan verschillen tussen opleidingsplaatsen en in het WO 2.0%. In de keuze doorgaan of stoppen konden we op het HBO geen effect van de studieomgeving vaststellen. In het WO kan 5.9% van de variantie in het doorgaan met de opleiding worden toegeschreven aan verschillen tussen opleidingsplaatsen. Wij moeten hieruit concluderen dat de verschillen zonder controle voor kenmerken van de studenten of kenmerken van de opleiding niet erg groot zijn. De door ons gevonden verschillen zijn min of meer vergelijkbaar met het percentage van de variantie in de oordelen van de visitatiecommissies dat aan het kenmerk universiteit kan worden toegeschreven (Hulshof, e.a. 1998).

We hebben vervolgens geprobeerd de individuele variantie en de variantie tussen opleidingsplaatsen te verklaren met individuele kenmerken en kenmerken van de studieomgeving.

De uitkomsten van de directe effecten van het model op het niveau van de student tonen aan dat

de ervaringen van de studenten met de studie niet altijd de richting hebben die men zou verwachten. Dit wordt veroorzaakt doordat we directe effecten presenteren, na controle voor de effecten van de andere variabelen in het model. Bivariate analyses gaven wel effecten in de verwachte richting te zien.

In het in de inleiding genoemde onderzoek van Elsevier worden verschillen tussen opleidingsplaatsen die zijn gebaseerd op meningen van studenten over de kwaliteit van hun studie gepresenteerd als verschillen in *kwaliteit* tussen opleidingsplaatsen. We weten echter dat studenten die meer tevreden zijn met hun studieomgeving vaak behoren tot de 'betere' studenten. Onze analyse toont aan dat de meningen van studenten over deze kwaliteitsaspecten niet zonder meer geïnterpreteerd kunnen worden als verschillen in kwaliteit tussen opleidingsplaatsen: er dient eerst gecontroleerd te worden voor de achtergronden van deze studenten.

In paragraaf 2 werden de omgevingskenmerken onderverdeeld in kenmerken die de aantrekkingskracht van een opleidingsplaats op een bepaald type student aanduiden en kenmerken die duiden op het integratief vermogen van opleidingsplaatsen. De theorie van Tinto, die in dit artikel getoetst is, veronderstelt dat niet alleen individuele processen een rol spelen bij de studievoortgang, maar dat ook de studieomgeving van invloed is op de studieresultaten. Van de studieomgeving zou vooral het integratief vermogen van instellingen van belang zijn. Uit de analyses blijkt dat er niet veel effecten te vinden zijn van beide soorten omgevingsvariabelen. De omvang van de groep modale studenten en het gemiddeld ouderlijk inkomen zijn de enige significante effecten van de variabelen die de aantrekkingskracht van opleidingsplaatsen indiceren. De gemiddelde tevredenheid met de studiekeuze is het enige significante effect van variabelen die het integratief vermogen van instellingen indiceren. De effecten van deze omgevingsvariabelen zijn klein, maar wel afdoende om de verschillen tussen de opleidingsplaatsen te verklaren. In het WO zijn er na invoering van deze omgevingskenmerken zelfs geen effecten meer van sector, in het HBO nog enkele. In dit artikel vinden we geen steun voor de hypothese dat naast individuele kenmerken de studieomgeving ook significant van invloed is op het studiesucces. De verschillen die er zijn tussen opleidingsplaatsen zijn klein, en voor een deel te wijten aan instroomverschillen tussen opleidingsplaatsen.

Het antwoord op de vraag uit de titel of het belangrijk is om uit te kijken naar een goede opleidingsplaats luidt uiteindelijk ontkennend. Voor de kans op studiesucces in Nederland maakt het nauwelijks iets uit waar je gaat studeren. De inspanningen van Ritzén, de vorige minister van Onderwijs, om studenten te laten "stemmen met hun voeten" waren dan ook niet gebaseerd op Nederlands onderzoek naar het effect van de opleidingsplaats op studiesucces. Stemmen met de voeten hield in dat studenten voor de beste opleidingsplaatsen zouden kiezen. Dit streven werd gematerialiseerd door jarenlange subsidies voor het maken van de *Keuzegids voor Hoger Onderwijs*. Men maakte hierbij de veronderstelling dat er een causale relatie bestaat tussen de *tevredenheid* van de studenten en de *kwaliteit* van de opleiding en dat kwaliteit in deze zin vervolgens leidt tot meer kans op succes. In dit artikel lieten we zien dat de tevredenheid van studenten over onderwijs en studeerbaarheid van de opleiding - na controle voor relevante factoren als de capaciteiten en de studieinzet van studenten - geen positieve bijdrage levert aan meer studiesucces. Initiële capaciteiten van studenten doen de - geringe - verschillen in studiesucces tussen opleidingsplaatsen verdwijnen. Zolang in Nederland studenten een vrije studiekeuze hebben en instellingen niet door ingangselectie onderlinge verschillen kunnen creëren zal er slechts een klein deel van de verschillen in studiesucces zijn te wijten aan de opleidingsplaats. En zelfs al komen er mogelijkheden tot ingangselectie, ook dan zullen verschillen tussen opleidingsplaatsen verdwijnen wanneer we controleren voor individuele verschillen. Kwaliteitsverschillen kunnen daarom niet worden aangetoond door verschillen in de output van opleidingsplaatsen. Slechts als we kunnen laten zien dat er verschillen bestaan in toegevoegde waarde, dus outputverschillen gecontroleerd voor inputverschillen, kunnen we spreken van kwaliteitsverschillen tussen opleidingsplaatsen.

Er zal veel moeten veranderen in het hoger onderwijs voordat we kwaliteitsverschillen die bestaan uit verschillen in toegevoegde waarde van opleidingsplaatsen kunnen constateren. Wij menen in dit artikel te hebben aangetoond dat de huidige discussie in het hoger onderwijs over kwaliteitsverschillen tussen instellingen misschien interessant is als men geïnteresseerd is in tevredenheid van studenten, maar niets te maken heeft met succeskansen van studenten. Onze conclusie is dat studenten er goed aan doen te zoeken naar een opleidingsplaats waarvoor ze over de benodigde capaciteiten beschikken en die voldoet aan hun persoonlijke programma van eisen, want voor studiesucces zullen ze zelf moeten zorgen.

LITERATUUR

Dronkers, J. (1982). Student-culture, social relations between students and grades. In E.A. Trotsenberg (ed.). *Higher education, a field of study* (p. 141-156). Frankfurt am Main: Peter Lang.

Elshout-Mohr, M., Jong, U. de & Bijtel, J. van den (1994). *Opdracht gestuurd leren. Vakwerk, jaargang 3, editie 0, 18-25*. Amsterdam: Universiteit van Amsterdam.

Esser, H. (1997). *The framing of marital relations and the risk of diverse*. Artikel gepresenteerd op de conferentie "Rational action theories in social analyses: applications and new developments". Stockholm.

Ferrari, J., Johnson, J. & McCown, W. (1995). *Procrastination and task avoidance: Theory, research and treatment*. New York: Plenum Pub. Corp..

Fruyt, F. de, & Mervielde, I. (1996). Personality and interests as predictors of educational streaming and achievement. *European Journal of Personality*, 405-425.

Goldstein, H. (1995). *Multilevel Statistical Models*. London: Edward Arnold:

Goldstein, H. and Rasbash, J. (1996). Improved approximations for multilevel models with binary responses. *Journal of the Royal Statistical Society, A*, 159, 385-443

Hout, J.F.M.J., van (1996). *Verschillen tussen wetenschapsgebieden*. (oratie) Den Haag: Vuga.

M.J.F. Hulshof, J.F.L.H. van den Broek en R.H. Slijkerman (1998). *Deskundigen over*

- kwaliteit. Analyse van Wo-visitatierapporten 1994-1997.* Nijmegen: IOWO.
- Jansen, (1996). *Curriculumorganisatie en studievoortgang. Een onderzoek onder zes studierichtingen aan de Rijksuniversiteit Groningen.* (proefschrift) Groningen: Gion.
- Jong, U. de, Koopman, & Roeleveld, J. (1991). *Snelwegen en slingerpaden in en om het Hoger Onderwijs.* Eindrapport project 'Studieloopbanen in het hoger onderwijs'. Den Haag Ministerie van O&W.
- Jong, U. de, Roeleveld, J. & Webbink, D. (1996). *De subjectieve factor: commitments en academische integratie in de studieloopbaan.* Serie Verder Studeren, deel 4; een panelstudie onder scholieren en studenten. Den Haag: Ministerie van OC&W.
- Jong, U. de, Roeleveld, J. & Webbink, D. (1997). *Studeren in de jaren negentig, Studiekeuze en studieloopbaan over de periode 1991-1995.* Verder Studeren Eindrapport; een panelstudie onder scholieren en studenten. Den Haag, Ministerie van OC&W.
- De Jong, U., J. Roeleveld, J., Webbink, D. & Verbeek, F. (1997). *Het Amsterdamse studieloopbaanmodel.* Verder Studeren, een panelstudie onder scholieren en studenten, deel 6. Den Haag, Ministerie van OC&W.
- Koppen, J.K. (1991). *Een kwestie van discipline: over de democratisering van het wetenschappelijk onderwijs.* (proefschrift) Amsterdam :Thesis Publishers.
- Leeuwen, A. van (1994). Universiteiten Test. Studenten geven rapportcijfers aan achttien opleidingen. *Elsevier, 50(43), 28-48.*
- Lindenberg, S.M. (1990). Homo socio-economicus: The emergence of a general model of man in social sciences. *Journal of Institutional Economics, 146, 727-748.*
- Oosterbeek, H. (1992). *Essays on human capital theory.* (proefschrift) Amsterdam: Thesis Publishers.
- Pascarella, E.T., & Terenzini, P.T. (1991). *How college affects students: Findings and insight from twenty years of research.* San Francisco: Jossey-Bass.
- Schouwenburg, H.C. (1994). *Uitstelgedrag bij studenten.* (proefschrift) Groningen: Rug.
- Smart, J.C. & Elton, C.F. (1975). Goal orientation of academic departments: a test of Biglan's model. *Journal of Applied Psychology, 60, 580-588.*
- Tinto, V. (1987). *Leaving college: Rethinking the causes and curses of student attrition.* Chicago: the University of Chicago.

NOTEN

1. Voor 83 procent van de studenten is dit ook de eerste opleiding in het hoger onderwijs.
2. In alle modellen is eerst gekeken of er sprake was van multicollineariteit. De onderlinge samenhang van variabelen beïnvloedt de effecten van variabelen niet.
3. Deze variabelen zijn: het percentage vrouwen in de opleidingsplaats, het percentage alloctonen in de opleidingsplaats, de gemiddelde leeftijd van studenten in de opleidingsplaats, het gemiddeld inkomen van de ouders van de studenten in de opleidingsplaats, de gemiddelde opleiding van de ouders van de studenten in de opleidingsplaats, het percentage modale studenten in de opleidingsplaats en het gemiddeld eindexamencijfer van studenten in de opleidingsplaats.
4. Deze variabelen zijn: de gemiddelde tevredenheid met de studiebegeleiding in de opleidingsplaats, de gemiddelde problematische aansluiting in de opleidingsplaats, de gemiddelde score op studeerbaarheid in de opleidingsplaats, de gemiddelde tevredenheid met de studie in de opleidingsplaats, de gemiddelde ervaren werkdruk in de opleidingsplaats, en voor de analyse van de afhankelijke variabele 'doorgaan' ook nog het gemiddelde van de behaalde cijfers in de opleidingsplaats en het gemiddeld percentage behaalde vakken in de opleidingsplaats.
5. De variabele 'doorgaan' is een binaire variabele. Om toch de variantiecomponenten op individueel en instelingsniveau te kunnen uitrekenen, negeren we hier het feit dat we in dit geval een logistische regressieanalyse

zouden moeten uitvoeren. In de latere analyses nemen we de variabele 'doorgaan' wel als binaire variabele in de analyse op.

6. Hier dient men zich te bedenken dat dit het directe effect is en dat dit zal afwijken van het totale effect, waarin ook de joint effects via de overige variabelen uit het model zijn opgenomen. Dat dit directe effect af zal wijken van het totale effect mogen we afleiden uit het feit dat de correlaties tussen extrinsieke en intrinsieke motivatie en percentage behaalde vakken resp. .02 en .14 zijn.

7. Ook hier doet zich het verschil tussen direct en totaal voor. De correlatie tussen inzet en doorgaan in het Wo is .19.

8. Idem, de correlatie tussen tevredenheid begeleiding en doorgaan is .16 in het Hbo.

9. Idem de correlatie studeerbaarheid en percentage behaalde vakken is .11 in het Hbo en .17 in het Wo.

10. Idem, de correlatie studeerbaarheid en doorgaan is -.11 in het Wo.

APPENDIX A:
OVERZICHT GEMIDDELD PERCENTAGE BEHAALDE VAKKEN
(AANTALLEN TUSSEN HAAKJES) PER SECTOR EN PER INSTELLING

WO									
	Econo- mie	Sociaal	Gezond- heid	Land- bouw	Natuur/ wiskunde	Rechten	Taal	Techniek	Totaal
EU Rotterdam	72 (14)	82 (12)	94 (11)			77 (13)	90 (25)		84 (75)
KU Brabant	78 (12)	90 (26)				89 (13)	90 (18)		88 (69)
KU Nijmegen		90 (43)	100 (5)		79 (41)	69 (18)	86 (38)		83 (145)
LU Wageningen				80 (24)	85 (14)				82 (38)
RU Groningen	82 (33)	86 (25)	88 (6)		79 (52)	85 (10)	77 (23)	74 (10)	81 (159)
RU Leiden		77 (13)	100 (3)		87 (43)	75 (15)	91 (26)		85 (100)
RU Limburg	79 (24)	90 (10)	93 (12)			79 (10)	100 (11)		86 (67)
RU Utrecht	78 (5)	80 (30)	94 (7)		83 (69)	73 (13)	76 (33)		81 (157)
TU Delft		86 (6)						75 (42)	76 (48)
TU Eindhoven	69 (15)							75 (47)	73 (62)
TU Twente	81 (20)	78 (13)						73 (42)	76 (75)
UvA Amsterdam	73 (6)	83 (31)	93 (8)		85 (50)	88 (12)	66 (22)		81 (129)
VU Amsterdam	80 (17)		95 (11)		82 (46)	84 (11)	83 (17)		84 (102)
Antwerpen			80 (4)						80 (4)
Leuven			74 (5)						74 (5)
	78 (146)	85 (209)	92 (72)	80 (24)	83 (315)	79 (115)	83 (213)	74 (141)	82 (1235)
HBO									
	Econo- mie	Sociaal	Gezond- heid	Land- bouw	Natuur/ wiskunde	Rechten	Taal	Techniek	Totaal
HSEN Enschede	88 (31)	96 (15)	89 (10)		90 (18)			69 (21)	86 (95)
VHI Leeuwarden				87 (7)					87 (7)
Pabo Amsterdam						99 (4)			99 (4)
ChrHSDN Ede		94 (19)							94 (19)
Hanze Groningen	78 (16)	84 (8)	95 (11)		88 (7)	90 (10)	81 (7)	70 (19)	82 (78)
HvanU Utrecht				100 (2)	75 (10)	83 (4)	77 (9)	92 (10)	83 (35)
HS-ES Rotterdam	84 (22)								84 (22)
ChrHSW Zwolle			82 (13)			95 (6)	90 (21)	84 (24)	87 (64)
HSDH Driebergen		99 (20)							99 (20)
HSH Diemen	91 (19)		93 (10)				73 (13)		86 (42)
HSWB Breda				91 (11)	82 (17)			86 (19)	86 (47)
HSEI Eindhoven			89 (13)		90 (10)		85 (12)	87 (15)	88 (50)
AHSDB Den Bosch				95 (12)					95 (12)
HES Amsterdam	89 (24)								89 (24)
HSD Emmen					70 (13)			78 (8)	73 (21)
HvA Amsterdam		95 (14)	100 (14)		83 (10)	93 (8)	85 (6)	86 (15)	91 (67)
HIS Rotterdam							94 (6)		94 (6)
HSG Nijmegen			92 (17)			72 (10)	81 (8)		86 (35)
HSDB Den Bosch	85 (20)								85 (20)
HSKatS Sittard			98 (12)			90 (6)	90 (5)		91 (23)
AHSD Delft				87 (9)					87 (9)
NHSL Leeuwarden			86 (5)				90 (6)		88 (11)
ChrAHS Dronen				84 (4)					84 (4)
HHS Den Haag		97 (18)	93 (11)						96 (29)
HSKatT Tilburg						95 (10)	90 (31)		92 (41)
ChrHNN Leeuwarden						87 (19)			87 (19)
HSR&O Rotterdam					92 (5)	90 (11)	79 (8)	86 (8)	87 (32)
Totaal	86 (132)	95 (94)	93 (116)	90 (45)	83 (90)	88 (88)	85 (132)	81 (139)	87 (836)

Waargenomen Eigen Competentie, Doeloriëntaties en Studieresultaten bij Beginners en Gevorderden

Anneke Vrugt, Frans J. Oort & Caroline Zeeberg
(Afdeling Psychologie, Faculteit Maatschappij- en Gedragwetenschappen, Universiteit van Amsterdam, Roetersstraat 15, 1018 WB Amsterdam)

SAMENVATTING

Op grond van eerder onderzoek wordt verwacht dat de opvattingen over de eigen competentie en sociale vergelijking zullen bijdragen aan de doelen die nagestreefd worden en dat deze doelen zullen doorwerken in de prestaties die behaald worden. Ook wordt verwacht dat een taakoriëntatie zal bijdragen aan de opvattingen over de eigen competentie en dat een ego-oriëntatie zal bijdragen aan de sociale vergelijking. De centrale vraag in het onderhavige onderzoek is of de relaties tussen de bovengenoemde motivationele variabelen bij beginners op een bepaald gebied even sterk ontwikkeld zijn als bij personen die reeds enigszins gevorderd zijn. Omdat beginners beperkte kennis hebben van de taakvereisten verwachten we dat de relaties tussen deze variabelen bij beginners zwak of afwezig zullen zijn, en dat deze relaties bij meer gevorderden relatief sterk zullen zijn. De resultaten bieden overwegend steun voor de verwachtingen. Bij beginners en gevorderden is geconstateerd dat de opvattingen over de eigen competentie een bijdrage hebben geleverd aan de nagestreefde doelen en dat deze doelen op hun beurt bijdragen aan de prestaties. De relaties tussen deze variabelen zijn bij beginners duidelijk minder sterk dan bij gevorderden. Verder is vastgesteld dat alleen bij gevorderden de taakoriëntatie heeft bijgedragen aan de opvattingen over de eigen competentie, hetgeen suggereert dat er bij gevorderden relaties tussen motivationele variabelen werkzaam zijn die bij beginners nog niet tot ontwikkeling zijn gekomen.

ABSTRACT

Based on previous research we expected that perceived self-efficacy and social comparison would contribute to the goals that are pursued and that these goals would affect the study results. We also expected that a task orientation would contribute to perceived self-efficacy and that an ego orientation would contribute to social comparison. The central question of the present study is whether the relationship between the above-mentioned motivational variables are developed as strongly in the case of freshmen as in the case of more advanced students in a certain area. Because freshmen have limited knowledge of task requirements, we expected that the relations between these variables would be weak or absent in the case of freshmen, and that they would be relatively strong in the case of more advanced students. The results largely support these expectations. It was found that with respect to freshmen and advanced students perceived self-efficacy contributed to pursued goals and these goals contributed in their turn to achievements. The relationships between these variables are clearly less strong for freshmen than for advanced students. It was further found that task orientation only contributed to perceived self-efficacy in the case of advanced students, which suggests

that for these students motivational variables are active which are still undeveloped in the case of freshmen.

De prestatie-motivatietheorie van Atkinson (Atkinson, 1964; Atkinson & Raynor, 1974) heeft de aanzet gegeven tot het psychologische onderzoek naar de determinanten van prestatiegericht gedrag. Atkinson baseert zich op verwachtingstheoretische uitgangspunten, namelijk dat een persoon in een decisiesituatie een keuze zal maken uit alternatieven op grond van verwachtingen over de relatie tussen zijn gedrag en de uitkomsten, en de aantrekkelijkheid van de te verkrijgen uitkomsten. In het model dat Atkinson en collega's hebben ontwikkeld ligt de nadruk op twee tegenovergestelde motieven - het motief om succes na te streven en het motief om falen te vermijden - die beschouwd worden als duurzame persoonlijkheidskenmerken. Volgens dit model geeft emotionele anticipatie de aanzet tot prestatiegericht gedrag. Personen met een sterke behoefte om succesvol te zijn anticiperen op trots waardoor zij gemotiveerd worden om te presteren. Personen met een sterke behoefte om falen te vermijden anticiperen op schaamte en proberen dat te vermijden door zich terug te trekken of niet in te zetten.

Mede door de resultaten van studies waaruit bleek dat de taakuitvoering van personen met veel testangst niet zozeer werd aangetast door emotionaliteit, de subjectieve ervaring van fysiologische arousal, maar vooral door de cognitieve preoccupatie met de eigen tekortkomingen en de gevolgen van falen (zie bijvoorbeeld, Liebert & Morris, 1967), is er vervolgens meer aandacht besteed aan de rol van cognitieve processen (zie bijvoorbeeld Weiner, 1986; Bandura, 1986, 1997).

Ook Bandura gaat ervan uit dat mensen gemotiveerd worden door, en zich bij hun handelingen laten leiden door de uitkomsten die naar verwachting voortkomen uit een bepaald gedragsverloop. Volgens hem is het echter van doorslaggevend belang of men meent over de kennis, vaardigheden en strategieën te beschikken om het relevante gedrag tot stand te brengen. Ook als aangenomen kan worden dat bepaalde activiteiten bij goede uitvoering resulteren in waardevolle uitkomsten, bijvoorbeeld het schrijven van een belangwekkend essay dat geplaatst wordt in een vooraanstaand tijdschrift, zal dit iemand die meent niet over de relevante bekwaamheden te beschikken, niet motiveren. Bandura heeft dan ook een theorie ontwikkeld waarin een centrale plaats wordt ingeruimd voor het subjectieve oordeel over de eigen capaciteiten.

Bandura's theorie (1986, 1989, 1997) over waargenomen eigen competentie ("perceived self-efficacy") beschrijft op welke manier iemands denkbeelden over de eigen competentie zijn of haar prestaties beïnvloeden en door welke factoren deze denkbeelden beïnvloed kunnen worden. Het begrip "waargenomen competentie" wordt door Bandura (1986) als volgt omschreven: "... People's judgement of their capabilities to organize and execute courses of action required to attain designated types of performances" (p. 395). Uit deze omschrijving komt naar voren dat het subjectieve oordeel over de eigen capaciteiten domeinspecifiek is. Het gaat dus niet om een persoonlijkheidskenmerk, zoals zelfvertrouwen. Dit betekent ook dat iemand zichzelf op een bepaald domein, zoals het studeren voor tentamens, heel competent, en op een ander domein, zoals het schrijven van een paper, veel minder competent kan achten. Uit Bandura's omschrijving kan ook afgeleid worden dat het oordeel over de eigen competentie tot stand komt in relatie tot een bepaald prestatieniveau. Iemand kan zichzelf bijvoorbeeld heel goed in staat achten een taak op gemiddeld niveau uit te voeren, maar menen dat hij of zij minder goed in staat is een excellente prestatie op dezelfde taak te leveren.

Uit verschillende studies komt naar voren dat de denkbeelden over de eigen competentie op een bepaald terrein van invloed zijn op de prestaties op dat gebied (Bandura, 1997; Vrugt, 1996). Multon, Brown en Lent (1991) hebben meta-analyses uitgevoerd naar de relatie tussen de opvattingen over de eigen competentie ten aanzien van school- en studievakken en prestaties. De resultaten van deze analyses wijzen uit dat de denkbeelden over de eigen competentie een positieve bijdrage leveren aan de school- en studieprestaties, ongeacht de gebruikte onderzoeksopzet en de manier waarop de eigen competentie is gemeten.

De wijze waarop de eigen competentie de prestaties beïnvloedt laat zich als volgt beschrijven. Een persoon met gunstige denkbeelden over de eigen competentie op een bepaald terrein zal een niveau van taakuitvoering ambiëren dat net iets boven het eigen bereik ligt. Gunstige denkbeelden over de eigen competentie leiden tot het nastreven van uitdagende, persoonlijke doelen en tot veel inzet om zo'n persoonlijke doel te realiseren. (Locke & Latham, 1990; Vrugt, Langereis & Hoogstraten, 1993, 1994, 1997). Deze factoren stimuleren metacognitieve activiteiten en de ontwikkeling van metacognitieve vaardigheden, hetgeen bevorderlijk is voor goede prestaties (Elshout & Veenman, 1992; Locke & Latham, 1990). Als iemand er ongunstige denkbeelden over de eigen competentie op nahoudt, zal hij of zij een minder hoog prestatieniveau nastreven. Twijfel en onzekerheid tijdens de taakuitvoering tasten de concentratie aan (Deffenbacher, 1978), en bij een moeilijke taak wordt de moed snel opgegeven. Deze processen dragen ertoe bij dat personen met ongunstige denkbeelden over hun competentie minder goede prestaties leveren (Tuckman & Sexton, 1992).

Bandura (1986, 1997) heeft op verschillende plaatsen beschreven dat vergelijking van de eigen competentie en prestaties met die van relevante anderen de motivatie in belangrijke mate kan beïnvloeden. Hij heeft echter geen systematische aandacht besteed aan de rol die de richting van de sociale vergelijking kan spelen, bijvoorbeeld vergelijking met een persoon die slechter af is (neerwaartse vergelijking), of vergelijking met een persoon die beter af is (opwaartse vergelijking). De theorie over neerwaartse vergelijking (Wills, 1981, 1991) stelt dat neerwaartse vergelijking vooral wordt ingegeven door de behoefte de eigenwaarde te versterken. Als men zichzelf vergelijkt met een overeenkomstige ander die in relevante opzichten slechter functioneert, kan dat aanleiding geven tot gunstige conclusies over het eigen functioneren. Onderzoek wijst uit dat neerwaartse vergelijking bijdraagt aan een verhoogd welbevinden (Gibbons & Gerrard, 1991), en aan relatief goed functioneren (Affleck, Tennen, Pfeiffer & Fifield, 1987; Vrugt, 1994). Ook zijn er aanwijzingen dat neerwaartse vergelijking, een gemiddelde leerling die tot de besten behoort in een zwak presterende klas, meer aanzet tot goede prestaties, dan opwaartse vergelijking, een gemiddelde leerling in een goed-presterende klas (zie bijvoorbeeld Guldemand, Meijnen & Boomsma 1989).

Volgens verschillende auteurs (zie Collins, 1996) kan ook het streven naar zelfverbetering aanleiding geven tot opwaartse vergelijking. Als degene die vergelijkt meent dat de discrepantie tussen zijn of haar kenmerken en die van de vergelijkingpersoon gering is, kan opwaartse vergelijking een gunstig effect hebben op de stemming en zelfevaluatie. Het superieure niveau van de vergelijkingpersoon lijkt dan binnen het eigen bereik te liggen en dat kan stimuleren om zelf ook dat niveau na te streven (Vrugt & Koenis, 1996).

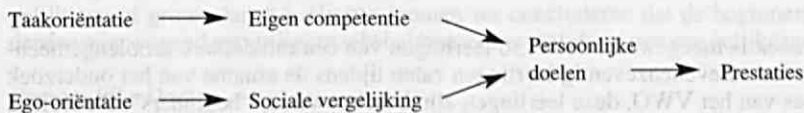
Volgens Bandura (1989, 1997) zijn ook de theorieën die men erop nahoudt over de determinanten van de eigen prestaties van invloed op de motivatie. In deze theorieën zijn twee gezichtspunten te onderscheiden die met verschillende termen worden aangeduid. Bandura (1989) heeft het over de opvattingen die mensen erop nahouden over intelligentie, als vaardigheid die ontwikkeld kan worden, of als stabiele entiteit die moeilijk te veranderen is. Elliot en Dweck (1988) onderscheiden twee doelen die individuen nastreven in prestatiegerichte situaties: leerdoelen en uitkomstdoelen. Als individuen er leerdoelen op nahouden, streven zij ernaar de eigen vaardigheden te vergroten en nieuwe taken te beheersen. Individuen met uitkomstdoelen proberen een positief oordeel over de eigen capaciteiten te handhaven en negatieve evaluaties te vermijden, omdat die negatieve evaluaties tot de conclusie zouden leiden dat hun capaciteiten onvoldoende zijn. Dit kan leiden tot het gebruik van strategieën om te voorkomen dat falen wordt toegeschreven aan gebrek aan capaciteiten (Covington, 1992). Als iemand bijvoorbeeld hard gewerkt heeft en desondanks faalt, wijst dat op onvoldoende capaciteiten. Door zich niet goed voor te bereiden kan, deze conclusie voorkomen worden. Duda en Nicholls (1992) gebruiken de term doeloriëntaties en maken daarbij onderscheid tussen taak- en ego-oriëntatie. Ook al verschillen de termen die gebruikt worden, de begripsomschrijvingen komen duidelijk overeen. Omdat ons onderzoek aansluit bij dat van Duda en Nicholls (1992) zullen wij eveneens gebruik maken van de termen taak- en ego-oriëntatie.

Uit verschillende studies (zie Hofmann & Strickland, 1995) komt naar voren dat taak- en ego-oriëntatie onafhankelijke factoren zijn. In aansluiting daarbij hebben Hofmann en Strickland (1995) vastgesteld dat taak- en ego-oriëntatie ook in interactie werkzaam kunnen zijn.

Uit de resultaten van experimenteel onderzoek waarin een taak- of ego-oriëntatie geïnduceerd is (Jourdan, Bandura & Banfield, 1991; Wood & Bandura, 1989) komt naar voren dat de activatie van een taakoriëntatie de opvattingen over eigen competentie versterkt. Bij Vrugt, Hoogstraten en Oort (1999) blijkt dat een taakoriëntatie ook in een niet-experimentele setting bijdraagt aan de eigen competentie. Ook in het onderzoek van Duda en Nicholls (1992) is vastgesteld dat een taakoriëntatie verband houdt met de beoordeelde eigen competentie. Personen met een ego-oriëntatie houden zich vooral bezig met het maken van een goede indruk, met het verslaan van anderen en het tonen van hun superioriteit. Deze gegevens zijn voor Staalvik (1997) aanleiding om te stellen dat personen met een ego-oriëntatie relatief sterk gericht zijn op de vergelijking met anderen. Hij heeft zich niet bezig gehouden met de vraag in hoeverre een ego-oriëntatie de richting van de vergelijking beïnvloedt.

Op grond van het bovenstaande kan verwacht worden dat de opvattingen over de eigen competentie en sociale vergelijking zullen bijdragen aan de persoonlijke doelen die nagestreefd worden en dat deze persoonlijke doelen zullen doorwerken in de prestaties die behaald worden. Ook kan verwacht worden dat een taakoriëntatie zal bijdragen aan de opvattingen over de eigen competentie en dat een ego-oriëntatie zal bijdragen aan de sociale vergelijking. De verwachte relaties tussen de onderzoeksvariabelen zijn grafisch weergegeven in het pad-diagram in Figuur 1.

Figuur 1: Paddiagram van de verwachte relaties tussen de onderzoeksvariabelen.



De centrale vraag in het onderhavige onderzoek is of deze verwachtingen in dezelfde mate van toepassing zijn op beginners op een bepaald gebied als op personen die reeds enigszins gevorderd zijn. Als scholieren of studenten beginnen met een bepaald school- of studievak waarmee zij weinig of geen eerdere ervaring hebben, hebben zij veelal weinig kennis van de taakvereisten en dus weinig zicht op hetgeen van hen verwacht wordt. Als men geen duidelijk beeld heeft van de taakvereisten, is het niet mogelijk een accuraat oordeel te vormen over eigen bekwaamheden om de taak naar behoren te vervullen. Beginners beschikken dus over relatief weinig aanknopingspunten om de opvattingen over de eigen competentie op te baseren en om adequate persoonlijke doelen te stellen. Hierdoor kan er ook geen goede "monitoring" van de taakuitvoering plaats vinden, is het niet duidelijk welke vaardigheden nodig zijn, hoeveel inzet en volharding nodig is en ook niet in welke opzichten de gehanteerde strategieën bijgesteld dienen te worden (Bandura, 1997). In aansluiting hierbij is gebleken dat de opvattingen over de eigen competentie en de persoonlijke doelen van eerstejaarsstudenten een geringe bijdrage leveren aan de prestaties die behaald worden (Vrugt, Langereis & Hoogstraten, 1993, 1997). Als scholieren of studenten een aantal jaren met een bepaald vak bezig zijn geweest hebben zij daarmee veel ervaring opgedaan, waardoor zij zich een goed beeld hebben kunnen vormen van de taakinhoud en de taakvereisten. Ook hebben zij in de loop van de tijd veel feedback ontvangen over hun kennis en vaardigheden, in de vorm van cijfers en beoordelingen. Hierdoor is duidelijk geworden welke vaardigheden en strategieën nodig zijn en wanneer deze ingezet dienen te worden om tot een goede taakuitvoering te komen. Het beeld van de eigen mogelijkheden en de persoonlijke doelen die nagestreefd worden, kunnen dan afgestemd worden op de taakvereisten. In aansluiting hierbij is gebleken dat de eigen competentie en de persoonlijke doelen van gevorderden op een bepaald terrein, van weten-

schappelijke stafleden op het gebied van onderzoek, goede voorspellers zijn van de prestaties die geruime tijd later behaald worden (Vrugt & Koenis, 1996). Op grond van het voorafgaande verwachten we dat de relaties tussen de motivationele variabelen bij beginners zwak of afwezig zullen zijn, en dat deze relaties bij meer gevorderden relatief sterk zullen zijn.

Eerder kwam aan bod dat neerwaartse en opwaartse vergelijking het functioneren in gunstige zin kunnen beïnvloeden. Volgens de theorie van Wills (1981, 1991) wordt neerwaartse vergelijking vooral ingegeven door de behoefte de eigenwaarde te versterken. Volgens Collins (1996) kan het streven naar zelfverbetering aanleiding geven tot opwaartse vergelijking. Omdat beginners nog weinig duidelijkheid hebben over hun eigen mogelijkheden in relatie tot de relevante taak, zullen zij onderhevig zijn aan onzekerheid en twijfel over hun functioneren op het taakdomein. Hierdoor zal bij beginners de behoefte om de eigenwaarde te versterken prevaleren. Uit onderzoek van Vrugt (1994) blijkt inderdaad dat beginners gemotiveerd worden door neerwaartse vergelijking. Gevorderden zullen minder onderhevig zijn aan twijfel over hun mogelijkheden met betrekking tot het relevante taakdomein, waardoor zij in de gelegenheid zijn om te streven naar zelfverbetering. De resultaten van Vrugt en Koenis (1996) wijzen uit dat gevorderden gemotiveerd worden door opwaartse vergelijking. We verwachten dan ook dat een ego-oriëntatie bij beginners zal negatief samenhangen met sociale vergelijking en bij gevorderden positief. Bij beginners zal neerwaartse vergelijking bijdragen aan de persoonlijke doelen die worden nagestreefd, bij gevorderden zal opwaartse vergelijking daaraan bijdragen.

METHODE

Deelnemers

Aan het onderzoek is meegewerkt door 150 leerlingen van een middelbare scholengemeenschap in Rotterdam. Zevenenzeventig leerlingen zaten tijdens de afname van het onderzoek in de eerste klas van het VWO, deze leerlingen zijn beschouwd als "beginners". De leeftijd van deze leerlingen varieerde van 12 tot 14 jaar. Drieenzeventig leerlingen zaten in de vijfde klas van het VWO, deze leerlingen zijn beschouwd als "gevorderden". De leeftijd van deze leerlingen varieerde van 16 tot 19 jaar.

Variabelen

De deelnemers hebben klassikaal achtereenvolgens een aantal schriftelijke vragen beantwoord over de opvattingen over de eigen competentie, sociale vergelijking, de doelen die zij nastreven en doeloriëntaties (Aan het eind van de vragenlijst hebben zij tevens hun sekse en leeftijd ingevuld). De vragen die gesteld zijn hebben betrekking op het vak Nederlands, omdat dit een verplicht vak is voor leerlingen in de eerste en in de vijfde klas. Deze vragen zijn enkele weken voorafgegaan aan een repetitieweek (waarin ook de repetities voor het vak Nederlands waren opgenomen) afgenomen. De cijfers voor de repetities in de repetitieweek wegen relatief zwaar voor het rapportcijfer. Het rapportcijfer is gebruikt als maat voor studieresultaten.

Doeloriëntaties

De doeloriëntaties zijn gemeten met behulp van de 21 vragen die Duda en Nicholls (1992) gebruikt hebben om doeloriëntaties ten aanzien van de studie te meten. De deelnemers werden verzocht bij zichzelf na te gaan wanneer ze zich het meest succesvol voelen bij het vak Nederlands. De bewering "Ik voel mezelf succesvol als:..." werd gevolgd door 21 stellingen die elk voorzien waren van vijfpuntsschalen met als extremen: "duidelijk oneens" (1) en "duidelijk eens" (5). Een voorbeeld van een stelling die taakoriëntatie representeert, is: "...ik mijn uiterste best doe". Een voorbeeld van een stelling die betrekking heeft op ego-oriëntatie is: "...ik anderen overtroef". Er is gebruik gemaakt van twee schalen, een schaal die de taak-

oriëntatie representeert en een schaal die de ego-oriëntatie representeert. De Cronbach's alfa voor de elf items die de ego-oriëntatie representeren is 0.90, voor de tien items die de taak-oriëntatie representeren is alfa 0.86. Bij de analyses is gebruik gemaakt van de somscores.

Opvattingen over de eigen competentie

In overleg met de docent Nederlands is het schoolvak Nederlands onderscheiden in een aantal deelgebieden die overeenkomen met de taken waarvoor de leerlingen bij dit vak gesteld worden: spelling, grammatica, literatuur, tekstverklaren, opstel, spreekbeurt/discussie. Over elke van deze zes deelgebieden is de vraag gesteld hoe goed de leerling(e) zichzelf vindt. Deze vragen zijn voorzien van vijfpuntsschalen met als extremen "heel goed" (5) en "niet goed" (1). De Cronbach's alfa voor deze schaal is 0.57. Bij de analyses is gebruik gemaakt van de somscores.

Sociale Vergelijking

De deelnemer is verzocht bij elk van zes deelgebieden van het vak Nederlands te denken aan een medeleerling(e) met wie hij of zij zich meestal vergelijkt en te beantwoorden hoe goed deze medeleerling(e) is in vergelijking met zichzelf. Daarbij is gebruik gemaakt van zevenpuntsschalen met als extremen "veel beter" (7) en "veel slechter" (1). De Cronbach's alfa voor deze schaal is 0.68. Bij de analyses is gebruik gemaakt van de somscores.

Verder is bepaald hoeveel leerlingen zichzelf vergeleken hebben met een slechtere, gelijke, of betere medeleerling(e). Daartoe is gebruik gemaakt van de gemiddelden over de zes items. Bij de beginners heeft 16% een score gelijk aan of kleiner dan 3.5, bij de gevorderden is dat 22%. Van de beginners heeft 59% een score van 3.5 tot 4.5, bij de gevorderden is dat 50%. Vijfentwintig procent van de beginners en 28% van de gevorderden heeft een score gelijk aan of groter dan 4.5. Hieruit kunnen we concluderen dat de beginners en de gevorderden overwegend gebruik gemaakt hebben van vergelijking met een gelijke medeleerling(e).

Persoonlijke Doelen

Om te bepalen welke persoonlijke doelen de leerlingen nastreven is gebruik gemaakt van vier vragen die ontleend zijn aan Wood en Locke (1989). Er is gevraagd welk cijfer de leerling(e) denkt, hoopt, daadwerkelijk en minimaal probeert te behalen. Daarbij is gebruik gemaakt van tienpuntsschalen, met als extremen 1 en 10. De Cronbach's alfa voor deze schaal is 0.75.

Studieresultaten

Het rapportcijfer voor het vak Nederlands is gebruikt als maat voor studieresultaten. Bij de toekenning van de cijfers is gebruik gemaakt van een schaal van 1 tot en met 10.

RESULTATEN

Er zijn drie eerste klassen en vier vijfde klassen. Om na te gaan of er in de analyse rekening gehouden moet worden met de meer-niveau structuur van de data, zijn intraklasse-correlaties uitgerekend. De via zelfrapportage verkregen scores blijken als onafhankelijk beschouwd te kunnen worden (de hoogste intraklasse-correlatie is 0.043). De rapportcijfers voor vijfde klassers blijken echter niet onafhankelijk te zijn van de leraar die de cijfers geeft. De intraklasse-correlatie is 0.38 voor de vijfde klassers en -0.01 voor de eerste klassers. Om deze afhankelijkheid weg te nemen zijn de rapportcijfers per leerkracht gestandaardiseerd.

Met het computerprogramma LISREL van Jöreskog en Sörbom (1996) is nagegaan of het padmodel een goede beschrijving van de data geeft, en of er verschillen zijn tussen beginners en gevorderden. Door onvolledig ingevulde vragenlijsten zijn voor de LISREL-analyses 141 leerlingen beschikbaar: 76 van Klas 1 (beginners) en 65 van Klas 5 (gevorderden). Model 1 is het padmodel zoals weergegeven in Figuur 1. Uit simultane analyse ("maximum likeli-

hood"-methode) van de data van beginners en gevorderden afzonderlijk blijkt dat de passing van Model 1 redelijk is. De Chi-kwadraat passingsmaat is niet significant ($CHISQ(16) = 21.77$, $p = 0.151$). Ten behoeve van de interpretatie van de genoemde passingsmaten zij het volgende opgemerkt. Een significante Chi-kwadraat waarde wijst op een significant verschil tussen enerzijds de geobserveerde en anderzijds de volgens het model verwachte varianties, covarianties en gemiddelden. Als de gevonden Chi-kwadraat waarde niet significant is, kan men concluderen dat het model een goede beschrijving van de data geeft. De relatieve passingsmaat RMSEA ("Root Mean Square Error of Approximation") houdt meer rekening met de omvang van het gepaste model. Het is een maat voor het verschil tussen model en data per vrijheidsgraad. Volgens een algemeen geaccepteerde vuistregel wijst een RMSEA-waarde onder 0.08 op "redelijke" passing en een RMSEA-waarde onder 0.05 op "goede" passing (zie voor referenties Jöreskog & Sörbom, 1996). De RMSEA voor Model 1 is 0.072. De passing van Model 1 mag dus redelijk genoemd worden.

De relaties tussen onderzoeksvariabelen worden in het padmodel gerepresenteerd door regressiecoëfficiënten. Teneinde na te gaan of de relaties tussen de variabelen voor beginners en gevorderden hetzelfde zijn, wordt Model 1 nog eens gepast met de restrictie van gelijke regressie-coëfficiënten. Dit gemodificeerde model blijkt minder goed te passen ($CHISQ(21) = 32.92$, $p = 0.054$; $RMSEA = 0.090$). Het verschil in Chi-kwadraat is zelf ook Chi-kwadraat verdeeld en kan gebruikt worden als toets op de gelijkheidsrestrictie. Dit verschil is significant ($CHISQ(5) = 11.15$, $p < 0.048$), zodat geconcludeerd moet worden dat de relaties tussen de variabelen in het padmodel verschillen voor beginners en gevorderden. Uit vergelijkbare analyses blijkt dat ook de intercepten voor beginners en gevorderden significant verschillen ($CHISQ(6) = 13.21$, $p < 0.041$), maar de residuele varianties en covarianties niet ($CHISQ(8) = 11.19$, $p < 0.191$).

Uit de redelijke passing van Model 1 zou men kunnen afleiden dat het padmodel in Figuur 1 zowel op beginners als op gevorderden van toepassing is; zij het dat de effecten verschillende waarden aannemen, zoals blijkt uit de significant slechtere passing van de modellen met gelijkheidsrestricties. Echter, nadere beschouwing van de passingsresultaten van Model 1 leert dat de bijdrage van de beginners aan de globale Chi-kwadraat veel groter is ($CHISQ(8) = 14.84$, $p = 0.062$; $RMSEA = 0.107$) dan de bijdrage van de gevorderden ($CHISQ(8) = 6.94$, $p = 0.534$; $RMSEA = 0.000$). Inspectie van LISREL's modificatie-indices wijst uit dat de passing van het model op de data van de beginners verbeterd kan worden door het model uit te breiden met directe effecten van ego-oriëntatie en taakoriëntatie op prestaties. Het aldus gemodificeerde model, Model 2, blijkt beter te passen ($CHISQ(6) = 8.72$, $p = 0.190$; $RMSEA = 0.078$). Het verschil in passing tussen Model 1 en Model 2 blijkt voor de beginners wel, en voor de gevorderden niet significant te zijn.

Tabel 1 geeft de ("maximum likelihood") schattingen voor alle parameters in Model 2 voor de beginners en Model 1 voor de gevorderden. Het quotiënt van de parameterschatting en haar standaardfout is bij benadering standaardnormaal verdeeld. In Tabel 1 zien we dat bij de beginners drie van de vijf verwachte effecten niet significant zijn, en bij de gevorderden twee van de vijf (bij een significantieniveau van 2.5 procent). Voor de duidelijkheid zijn in Figuur 2 en Figuur 3 voor respectievelijk beginners en gevorderden de regressiecoëfficiënten en gestandaardiseerde parameterschattingen weergegeven.

Verwacht werd dat een taakoriëntatie zou bijdragen aan de opvattingen over de eigen competentie en een ego-oriëntatie aan de sociale vergelijking. Ook werd verwacht dat opvattingen over de eigen competentie en sociale vergelijking zouden bijdragen aan de persoonlijke doelen en dat deze doelen zouden doorwerken in de prestaties. Uit Tabel 1 blijkt, bij de regressiecoëfficiënten die betrekking hebben op de gevorderden, dat een taakoriëntatie bijdraagt aan de opvattingen over de eigen competentie. Ook blijkt dat de opvattingen over de eigen competentie bijdragen aan de persoonlijke doelen en dat die doelen op hun beurt doorwer-

kenn in de prestaties. Deze resultaten zijn in overeenstemming met de verwachtingen. Een ego-oriëntatie heeft echter geen bijdrage geleverd aan de sociale vergelijking, en de sociale vergelijking heeft geen bijdrage geleverd aan de persoonlijke doelen. Deze resultaten zijn niet in overeenstemming met de verwachtingen. Bij de beginners zijn er, zoals verwacht, minder significante effecten. De opvattingen over de eigen competentie dragen bij aan de persoonlijke doelen, en deze doelen vinden hun neerslag in de prestaties.

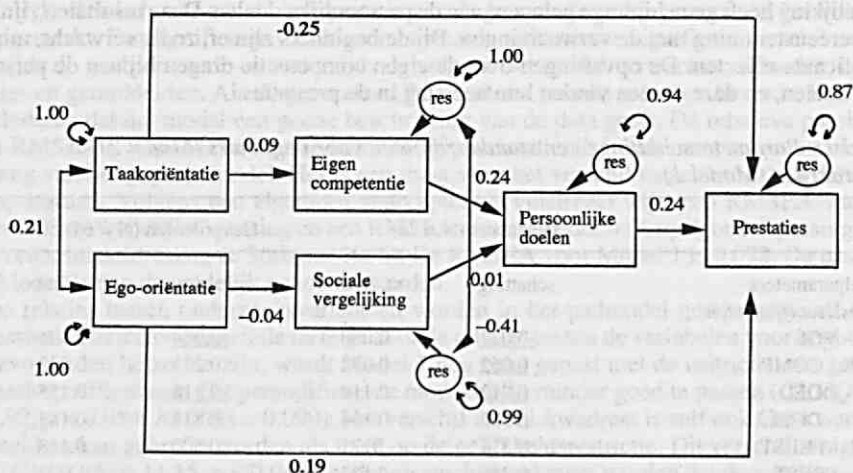
Tabel 1: Parameterschattingen en standaardfouten voor beginners (Model 2) en gevorderden (Model 1).

Modelparameters	Beginners (N = 76)		Gevorderden (N = 65)	
	schatting	st. fout	schatting	st. fout
<i>Regressiecoëfficiënten</i>				
EGO - VGL	<u>-0.027</u>	0.080	<u>-0.084</u>	0.096
TAAK - COMP	<u>0.062</u>	0.072	0.273	0.092
VGL - DOEL	<u>0.012</u>	0.114	<u>-0.116</u>	0.155
COMP - DOEL	0.305	0.154	0.817	0.183
DOEL - PREST	0.371	0.171	0.500	0.118
EGO - PREST	<u>0.210</u>	0.122		
TAAK - PREST	<u>-0.332</u>	0.149		
<i>Variaties</i>				
EGO	0.771	0.126	0.636	0.112
TAAK	0.523	0.085	0.447	0.079
<i>Covariantie</i>				
EGO - TAAK	<u>0.136</u>	0.075	<u>-0.080</u>	0.067
<i>Residuele variaties</i>				
VGL	0.442	0.072	0.438	0.077
COMP	0.241	0.039	0.284	0.050
DOEL	0.357	0.058	0.593	0.105
PREST	0.827	0.135	0.742	0.131
<i>Residuele covariantie</i>				
VGL - COMP	-0.134	0.041	-0.135	0.047
<i>Intercepten</i>				
EGO	3.274	0.101	3.504	0.100
TAAK	3.553	0.084	3.260	0.084
VGL	4.231	0.273	4.348	0.346
COMP	2.930	0.261	2.321	0.308
DOEL	6.730	0.806	4.815	1.006
PREST	-2.381	1.440	-3.479	0.827

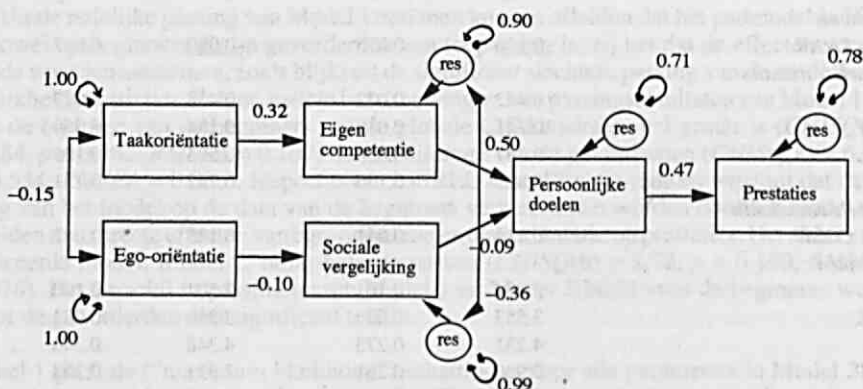
Noot: $CHISQ(6) = 8.72$ ($p = 0.12$) voor beginners ($N_1 = 76$), $CHISQ(8) = 6.94$ ($p = 0.54$) voor gevorderden ($N_2 = 65$); globale passing ($N = N_1 + N_2 = 141$): $CHISQ(14) = 15.66$ ($p = 0.33$); onderstreepte coëfficiënten zijn niet significant (i.e. quotiënt van schatting en standaardfout < 1.96). EGO = ego-oriëntatie, TAAK = taakoriëntatie, COMP = competentie, VGL = sociale vergelijking, DOEL = nagestreefde doelen, PRES = prestatie.

Bij de beginners heeft zich ook een niet-verwachte effect voorgedaan: de taakoriëntatie heeft een directe bijdrage geleverd aan de prestaties. De bijdrage van de taakoriëntatie aan de prestaties is negatief. Verder zien we in Tabel 1 dat de covariantie tussen de taak- en ego-oriëntatie niet significant is. De negatieve correlatie tussen de residuele factoren van eigen competentie en sociale vergelijking is wel significant. Deze resultaten zijn in overeenstemming met de theoretische uitgangspunten.

Figuur 2: Paddiagram van de relaties tussen onderzoeksvariabelen voor beginners (Model 2) met regressie-coëfficiënten en gestandaardiseerde parameterschattingen



Figuur 3: Paddiagram van de relaties tussen onderzoeksvariabelen voor gevorderden (Model 1) met regressie-coëfficiënten en gestandaardiseerde parameterschattingen



DISCUSSIE

Aan ons onderzoek is deelgenomen door scholieren die op basis van hun ervaring met het studievak Nederlands door ons zijn onderscheiden in beginners en gevorderden. Deze beginners en gevorderden verschillen niet alleen wat betreft de ervaring met het studievak, maar ook wat betreft leeftijd en ontwikkelingsfase. De beginners vallen in de leeftijdscategorie van 12 tot 14 jaar en de gevorderden in de categorie van 16 tot 19 jaar. De laatste groep verkeert in een andere fase wat betreft de ontwikkeling van het abstracte denken (Piaget, 1970) en de identiteitsontwikkeling (Erikson, 1968). Hierdoor kan niet uitgesloten worden dat de geconstateerde verschillen tussen beginners en gevorderden ook verband houden met verschillen in ontwikkelingsfase. Als onderzoeksdeelnemers in dezelfde leeftijdscategorie kunnen wor-

den onder gebracht en alleen verschillen in vakkennis, zal de ontwikkelingsfase geen rol spelen. De literatuur biedt echter geen aanknopingspunten om te veronderstellen dat de relaties tussen de opvattingen over de eigen competentie, andere motivationele variabelen en prestaties in een specifiek domein beïnvloed worden door bovengenoemde verschillen in ontwikkelingsfase (zie bijvoorbeeld Bandura, 1997).

De resultaten bieden voor een belangrijk deel steun voor de verwachtingen. Bij beginners en bij gevorderden is geconstateerd dat de opvattingen over de eigen competentie een bijdrage hebben geleverd aan de persoonlijke doelen en dat deze doelen op hun beurt hebben bijgedragen aan de prestaties. In dit opzicht verschillen beginners en gevorderden niet. Wel zijn er duidelijke verschillen in de sterkte van de relaties tussen de motivationele variabelen. De relaties tussen de eigen competentie en de persoonlijke doelen, en tussen de persoonlijke doelen en de prestaties zijn bij beginners veel minder sterk dan bij gevorderden. Dit gegeven onderschrijft de gedachtengang dat beginners door hun beperkte ervaring met het taakgebied nog niet zoveel zicht hebben op de taakvereisten en op hetgeen er van hen verwacht wordt, waardoor het beeld van de eigen mogelijkheden en doelen nog niet goed is afgestemd op de taakvereisten. Bij gevorderden zijn de opvattingen over manier waarop goede prestaties behaald worden, over de eigen mogelijkheden en de persoonlijke doelen duidelijk beter afgestemd op elkaar en op de taakvereisten.

Ook is vastgesteld dat de taakoriëntatie alleen bij gevorderden, en niet bij beginners, een bijdrage heeft geleverd aan de opvattingen over de eigen competentie. Als men gelooft dat hard werken en het begrijpen van de stof belangrijke determinanten zijn van goede prestaties, verschaft dit aanknopingspunten voor de acties die ondernomen kunnen worden. Met toenemende ervaring zullen de benodigde strategieën en vaardigheden ontwikkeld worden, waardoor de opvatting dat hard werken en het begrijpen van de stof van belang zijn voor het behalen van goede prestaties versterkt zal worden. Door deze processen zal de beheersing die men ervaart bij de uitvoering van taken verhoogd worden, hetgeen zijn neerslag vindt in de waargenomen eigen competentie, en doorwerkt in de prestaties. De bijdrage van taakoriëntatie aan de opvattingen over de eigen competentie die bij gevorderden is vastgesteld geeft aan dat er bij gevorderden bepaalde relaties tussen motivationele variabelen werkzaam zijn, die bij beginners nog niet tot ontwikkeling zijn gekomen.

Deze conclusie wordt ondersteund door een niet-verwachte effect dat zich bij beginners heeft voorgedaan: de taakoriëntatie heeft een directe negatieve bijdrage geleverd aan de prestaties. Een sterke taakoriëntatie heeft dus aanleiding gegeven tot relatief slechte prestaties. In eerder onderzoek is, zover als wij weten, geen negatieve relatie vastgesteld tussen een taakoriëntatie en prestaties. Dit resultaat is moeilijk te verklaren. Misschien hangt het samen met het feit dat beginners nog niet goed weten welke strategieën en vaardigheden van hen verwacht worden. Mensen met een taakoriëntatie menen dat hard werken en het begrijpen van de stof zullen resulteren in goede prestaties. Als men de strategieën en vaardigheden voor het tot stand brengen van die prestaties echter nog niet voldoende ontwikkeld heeft, kan een taakoriëntatie aanleiding geven tot een inzet die niet gecombineerd wordt met de juiste aanpak. Dit kan tot verwarring leiden en het prestatieniveau aantasten.

Opvallend is dat niet alleen bij de beginners, maar ook bij de gevorderden een ego-oriëntatie niet heeft bijgedragen aan sociale vergelijking en dat de sociale vergelijking geen bijdrage heeft geleverd aan de nagestreefde doelen. Staalvik (1997) heeft gesuggereerd dat een ego-oriëntatie aanleiding geeft tot sociale vergelijking, maar heeft dit niet onderzocht. Keet (1998) heeft deze kwestie bestudeerd bij scholieren in de vierde klas van de middelbare school, in het kader van het vak Nederlands. Deze leerlingen kunnen ook beschouwd worden als gevorderden. Zij heeft vastgesteld dat een ego-oriëntatie verband hield met het belang dat men hechtte aan sociale vergelijking, en ook dat een ego-oriëntatie verband hield met de richting van de vergelijking. De meeste leerlingen vergeleken zichzelf met een betere medeleerling(e). Deze resultaten komen wel overeen met onze verwachtingen, maar verschillen duidelijk van de resultaten van het huidige onderzoek, ook omdat in ons onderzoek de gevorderden

(en de beginners) zichzelf in meerderheid vergeleken hebben met een gelijke medeleerling(e). Het is vooral nog niet duidelijk waaraan de verschillen tussen de gevorderden in het onderhavige onderzoek en in het onderzoek van Keet (1998) toegeschreven kunnen worden.

Voor de volledigheid dient opgemerkt te worden dat het onderhavige onderzoek een correlatief onderzoek is waarin modellen getoetst zijn die uitgaan van causale relaties. De steun die gevonden is voor de modellen verschaft als zodanig onvoldoende evidentie om te kunnen concluderen dat deze causale relaties ook daadwerkelijk werkzaam zijn. Hiervoor is vervolgonderzoek met een experimentele opzet nodig.

REFERENTIES

- Atkinson, J. W. (1964). *An introduction to motivation*. Princeton: van Nostrand.
- Atkinson, J. K. & Raynor, J. O. (1974). *Motivation and achievement*. New York: Wiley.
- Affleck, G., Tennen, H., Pfeiffer, C., Fifield, J., et al. (1987). Downward comparison and coping with serious medical problems. *American Journal of Orthopsychiatry*, 57, 570-578.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall, inc.
- Bandura, A. (1989). Self-regulation of motivation and action through internal standards and goal systems. In L. A. Pervin (Ed.), *Goal concepts in personality and social psychology*. Hillsdale NJ.: Lawrence Erlbaum Associates.
- Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: W. H. Freeman and Company.
- Collins, R. L. (1996). For better for worse: the impact of upward social comparison on self-evaluations. *Psychological Bulletin*, 119, 51-69.
- Covington, M. V. (1992). *Making the grade: a self-worth perspective on motivation and school reform*. Cambridge: Cambridge University Press.
- Deffenbacher, J. L. (1978). Worry, emotionality, and task-generated interference in test anxiety: an empirical test of attentional theory. *Journal of Educational Psychology*, 70, 2, 248-254.
- Duda, J. L., & Nicholls, J. G. (1992). Dimensions of achievement motivation in school work and sport. *Journal of Educational Psychology*, 84, 290-299.
- Elliot, E. S., & Dweck, C. S. (1988). Goals: an approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54, 5-12.
- Elshout, J. J., & Veenman, M. V. J. (1992). Relation between intellectual ability and working method as predictors of learning. *Journal of Educational Research*, 85, 134-143.
- Erikson, E. (1968). *Identity: youth and crisis*. New York: Norton.
- Gist, M. E. (1989). The influence of training method on self-efficacy and idea generation among managers. *Personnel Psychology*, 42, 787-805.
- Gibbons, F. X., & Gerrard, M. (1991). Downward comparison and coping with threat. In: J. Suls & T. A. Wills (Eds.), *Social comparison: contemporary theory and research*. London: Lawrence Erlbaum.
- Guldemond, H., Meijnen, G. W., & Boomsma, A. (1989). Gezin, school en ontwikkelingen in intelligentie tussen zes en twaalf jaar. *Nederlands Tijdschrift voor de Psychologie*, 44, 181-191.
- Hofmann, D. A. & Strickland, O. J. (1995). Task performance and satisfaction: Evidence for a task- by ego-orientation interaction. *Journal of Applied Social Psychology*, 25, 495-511.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8 user's guide* (2nd ed.). Chicago, IL: Scientific Software International, Inc.
- Jourden, F. J., Bandura, A., & Banfield, J. T. (1991). The impact of conceptions of ability on self-regulatory factors on motor skill acquisition. *Journal of Sport and Exercise Psychology*, 8, 211-226.
- Keet, A. (1998). *De invloed van waargenomen eigen competentie, doelen, doeloriëntaties, sociale vergelijking en gevoelens op schoolprestaties*. Doctoraal werkstuk, vakgroep Sociale Psychologie, Universiteit van Amsterdam.
- Liebert, R. M. & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, 20, 975-978.
- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs NJ: Prentice Hall.
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: a meta-analytic investigation. *Journal of Counseling Psychology*, 38, 30-38.

- Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (vol. 1). New York: Wiley.
- Schunk, D. H., & Swartz, C. W. (1993). Goals and progress feedback: effects on self-efficacy and writing achievement. *Contemporary Educational Psychology*, 18, 337-354.
- Staalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: relations with task and avoidance orientation, achievement, self-perceptions and anxiety. *Journal of Educational Psychology*, 89, 71-81.
- Tuckman, B. W., & Sexton, T. L. (1992). Self-believers are self-motivated, self-doubters are not. *Personality and Individual Differences*, 13, 425-428.
- Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90, 245-271.
- Wills, T. A. (1991). Similarity and self-esteem in downward comparison. In: J. Suls & T. A. Wills (Eds.), *Social comparison: contemporary theory and research*. London: Lawrence Erlbaum.
- Vrugt, A. (1994). Perceived self-efficacy, social comparison, affective reactions and academic performance. *British Journal of Educational Psychology*, 64, 465-472.
- Vrugt, A. (1996). Perceived self-efficacy, work motivation and well-being. In: M. J. Schabracq, J. A. M. Winnubst & C. L. Cooper (Eds.), *Handbook of work and health psychology*. London: John Wiley & Sons.
- Vrugt, A., Hoogstraten, Joh., & Oort, F. J. (1999). Doeloriëntaties, waargenomen eigen competentie en studieresultaten. *Tijdschrift voor Onderwijsresearch*, 23, 210-223.
- Vrugt, A., & Koenis, S. (1996). Waargenomen eigen competentie, persoonlijke doelen, sociale vergelijking en wetenschappelijke productiviteit. *Gedrag en Organisatie*, 10, 134-144.
- Vrugt, A., Langereis, M., & Hoogstraten, Joh. (1993). De veronderstelde bekwaamheid als voorspeller van studieresultaten. *Tijdschrift voor Onderwijs Research*, 18, 42-51.
- Vrugt, A., Langereis, M. & Hoogstraten, Joh. (1993). Perceived self-efficacy, perceived learnability of capabilities and academic performance. In: W. Weblor & J. Koppen (Eds.), *Strategies for increasing access to and performance in higher education*. Amsterdam: Thesis Publishers.
- Vrugt, A., Langereis, M., & Hoogstraten, Joh. (1997). Academic self-efficacy and malleability of relevant capabilities as predictors of exam performance. *Journal of Experimental Education*, 66, 61-74.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer Verlag.
- Wills, T. A. (1991). Similarity and self-esteem in downward comparison. In: J. Suls & T. A. Wills (Eds.), *Social comparison: contemporary theory and research*. London: Lawrence Erlbaum.
- Wood, R., & Bandura, W. (1989). Impact of conceptions of ability on self-regulatory mechanisms and complex decision making. *Journal of Personality and Social Psychology*, 56, 407-415.
- 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.

Boekbesprekingen

J. Buitink

In-functie opleiden en in-functie leren van aanstaande docenten

Proefschrift, Rijksuniversiteit Groningen, 1998, ISBN 90-367-0892-3

Achtergrond en vraagstelling

In het proefschrift wordt verslag gedaan van een onderzoek naar het leren lesgeven van aanstaande wiskunde-leraren tijdens een in-functie-opleiding aan de universitaire lerarenopleiding van de Rijksuniversiteit Groningen in het schooljaar 1993-1994. Aanleiding tot het onderzoek vormde de in 1990 ingestelde, en inmiddels weer afgeschafte, faciliteitenregeling voor de zogenoemde 'tekortvakken': in deze vakken werkten leraren-in-opleiding (LIO's) gedurende het gehele opleidingsjaar als docenten in een school, waarbij zij volledige verantwoordelijkheid droegen voor een of meer klassen en in een team functioneerden als volwaardige collega's, met alle rechten en plichten. Zij ontvingen tevens een salaris. Voor door het opleidingsinstituut te verzorgen opleidingsactiviteiten op het gebied van de algemene onderwijskunde en vakdidactiek werden de LIO's anderhalve dag per week uitgeroosterd. Naast werkcolleges, het maken van producten e.d. werd door het opleidingsinstituut het leerproces van de LIO's bewaakt en vond stagebegeleiding plaats.

Tegen bovenstaande achtergrond worden in het proefschrift twee hoofdvragen beantwoord. Beide vragen spitsen zich toe op de werktheorie van de LIO's, respectievelijk de theorie die in de praktijk door de LIO's gebruikt wordt. De eerste vraag richt zich op de aard van de werktheorie van de LIO's en de veranderingen die daarin optreden gedurende de in-functie opleiding. Het gaat hierbij vooral om veranderingen in de inhoud van de werktheorie, de rijkheid van de werktheorie en de relaties die de LIO's leggen tussen begrippen in hun werktheorie. De tweede vraag richt zich op de manier waarop LIO's hun werktheorie veranderen. Bij deze vraag gaat het vooral om de rol van de leeromgeving, de doelbeleving door de LIO's, de wijze waarop de LIO's vinden dat zij leren lesgeven en de fasering van het leerproces.

Onderbouwing en opzet van de studie

In de studie wordt onderwijzen opgevat als complex handelen. De invloed van de context op de wijze waarop een handeling wordt uitgevoerd, maakt dat onderwijzen als complex handelen kan worden gekarakteriseerd. Van extra complexiteit is sprake door de zwakke structurering van de onderwijskundige theorie op het gebied van (leren) onderwijzen. Leren onderwijzen wordt derhalve opgevat als het leren van complexe handelingen, hetgeen het best kan plaatsvinden in de context waarin deze handelingen moeten worden toegepast. In het proefschrift wordt gesproken van leren in een authentieke context, dat wil zeggen dat de context waarin dit leren plaatsvindt identiek is aan de alledaagse werksituatie van de leraar.

Het begrip werktheorie neemt in de studie een centrale plaats in (zie ook het Raamplan voor de Universitaire Lerarenopleiding, 1996). Daarmee wordt aangesloten bij recente opvattingen over cognities en praktijkkennis van leraren. Het onderzoek naar cognities en praktijkkennis van leraren is vaak kleinschalig en kwalitatief van aard. Door Buitink zijn acht LIO's gedurende hun opleiding intensief gevolgd. Op verschillende momenten tijdens de opleiding zijn van deze LIO's gegevens verzameld. Leerverslagen, concept maps en planning- en evaluatieverslagen zijn gebruikt voor het krijgen van inzicht in de aard van de werktheorie en de veranderingen daarin (onderzoeksvraag 1). Voor het krijgen van inzicht in de wijze waarop de werktheorie van LIO's verandert (onderzoeksvraag 2), is gebruik gemaakt van interviews en concept maps. In het onderzoek is zowel sprake van methoden- als data-triangulariteit.

In het proefschrift worden de acht LIO's als afzonderlijke cases onafhankelijk van elkaar en volgens een identieke systematiek beschreven (intra case-studies). Vervolgens zijn ze

onderling volgens een vastgestelde systematiek geanalyseerd (inter case-study) om bepaalde patronen op te sporen in termen van overeenkomsten en verschillen. Omdat de acht cases uiteindelijk worden gezien in hun onderlinge samenhang, wordt in het proefschrift gesproken van een meervoudige case-study. Analyse en verwerking van de ruwe data vonden plaats door het toekennen van codes aan teksteenheden. Aan de basis hiervan lag een ontwikkeld categorieënsysteem voor elke onderzoeksvraag. Op deze wijze konden de data nauwkeurig in (within- en cross-case) displays worden weergegeven.

Belangrijkste resultaten

Met betrekking tot de aard van de werktheorie van LIO's en de veranderingen daarbinnen blijkt dat LIO's al vroeg in de opleiding beschikken over redelijk uitgekristalliseerde werktheorieën. Zij hebben bijvoorbeeld al gauw in de gaten dat lesgeven een samenspel is van een groot aantal factoren. Alle LIO's blijken hun werktheorie te ontwikkelen tot een rijke werktheorie bestaande uit veel categorieën van het lesgeven, waarin gaandeweg de opleiding een overzichtelijke structuur ontstaat. De inhoud van de werktheorie weerspiegelt bij zeven van de acht LIO's meer dan de alledaagse onderwijspraktijk. Bij een aantal van hen bestaat ook nadrukkelijk aandacht voor zelfstandig leren.

Het leerproces van de LIO's vertoont grote individuele verschillen wat betreft doelbeleving, de opstelling in het leerproces, de fasering van het lesgeven en de beleving van de leeromgeving. De LIO's stellen zich of actief of passief op in het leerproces. Het komt voor dat LIO's hun opstelling in het leerproces tijdens de opleiding veranderen van passief naar actief. Drie groepen LIO's worden onderscheiden op grond van hun manieren van leren. Gesproken wordt van drie leerstijlen, te weten: de aanpassingsstijl, de onderzoekende stijl en de vernieuwende stijl. Er blijkt geen samenhang te zijn tussen door LIO's gehanteerde leerstijlen en de aard van de door hen ontwikkelde werktheorie. Met andere woorden, het lijkt voor het leereffect van de opleiding niet veel uit te maken welke leerstijl een LIO hanteert.

Aan het eind van het proefschrift wordt gesteld dat het bereikte leereffect in belangrijke mate is toe te schrijven aan de wijze waarop de leeromgeving is ingericht. Daarbij wordt gewezen op activiteiten die zowel in de school als in het opleidingsinstituut plaatsvinden. De relatie tussen leeromgeving en leerstijl is minder duidelijk. Diverse maatregelen worden tenslotte genoemd op grond waarvan in een in-functie opleiding een meer expliciete invloed op de leerstijl van LIO's kan worden uitgeoefend.

Commentaar

Naar het leerproces van leraren-in-opleiding is (in Nederland) tot op heden weinig onderzoek gedaan. In het onderzoek spitst zich dit toe op de constructie van kennis en de wijze waarop dat plaatsvindt. Als zodanig is sprake van een studie die een trend zet en navolging verdient. De context van een in-functie opleiding wordt een steeds meer voor de hand liggende context om het leren onderwijzen te bestuderen. Hoewel deze op dit moment nogal afwijkt van de meeste andere universitaire lerarenopleidingen, is ook in deze opleidingen in toenemende mate sprake van LIO-trajecten waarin (betaald) werken en leren hand in hand gaan en het werken alle kenmerken vertoont van een leraar in functie met alle verantwoordelijkheden van dien. In theoretisch opzicht sluit het onderzoek goed aan bij recente opvattingen in het domein van het (leren) onderwijzen, waaronder de opvatting dat onderwijzen een complexe activiteit is. De relatie tussen het begrip werktheorie en de huidige opvattingen over de kennisbasis van en voor leraren had echter meer uitgewerkt kunnen worden.

Zoals gezegd, worden de resultaten van het onderzoek voor een belangrijk deel toegeschreven aan de leeromgeving. In het proefschrift wordt deze omgeving als 'krachtig' en 'cruciaal' aangeduid. Als reden hiervoor wordt genoemd het feit dat: (1) de leeromgeving sturing geeft aan het leerproces, (2) de leeromgeving neutraal tot positief wordt beoordeeld door de LIO's en (3) aspecten die in het algemeen kenmerkend zijn voor een krachtige leeromgeving (ontleend aan De Corte, 1992) ook grotendeels voorkomen in de Groningse in-functie oplei-

ding. Aangezien de leeromgeving zelf geen object van onderzoek is geweest, lijken dit tamelijk vage indicatoren te zijn om deze omgeving zo stellig als krachtig en cruciaal te kwalificeren.

Sprake is van case-study onderzoek: elke afzonderlijke case wordt tot in de kern geanalyseerd; het specifieke geval is het vertrekpunt. Bij de reconstructie van gegevens spelen twee aspecten van dataverwerking een rol: (1) de reductie van de hoeveelheid data per case tot overzichtelijke cases en (2) de interpretatie van data, waarbij de onderzoeker zijn kennis gebruikt om de cases te beschrijven en theoretische en praktische gegevens aan elkaar verbindt. Aan het eerste aspect, de reductie van data, is door middel van een overzichtelijk coderingssysteem en het onderbrengen van gegevens in displays voldaan. De wijze waarop dit is gebeurd is goed navolgbaar (Smaling, 1987). Als het gaat om de interpretatie van de data, dan is deze navolgbaarheid minder duidelijk. Dit betreft niet zozeer de beantwoording van de eerste onderzoeksvraag naar de aard van de werktheorie van LIO's en de veranderingen die daarin optreden, maar vooral de vraag naar de kenmerken van de manier waarop de werktheorie van LIO's verandert. Er wordt naar verhouding veel theorie gebruikt voor de interpretatie van data, waardoor generalisaties worden gemaakt waarin de afzonderlijke cases niet of moeilijk meer herkenbaar zijn. Aan deze eis van case-study onderzoek lijkt op onderdelen onvoldoende recht te zijn gedaan.

Tot besluit

Ondanks bovenstaande kanttekeningen is sprake van een verdienstelijk proefschrift. De studie heeft de praktijk als aangrijpingspunt genomen. De onderzoeker heeft zich daarmee niet in een eenvoudige positie gemanoeuvreerd, zoals ook in het proefschrift wordt beschreven: de opleiding zoals beoogd week op onderdelen aanzienlijk af van de wijze waarop die uiteindelijk is gerealiseerd. Het hoeft geen betoog dat het in een situatie waarin onderzoek afhankelijk is van ontwikkelingen of veranderingen in de praktijk niet eenvoudig is het doel van het onderzoek overeind te houden. Dat dit is gelukt is waarschijnlijk mede te danken aan het feit dat de onderzoeker tevens opleider van de onderzochte LIO's was. Hoewel deze dubbele positie zich bij veel onderzoek in lerarenopleidingen voordoet, is ook dit niet altijd eenvoudig. Het zou daarom interessant zijn wanneer over het effect van deze dubbele betrokkenheid van onderzoekers meer wordt geschreven in proefschriften.

Literatuur

- Corte, E. de (1992). Design and evaluation of powerful learning environments. In B.P.M. Creemers & G.J. Reezigt (Eds.), *Evaluation of educational effectiveness*. Groningen: ICO.
- Smaling, A. (1987). *Methodologische objectiviteit en kwalitatief onderzoek*. Lisse: Swets & Zeitlinger.
- VSNU (1996). *Raamplan voor de universitaire lerarenopleiding*. Utrecht: VSNU.

D. Beijaard

ICLON, Universiteit Leiden

M. Gemmeke (1998)

Politieke betrokkenheid van kinderen op de basisschool

Thesis Publishers Amsterdam

ISBN 90-9011-668-0

Achtergrond en vraagstelling

Het thema van het proefschrift *Politieke betrokkenheid van kinderen op de basisschool* is wetenschappelijk en maatschappelijk van belang. Centraal staan concepten als burgerschap, politieke vorming en politieke socialisatie. De auteur richt zich op de vraag hoe mensen zich ontwikkelen (of niet) tot zoiets als een goed burger (Van Gunsteren, 1992). Velen maken zich

zorgen over bijvoorbeeld de inburgering van nieuwkomers in onze samenleving en de dalende opkomst bij verkiezingen voor diverse politieke gremia.

In theorieën van cognitieve ontwikkeling ligt besloten dat vanaf een bepaalde biologische leeftijd politieke bewustwording mogelijk wordt. Vanuit socialisatie-theorie is het interessant te kijken naar pogingen van bijvoorbeeld gezinsleden of onderwijsgeevenden om deze politieke betrokkenheid te beïnvloeden. Ook rijst de vraag of factoren als geslacht, etniciteit en de sociale positie van ouders verschillen kunnen verklaren in politieke socialisatie en betrokkenheid.

In onderwijskundig onderzoek krijgt de ontwikkeling van specifiek politieke competenties van leerlingen relatief weinig aandacht. De politieke wetenschappen daarentegen richten zich wel op politieke cognities, affecties en gedragingen, maar baseren hun bevindingen overwegend op onderzoek onder kiesgerechtigden - dus volwassenen. Het is prijzenswaardig dat Gemmeke tracht inzichten uit beide disciplines te verenigen en zich richt op kinderen van 8 tot 13 jaar oud: een populatie waarvan nauwelijks bekend is in welke mate deze bewust omgaat met politieke materie.

De auteur "*beoogt inzicht te geven in de vraag welke kinderen onder welke invloeden betrokken raken bij de politiek*" (pp.2). Er zijn drie hoofdvragen (pp.22): In welke mate bepalen de cognitieve ontwikkeling, gender en leerkrachten de ontwikkeling van politieke betrokkenheid bij kinderen van 7 tot 13 jaar? Verschilt de ontwikkeling van politieke betrokkenheid naar etnische groep? Verschilt de mate van politieke betrokkenheid naar klassenomgeving?

Opzet van de studie

De conceptualisering van politieke betrokkenheid getuigt van een gedegen analyse van een grote hoeveelheid literatuur. Terecht wordt gesteld "*dat door vrijwel geen van de auteurs [...] aandacht wordt besteed aan de conceptuele en operationele overlap tussen begrippen*" (pp. 48). Gemmeke ontwikkelt een instrument met drie schalen: cognitieve, emotionele en gedragsmatige politieke betrokkenheid. De cognitieve betrokkenheid wordt gemeten door enerzijds te vragen naar de subjectieve bekendheid met 'problemen' als de Europese eenwording of oorlog en anderzijds het kunnen herkennen van Ruud Lubbers op een foto, het geven van het juiste antwoord op de vraag wie er het meeste te zeggen heeft over wat er in Nederland gebeurt en het kunnen herkennen van politieke partijen uit een lijst met afkortingen. De emotionele betrokkenheid is gemeten door te vragen in welke mate de 'problemen' belangrijk zijn, bezorgdheid oproepen, of saai worden geacht. Voorbeelden van gedragsmatige politieke betrokkenheid zijn het lezen over politiek in de krant, het volgen van nieuws en actualiteitenprogramma's en het praten met anderen over politiek.

De data-verzameling bestaat uit twee fasen: een vooronderzoek bestaande uit diepte-interviews met 22 leerlingen op vier Amsterdamse basisscholen. Deze fase leidt tot de conclusie dat de oudste leerlingen op basisscholen een mening hebben en deze kunnen uiten als het gaat om zaken als oorlog, discriminatie en het milieu. In de hoofdfase is een schriftelijke enquête afgenomen bij 475 leerlingen in de groepen 4, 6 en 8 op acht Amsterdamse basisscholen. Het gaat om een bewust geselecteerd aantal van vier 'zwarte' en vier 'witte' basisscholen in Amsterdam: het doel is tot uitspraken te komen over verschillen in betrokkenheid tussen autochtone en allochtone kinderen. Om een link te leggen tussen kenmerken op leerkrachten leerlingniveau worden ook de twintig betrokken groepsleerkrachten benaderd met een schriftelijke enquête: de betreffende analyse vindt plaats via multivariate regressie.

Belangrijkste resultaten

In welke mate bepalen de algemene cognitieve ontwikkeling en gender de ontwikkeling van politieke betrokkenheid bij kinderen van 7 tot 13 jaar? Gemmeke operationaliseert cognitieve ontwikkeling door te kijken naar verschillen tussen de basisschool-leerjaren (groepen 4, 6 en 8); de oudere leerlingen hebben meer politieke kennis, kijken vaker naar het journaal en de televisie in het algemeen en zeggen vaker bekend te zijn met discriminatie, werkloosheid

en milieuvervuiling. Op het emotionele vlak verschilt de politieke betrokkenheid van de jaargroepen niet. De politieke betrokkenheid van jongens en meisjes verschilt niet.

Verschilt de ontwikkeling van politieke betrokkenheid naar etnische groep? Grofweg gezegd is uitkomst van het onderzoek dat de cognitieve kant van autochtonen beter ontwikkeld is, terwijl bij allochtonen de emotionele betrokkenheid bij de maatschappelijke problemen groter is. Allochtone en autochtone kinderen informeren zich evenveel over maatschappelijke problemen en kijken even vaak naar het (jeugd)journaal.

Verschilt de mate van politieke betrokkenheid naar klassenomgeving, oftewel is er nu sprake van politieke socialisatie op de basisschool? Zo ja, zijn het juist de leerlingen uit achterstandsmilieus die hiervan profiteren - doordat zij thuis wellicht nauwelijks met politieke kennis en houdingen geconfronteerd worden? Of profiteren geprivilegerde kinderen het meest van les over politiek? De voorzichtige conclusie luidt dat de politieke betrokkenheid van leerlingen toeneemt bij een stimulerende leerkracht. De leerkracht heeft meer invloed op de politieke betrokkenheid van allochtone kinderen dan op die van autochtonen.

Commentaar

Onderzoek onder 12 tot 18-jarigen maakt duidelijk dat zij veel meer geïnteresseerd zijn in problemen als milieuvervuiling en discriminatie dan in de nationale politieke instellingen, actoren en processen (o.a. Dekker & Portengen, 1995). Het is ook zo dat met name jongens en degenen uit de hogere sociale klassen geïnteresseerd zijn in de blijkbaar 'saaie' landelijke politiek. Het antwoord op de vraag in welke mate welke jongeren wel of niet voldoende politiek betrokkenheid tonen, hangt dus mede af van de definitie van politiek. Gemmeke zegt politieke en maatschappelijke betrokkenheid te willen scheiden, maar ze noemt oorlog, armoede, milieuvervuiling, criminaliteit, immigratie, discriminatie, grenzen en moslims *politieke* problemen. De reden: deze problemen zijn onderwerp van overheidsbeleid (pp. 98). Is er echter een maatschappelijk probleem te bedenken waar de overheid zich niet mee bemoeit? Terwijl de auteur een smalle definitie van politiek voorstaat, is deze waterscheiding tussen politiek en maatschappij flinterdun. Hierdoor weten we nu dat 8-13 jarigen betrokkenheid tonen bij iets als de oorlog in Bosnië, maar minder duidelijk is in welke mate zij betrokken zijn bij de 'saaie' landelijke politieke instellingen en actoren. Al met al had een titel van het proefschrift als '*Maatschappelijke betrokkenheid van kinderen op de basisschool*' toepasselijker geklonken.

Ook beperkingen in het onderzoeksdesign spelen het werk parten. In het proefschrift van Van den Broek (1996) over burgerschap en socialisatie wordt gesteld, dat voor het onderscheiden van cohort, leeftijds- en periode-effecten zelfs in geval van meerdere metingen in de tijd een methodologisch complexe procedure nodig is. Gemmeke trekt conclusies over de *ontwikkeling* van betrokkenheid op basis van één meting. Het is echter niet hard te maken dat verschillende groepsgegevens in betrokkenheid veroorzaakt zijn door de algemene cognitieve ontwikkeling. Ten eerste ontbreekt sowieso een meting van iets als het IQ. Ten tweede kunnen verschillen ook ontstaan doordat bijvoorbeeld de oudere leerlingen iets hebben meegemaakt, waardoor zij op een andere manier betrokken zijn geraakt bij de politiek en maatschappelijke problemen. Gemmeke beseft dit en bespreekt belangrijke gebeurtenissen in de wereld, Nederland en Amsterdam in het begin van de jaren negentig: de periode voorafgaand aan de meting. Ze stelt dat alle kinderen deze gebeurtenissen hebben meegemaakt en daarom groepsverschillen niet toegeschreven kunnen worden aan historische momenten. De kinderen die in 1993 in groep 8 zaten, zullen de oorlogen in voormalig Joegoslavië, Irak en Somalië echter *bewuster* hebben gemaakt dan de kinderen in groep 4. Theoretisch is het daarom mogelijk dat groepsverschillen niet puur betrekking hebben op cognitieve vermogens van dat moment, maar ook op effecten van het lid zijn van een geboorte-cohort. Kortom: cohort- en leeftijds-effecten zijn niet van elkaar te scheiden in dit design.

De gevonden verschillen in politieke betrokkenheid tussen de allochtonen en autochtonen kunnen verder rechtstreeks samenhangen met de steekproef van scholen: het gaat immers om specifiek witte en zwarte scholen in een grootstedelijk gebied. Zo zal een deel van de ouders

op de vier onderzochte zwarte scholen waarschijnlijk helemaal geen stemrecht hebben. Is het dan vreemd te veronderstellen, dat de nationale politiek meer een issue is voor de autochtonen? Het resultaat dat de 'witten' meer cognitief betrokken zijn bij de politiek is dan verklaarbaar, evenals het gegeven dat 'zwarte' leerlingen emotioneler worden als het gaat over zaken als armoede, discriminatie, moslims en grenzen. Dat in dit kleinschalige onderzoek scherpe verschillen worden gevonden, betekent niet dat dit verband tussen betrokkenheid en etniciteit in de rest van Nederland of in Amsterdamse scholen met een gemengde leerlingpopulatie opgaat. Het is zelfs waarschijnlijk dat deels een klasse-effect is gemeten, daar waar gesproken wordt over verschillen tussen etniciteiten. Het soms gehoorde adagio dat voor onderzoek naar verbanden tussen variabelen de representativiteit van een steekproef minder belangrijk is, gaat hier niet op.

Conclusie

Onderzoek naar de politieke socialisatie van kinderen is in (inter)nationaal opzicht schaars. Het combineren daarbij van inzichten uit de politicologie en onderwijskunde is een loffelijk streven. Het ontwikkelen van meet-instrumenten voor dergelijk onderzoek onder kinderen van 7 tot 13 jaar oud is een lastige zaak. Op deze punten vormt het proefschrift van Gemmeke een interessante en goede aanzet. Het is jammer dat de tekortkomingen in het onderzoeksdesign niet voldoende worden erkend en de auteur de neiging heeft conclusies te trekken die los van de data staan. Met minder pretenties had de inhoud meer aandacht gekregen.

Literatuur

- Van den Broek, A. (1996). *Politics and generations: cohort replacement and generation formation in political culture in the Netherlands*. Tilburg University Press: Tilburg.
- Dekker, H. & R. Portengen (1995). Explaining voting intentions among Dutch youth. In *Politics, Groups and the Individual*. 5(2): 81-100.
- Van Gunsteren, H. (1992). *Eigentijds burgerschap*. SDU uitgeverij, Den Haag.

Renze Portengen
ITS Nijmegen

Maarten H.J. Wolbers

Diploma-inflatie en verdringing op de arbeidsmarkt. Een studie naar ontwikkelingen in de opbrengsten van diploma's in Nederland.
Proefschrift K.U. Nijmegen. Amsterdam: Thesis, 1998.

Tot de spannendste discussies in de sociale wetenschappen behoort de vraag in welke gevallen er sprake is van competitie om *positionele goederen*. Dit zijn goederen waarvan de waarde afhangt van de relatieve mate waarin iemand ze bezit, dus bijvoorbeeld het relatieve welstandsniveau, de relatieve statuspositie etc. Kenmerkend voor deze goederen is dat hun schaarste niet kan worden teruggedrongen, want zodra iemand er meer van krijgt gaat iemand anders er noodzakelijkerwijs op achteruit. Het mogelijke bestaan van positionele goederen - het is bepaald geen onbetwist concept - zet de economische logica op zijn kop. In normale omstandigheden leidt concurrentie tot een efficiëntere besteding van middelen en daarmee tot een grotere welvaart. In het geval van positionele goederen leidt concurrentie echter slechts tot verspilling van middelen en zou iedereen erbij winnen als aan de concurrentie een halt werd toegeroepen. Er is in die situatie sprake van een mogelijkheid tot een Pareto-optimale verbetering, zoals de economen dat noemen. De vraag of ook onderwijs, of preciezer gezegd: het te bereiken opleidingsniveau, een positioneel goed is, houdt economen en sociologen al geruime tijd bezig. Maarten Wolbers stelt in zijn proefschrift uitdrukkelijk van wel (p. 128). De onuitgesproken, maar voor iedere lezer overduidelijke strekking van zijn verhaal is dan

ook, dat iedereen in Nederland er op vooruit zou gaan als iedereen een jaartje minder onderwijs zou volgen.

Wolbers' onderzoek naar diploma-inflatie en verdringing op de arbeidsmarkt staat in de traditie van het onderzoek van Fred Huijgen en anderen naar de ontwikkeling van de functieniveaustructuur van de Nederlandse beroepsbevolking. Dit onderzoek, waarvan de eerste rapportages nu bijna twintig jaar oud zijn, heeft altijd uitgewezen dat er sprake is van een groeiende kloof tussen het opleidingsniveau en het functieniveau van de Nederlanders. Hoewel ook het gemiddelde functieniveau sinds 1960 (Algemene Volkstelling: de oudste gegevens) fors is toegenomen, neemt het opleidingsniveau sneller toe, met als gevolg dat een toenemend aantal mensen werk moet aanvaarden beneden hun niveau. De laatste rapportage langs deze lijn stamt van 1998 (Asselberghs et al.) en wees op een weer verder toegenomen kloof.

Het was een goed idee van Wolbers en zijn promotoren om door middel van *cohortanalyse* te bezien of met dezelfde gegevens de ontwikkeling over een nog langere periode kan worden gereconstrueerd. Er worden dan geen dwarsdoorsnedes vergeleken, zoals bij Huijgen c.s. het geval is, maar leeftijds groepen, die in de verschillende gegevensbestanden uiteraard aanwezig zijn. Onder een aantal acceptabele aannames kunnen de onderzoekers zo terugkijken in de tijd en het beeld reconstrueren vanaf het cohort dat in 1911-1919 zijn intrede deed op de arbeidsmarkt. Behalve (1) deze verlengde periode claimt Wolbers dat zijn onderzoek op nog vier punten een stap vooruit brengt: (2) een controle op de invloed van sociaal milieu op de samenhang tussen opleidingsniveau en functieniveau; (3) een toets of de diploma-inflatie alleen komt door de toegenomen kloof of ook door verscherpte selectie-eisen van werkgevers; (4) een analyse of men alleen met zijn eigen intredcohort concurreert of tijdens de beroepsloopbaan ook nog met jongere, hoger opgeleide cohorten; en (5) de relatie tussen verdringing en werkloosheid. Met deze vijf verbeterstappen is meteen de structuur van het proefschrift weergegeven. Elke stap beslaat een hoofdstuk, waarvan de resultaten (soms met een co-auteur) al eerder in artikelvorm waren gepubliceerd.

De hoofdconclusie is dat over de hele periode bezien alle diploma's in absolute zin minder waard zijn geworden. De diploma-inflatie is vooral sterk op de lagere opleidingsniveaus, waar de opeenvolgende cohorten met hetzelfde papiertje een steeds lager functieniveau bereiken. In relatieve ('positionele') zin is de betekenis van opleiding uiteraard wel toegenomen, omdat een hoger opleidingsniveau steeds meer een vereiste is om de hogere functieniveaus te kunnen bereiken. Deze uitkomsten zullen de kenners van het werk van Huijgen niet verbazen. Gegeven diens methode en dezelfde data zitten de uitkomsten eigenlijk vanaf het begin in de kaarten. Je zou kunnen zeggen dat met het onderzoek van Wolbers de door Huijgen c.s. vermoede trend over een langere periode wordt bevestigd.

Voor de andere onderzoeksvragen zijn ook een aantal andere databestanden te hulp geroepen, zoals de Familie-enquête, het Nederlandse Stratificatie- en Mobiliteitsbestand en het OSA-arbeidsaanbodpanel. De resultaten zijn soms meer, soms minder eenduidig. Het is duidelijk dat de waargenomen diploma-inflatie vooral door de ongelijke ontwikkeling van opleidingsniveau en functieniveau wordt veroorzaakt en in mindere mate door veranderingen in de samenhang zelf. Het is ook duidelijk dat de diploma-inflatie zich vooral manifesteert bij arbeidsmarktintrede, maar of de relatieve opleidingspositie ook de kansen tijdens de beroepsloopbaan bepaalt is lastiger interpreteerbaar. Interessant is met name de analyse met betrekking tot de mogelijk versturende invloed van het sociale milieu van herkomst. Wolbers stelt terecht dat wanneer in het verleden de relatie tussen opleiding en beroep meer dan tegenwoordig werd gedragen door een milieu-effect, de mate van diploma-inflatie zou worden overschat. Hij vindt hiervoor ook in enige mate bevestiging, maar lang niet zoveel dat hiermee de gevonden diploma-inflatie zou kunnen worden wegverklaard. Toch is dit punt wel belangrijk, want het sociale milieu is bepaald niet de enige factor die het verband tussen opleiding en functieniveau kan beïnvloeden en in de tijd vertekenen. Minstens zo belangrijk lijken mij zaken als de persoonlijke begaafdheid of het IQ, waarover zo dadelijk nog een opmerking.

Wolbers toont zich in dit proefschrift een bekwaam onderzoeker. Geavanceerde technie-

ken als loglineaire analyse, cumulatieve logitanalyse en event-history analyse worden uit de kast gehaald teneinde de onderzoeksvragen zo goed mogelijk te beantwoorden en uit de gegevens het maximale rendement te halen. Met dit proefschrift heeft de Huijgen-benadering zo ongeveer zijn methodologische hoogtepunt bereikt, maar is, als ik het goed zie, ook aan het eind van zijn mogelijkheden gekomen. Een belangrijk probleem is dat de resultaten – of beter gezegd: de interpretaties van de resultaten – door een groot deel van het forum van onderwijs- en arbeidsmarktonderzoekers niet langer worden geaccepteerd. Er worden in hoofdzaak drie tegenwerpingen gemaakt tegen die interpretaties, die in oplopende mate van belangrijkheid luiden:

1. De functieniveaus zijn in de tijd niet vergelijkbaar vanwege belangrijke inhoudelijke veranderingen in beroepen. Omdat het functieniveau (gemeten op een schaal van 1 t/m 7) een grove maat is, worden wezenlijke veranderingen al gauw over het hoofd gezien.
2. De achterliggende factor van begaafdheid of IQ wordt genegeerd. Het onderwijs levert weliswaar een belangrijke prestatie door mensen van een gegeven begaafdheidsniveau naar een steeds hoger opleidingsniveau te tillen, maar daarmee staan de respectievelijke diploma's in de tijd bezien nog niet voor dezelfde mate van menselijk kapitaal (Wielers en Glebbeek 1995).
3. Het economische rendement van opleiding komt niet alleen tot uitdrukking in het bereiken van hogere functies, maar ook in het beter vervullen van lagere. Iemand met een 'te hoge' opleiding kan dan toch een meerwaarde hebben in de betreffende functie, hetgeen tot uitdrukking komt in een hogere beloning.

De betekenis van de laatste tegenwerping is dat de discussie over diploma-inflatie en verdringing alleen met inkomensgegevens kan worden beslecht. Met name economen doen daarom pogingen gegevensbronnen aan te boren waarin deze inkomensgegevens beschikbaar zijn. Dat een hogere opleiding op een lager functieniveau financieel rendeert, blijkt daarin een constant terugkerend resultaat. Groot en Maassen van den Brink, die onlangs een overzicht verzorgden van (inter-)nationaal onderzoek waarin met inkomensgegevens is gewerkt, komen op grond hiervan tot de krachtige stellingname dat "verhalen over 'diploma-inflatie' en verdringing naar het rijk der fabelen (dienen) te worden verwezen" (1998, p.91). Gevreesd moet worden dat zij het onderzoek van Wolbers dezelfde weg zouden willen laten gaan.

Maarten Wolbers kent deze tegenwerpingen natuurlijk allemaal wel, maar zat met het feit dat voor de lange-termijn reconstructie die hij wenste geen inkomenscijfers beschikbaar zijn. Vandaar dat hij de weg volgde om de economische opbrengst van opleidingen af te meten aan het bereikte functieniveau. Daarbij kiest hij ervoor achter zijn aanpak en methode te staan en niet weg te vluchten in slappe conclusies. Integendeel: het proefschrift mondt uit in de onverbloemde steun aan de verdringingshypothese en in feite (mijn gevolgtrekking) de impliciete stelling dat zonder enige schade voor de economie iedereen twee jaar minder opleiding zou kunnen volgen. Desalniettemin moet worden vastgesteld dat de hier gevolgde benadering niet voldoet om uit deze discussie te komen. Over wat ze kunnen verzinnen om er wel uit te komen, zullen onderwijs- en arbeidsmarktonderzoekers zich ongetwijfeld nog vele jaren het hoofd breken.

Literatuur

- Asselberghs, K., Batenburg, R., Huijgen, F. en Witte, M. de (1998). *De kwalitatieve structuur van de werkgelegenheid in Nederland, deel IV*. Den Haag: Organisatie voor Strategisch Arbeidsmarktonderzoek.
- Groot, W. & Maassen van den Brink, H. et al. (1998). *Overscholing en verdringing op de arbeidsmarkt*. Den Haag: Welboom.
- Wielers, R. & Glebbeek, A. (1995). Graduates and the labour market in the Netherlands: three hypotheses and some data. *European Journal of Education*, 30, 11-30.

Arie Glebbeek

Rijksuniversiteit Groningen, Vakgroep Sociologie/ICS
Grote Rozenstraat 31, 9712 TG Groningen.

Paulien Meijer

Teachers' practical knowledge. Teaching reading comprehension in secondary education (Dissertatie Rijksuniversiteit Leiden 1999)

De laatste tijd verschijnen steeds meer publicaties over onderzoek naar praktijkkennis van leraren en aanstaande leraren. Daarbij wordt voortgebouwd op een relatief jonge onderzoekstraditie waarbij cognities van (aanstaande) leraren object van onderzoek zijn. Veel onderzoek op dit gebied wordt in Nederland uitgevoerd binnen onderzoeksprogramma's van universitaire lerarenopleidingen. Dit is niet verwonderlijk als we in ogenschouw nemen dat bij het opleiden van leraren het belang van de praktijkkennis van de leraren en de praktijkkennis van de leraar-in-opleiding steeds groter is. Binnen de opleiding is het immers zaak dat over achterliggende principes en theorieën van dat wat in de klas gebeurt en dat wat de leraar in de klas doet gesproken wordt, met name om te voorkomen dat alleen het handelen centraal komt te staan in de opleiding van leraren. Daarnaast komt in de opleiding het accent meer en meer te liggen op het ontwikkelen van de eigen praktijkkennis door de aanstaande leraar. Het kunnen achterhalen en expliciteren van deze eigen praktijkkennis is daarbij een belangrijk aspect.

In deze ontwikkeling past het onderzoek van Paulien Meijer. Zij onderzocht de praktijkkennis van leraren over het onderwijzen van tekstbegrip in het voortgezet onderwijs. Twee doelen had ze daarbij voor ogen. Ten eerste het in kaart brengen van de inhoud van de genoemde praktijkkennis van leraren in de bovenbouw van het VWO met het oog het kunnen vaststellen van een eventueel gezamenlijke, gedeelde, praktijkkennis. Het tweede doel was het ontwikkelen en uitproberen van instrumenten om zicht te krijgen op de praktijkkennis van docenten.

Gezien deze twee doelen is het onderzoek van Paulien Meijer zeker relevant. Er is nog weinig bekend over de inhoud van praktijkkennis van leraren. In veel nascholings- en schoolontwikkelingsprogramma's wordt echter wel uitgegaan van kennis waarvan men veronderstelt dat die bij leraren aanwezig is. Ook voor de opleiding van leraren is het voor aanstaande leraren relevant te weten wat de inhoud van de praktijkkennis van leraren is.

Ook het doel om instrumenten te ontwikkelen is van belang. In ander onderzoek van de laatste tijd waar cognities van (aanstaande) leraren onderzocht werden (Dolk, 1997 en Buitink, 1998) wordt ook gewezen op het belang om passende instrumenten te hebben om cognities in de ruimste zin van het woord in kaart te brengen. Deze instrumenten worden dan idealiter in combinatie met elkaar gebruikt.

Opzet van het onderzoek

Zoals gezegd wordt in het onderzoek de praktijkkennis van docenten onderzocht. De term praktijkkennis verwijst naar de kennis van docenten. Deze kennis wordt onderscheiden van kennis voor docenten. Praktijkkennis voor docenten wordt in het onderzoek aangeduid met de term 'formele kennis'. Praktijkkennis van docenten wordt gedefinieerd als het geheel van kennis, opvattingen en interactieve cognities van docenten. Interactieve cognities zijn in deze context die zaken waar een docent aan denkt of op let tijdens het lesgeven.

Het onderzoek bestaat uit twee delen. Het eerste deel is een kwalitatief onderzoek. In dit deel werden 13 leerkrachten onderzocht. Met drie instrumenten werd de praktijkkennis van deze leerkrachten beschreven. Deze instrumenten vulden elkaar aan, met elk instrument werd een ander deel van de praktijkkennis in kaart gebracht (methodentriangulatie). Met behulp van het eerste instrument, een semi-gestructureerd interview, werden allerlei aspecten van het onderwijzen van tekstbegrip in de bovenbouw van het vwo beschreven. Het tweede instrument was het gebruik van concept mapping. Docenten werden gevraagd de begrippen rond het onderwerp "tekstbegrip onderwijs in de bovenbouw van het vwo" te ordenen in een schema en dit schema vervolgens toe te lichten. Met het derde instrument, een "stimulated recall" interview, werden van de interactieve cognities bestudeerd. Dit interview werd gehouden aan de hand van een video-opname van een les tekstbegrip.

In het tweede deel van het onderzoek werd met behulp van een vragenlijst de kennis en opvattingen van een grotere groep (69) leerkrachten bestudeerd. Bij het ontwikkelen van deze vragenlijst werd gebruik gemaakt van de uitkomsten van het eerste deel van het onderzoek.

Belangrijkste resultaten

Uit de analyses van de gegevens van het eerste deel van het onderzoek kwam naar voren dat de praktijkkennis van leerkrachten beschreven kan worden met behulp van tien categorieën. Door combinatie van de categorieën werden twee groepen patronen onderscheiden. De eerste groep patronen betrof de wijze waarop docenten leerlingen benaderen, bijvoorbeeld individueel of als groep. De tweede groep patronen had betrekking op de inhoud van de les. Door de patronen te combineren werden drie typen praktijkkennis onderscheiden te weten praktijkkennis (I) gericht op het vak, (II) gericht op de leerlingen en (III) gericht op leerprocessen en begripsvorming van leerlingen. Wat betreft de inhoud van de typen praktijkkennis bleken grote verschillen te bestaan tussen de leerkrachten. Er was in die zin niet echt sprake van een gezamenlijke praktijktheorie van leraren in het vwo.

Ook de resultaten van het tweede deel van het onderzoek wijzen erop dat leerkrachten weinig overeenkomsten vertonen in hun praktijkkennis. Wel bleek het mogelijk de verschillen tussen de leerkrachten te ordenen. Er werden vier groepen leerkrachten onderscheiden te weten (I) leerkrachten met een visie op tekstbegrip als bestaande uit grote elementen, (II) docenten met een visie op tekstbegrip als bestaande uit kleine, aparte, onderdelen, (III) docenten die tekstbegrip zien als het relateren van leerlingen en teksten en (IV) docenten met een lage waardering voor tekstbegriponderwijs.

Gezien de resultaten van de twee deelonderzoeken werd met betrekking tot het eerste onderzoeksdoel geconcludeerd dat de praktijkkennis van docenten met betrekking tot het onderwijzen van tekstbegrip in de bovenbouw van het vwo overheerst werd door variatie. Er waren nauwelijks onderdelen waar alle docenten het over eens waren. Uit het feit echter dat er in het eerste deel van het onderzoek relatief uniforme typen van praktijkkennis werden gevonden en in het tweede deel van het onderzoek relatief uniforme groepen docent werd geconcludeerd dat praktijkkennis van docenten slechts voor een deel uniek is.

Wat betreft het tweede onderzoeksdoel wordt geconcludeerd dat de instrumenten die gebruikt zijn in het kwalitatieve onderzoek geschikt zijn om de praktijkkennis in kaart te brengen. De combinatie van instrumenten is van belang, waarbij in het onderzoek de concept maps alleen in combinatie met het semi-gestructureerd interview begrepen konden worden. De gestructureerde vragenlijst die bij het tweede deel van het onderzoek gebruikt werd, is een problematischer instrument. Er blijkt een spanningsveld tussen het ontwikkelen van 'n vragenlijst vanuit de theorie enerzijds en vanuit de kwalitatieve data anderzijds. Dat maakt het gebruik van de gegevens die met de vragenlijst verkregen worden beperkt. Wel kan de vragenlijst gezien worden als een instrument dat meer gedetailleerde informatie oplevert over de resultaten van het kwalitatieve onderzoek.

Commentaar

Paulien Meijer onderzoekt de praktijkkennis van docenten. Ze beperkt zich hierbij tot een cognitieve invulling van dit begrip. Praktijkkennis omvat kennis, 'beliefs' en interactieve cognities. Zaken als beelden, 'images', 'gestalts' etc. laat ze buiten beschouwing (zie bijvoorbeeld het proefschrift van Maarten Dolk, 1997). Het is steeds weer de vraag wat we beschrijven als we proberen 'cognities' in kaart te brengen. Er zijn andere instrumenten die weer andere aspecten van praktijkkennis 'boven water halen'. Kortom, het begrip praktijkkennis, hoe we het ook definiëren, blijft een lastig begrip, zowel wat betreft een eenduidig concept als wat betreft de wijze waarop we het kunnen onderzoeken. Paulien Meijer omzeilt dit probleem, hoewel ze het wel via een omweg aanroert, onder andere door te kiezen voor methodentriangulatie waarbij instrumenten elkaar aanvullen om de praktijkkennis zo volledig mogelijk te kunnen beschrijven.

Ook een ander belangrijk aspect van het onderzoek wat ze heeft uitgevoerd benadert ze indirect, en wel de in het onderzoek te nemen maatregelen om de kwaliteit te waarborgen. Zeker in kwalitatief onderzoek is het belangrijk aan te geven welke maatregelen een onderzoeker genomen heeft in het kader van de kwaliteit van de data en de interpretatie ervan (zie bijvoorbeeld Smaling, 1996). Pas in haar slotopmerkingen gaat ze expliciet in op het aspect van wat zij 'objectiviteit' noemt. In het hoofdstuk over onderzoeksvragen, instrumenten en procedures had ze expliciet op dit aspect in kunnen gaan. Er lijkt hier sprake van een belangrijk omissie. Aansluitend kan opgemerkt worden dat de navolgbaarheid van een aantal van haar redeneringen en conclusies, met name in hoofdstuk 3, voor de lezer niet altijd even groot is.

Haar conclusie dat de uitkomsten van het onderzoek belangrijk is voor onder andere het opleiden van leraren is een juiste conclusie. Ze kan echter verder gaan bij die conclusie. Het is in de lerarenopleiding niet alleen van belang dat een aanstaande leraar de praktijkkennis van leerkrachten onderzoekt, maar ook de eigen praktijkkennis (zie bijvoorbeeld Buitink, 1998). En de instrumenten die Paulien Meijer heeft ontwikkeld kunnen dan ook in dat licht gezien worden. Zijn de in het onderzoek gehanteerde instrumenten bruikbaar voor aanstaande leraren om de eigen praktijkkennis te onderzoeken en zo nee, welke aangepaste of andere instrumenten zijn daarvoor wel geschikt.

Tenslotte zij opgemerkt dat ze in het laatste hoofdstuk van haar dissertatie een zeer relevant punt aanroert, namelijk de vraag of er sprake is van goede en minder goede praktijkkennis. Het antwoord dat ze daarop geeft is zeer plausibel en kan belangrijker worden naarmate er meer sprake is van assessment van (aanstaande) leraren waarbij de praktijkkennis onderdeel is van die assessment.

Literatuur

- Buitink, J. (1998). *In-functie-opleiden en in-functie-leren van aanstaande leraren*. Groningen: UCLO. Dissertatie
- Dolk, M (1997). *Onmiddellijk onderwijsgedrag: over denken en handelen van leraren in onmiddellijke onderwijsituaties*. Utrecht: IVLOS. Dissertatie
- Smaling, A. (1996). De methodologische kwaliteit van kwalitatief onderzoek. *Mogelijke maatregelen. Verpleegkunde*, 11 (4), 240-254

Jaap Buitink, UCLO
Rijksuniversiteit Groningen, Oude Boteringestraat 48-50
9712 GL Groningen, e-mail: j.buitink@uclou.rug.nl

Peter Bimmel.

Training en transfer van leesstrategieën: training in de moedertaal en transfer naar een vreemde taal, een effectstudie bij leerlingen uit het voortgezet onderwijs.

Dissertatie Universiteit Utrecht, 's Hertogenbosch: Malmberg, 1999. ISBN: 90 208 7818 2.

Samenvatting van de inhoud

In zijn proefschrift doet Bimmel verslag van twee studies naar het effect van onderwijs in leesstrategieën bij leerlingen uit het derde jaar van het voortgezet onderwijs. In de inleiding (hoofdstuk 1) beschrijft Bimmel binnen welk kader het onderzoek geplaatst dient te worden en wordt de vraagstelling beschreven. De vraag is in hoeverre onderwijs in leesvaardigheid oftewel het gebruik van leesstrategieën bij het van Nederlands van invloed is op de wijze waarop leerlingen leestaken bij het vak Engels aanpakken. Uit zijn literatuurstudie concludeert Bimmel dat het in principe mogelijk moet zijn transfer te bewerkstelligen, maar dat een zekere taalvaardigheid een voorwaarde is. Daarnaast zijn er een aantal transferbevorderende maatregelen nodig, zoals: laten zien van nut van strategiegebruik in verschillende contexten, gebruik van contextrijke en betekenisvolle leertaken en er moet gezorgd worden voor decontextualisatie en recontextualisatie.

Hoofdstuk 2 en 3 geven het theoretisch kader. De aspecten die van belang zijn voor het ontwikkelen van onderwijsprogramma's op het gebied van leesvaardigheid worden beschreven. Een goed programma moet niet alleen aandacht besteden aan het ontwikkelen van de cognitieve maar ook de metacognitieve competentie. In zijn programma's heeft Bimmel getracht deze aspecten van leesvaardigheid samen te brengen. Tijdens de training in de eerste studie kregen de leerlingen vier leesstrategieën aangeboden, namelijk: 1) koppen snellen (skimmen), 2) begin en einde alinea lezen om tekstinhoud te voorspellen, 3) sleutelfragmenten onderstrepen (belangrijke informatie selecteren) en 4) schanierwoorden vinden (structuur van tekst achterhalen). Dit bleek achteraf een goede keuze. Daarom werden in de tweede studie de eerste twee strategieën vervangen door 1) vragen stellen bij een tekst en 2) maken van een semantische kaart (schema van de tekst).

In zijn studies maakte Bimmel gebruik van de uitleg- en de bewustmakingsmethode. De uitlegmethode kent drie fasen. In de oriëntatiefase nemen leerlingen onder meer kennis van een strategie en lezen informatie over de toepassing ervan. In de uitvoeringsfase wordt de strategie toegepast en wordt gecontroleerd op een correcte uitvoering. In de evaluatiefase wordt tijdens een klassikale bespreking ingegaan op onder meer belangrijke informatie over de strategie, het doel ervan en de activiteiten die zijn uitgevoerd. De bewustmakingsmethode kent dezelfde drie fasen. Alleen wordt in de tweede en derde fase in tweetallen gereflecteerd op de wijze waarop de taak is uitgevoerd.

In studie 1 (*hoofdstuk 4*) is nagegaan in hoeverre twee experimentele programma's effect hebben op de beheersing van de strategieën, de algemene leesvaardigheid Nederlands en Engels, de kennis die leerlingen hebben omtrent lezen en leesstrategieën en er is gekeken naar samenhangen van de effecten met leesattitude en leerstijl.

Twee experimentele groepen (beide $N=36$) zijn vergeleken met een controlegroep ($N=72$). Daarbij is gebruik gemaakt van een pretest-posttest-control-group-design. De ene experimentele groep kreeg het uitlegprogramma en de andere het bewustmakingsprogramma. De lessen werden, op zeven zaterdagochtenden, gegeven door ervaren docenten in zes groepen van twaalf leerlingen. De controlegroep volgde het reguliere onderwijs.

De opzet van het experiment was als volgt: 1) voormeting: beheersing vier leesstrategieën, leesvaardigheid Nederlands en Engels, woordenschat Engels en vragenlijst: kennis van lezen, 2) trainingssessie 1 t/m 5: training vier strategieën bij Nederlands, 3) eerste nameting: beheersing leesstrategieën, leesvaardigheid Nederlands en Engels, vragenlijst: kennis van lezen, leerstijlvragenlijst, 4) trainingssessie 6 en 7: training vier strategieën bij Engels, 5) tweede nameting: leesvaardigheid Engels en leesattitude vragenlijst.

Alvorens de resultaten te bespreken gaat Bimmel in op de betrouwbaarheid van de verschillende meetinstrumenten. Deze bleek van de verschillende bestaande toetsen in orde. Van de zelf geconstrueerde strategietoetsen liet de betrouwbaarheid vaak te wensen over en van de attitudevragenlijst was hij prima. Door middel van co-variantie-analyses zijn vervolgens de effecten van de beide programma's bepaald. De resultaten waren teleurstellend. Er werd alleen een effect gevonden op de 'vragenlijst: kennis van lezen'. De beide experimentele groepen hadden meer kennis opgedaan tijdens de trainingsperiode dan de controlegroep. Bimmel formuleert een tiental interfererende factoren die het effect van de beide programma's waarschijnlijk hebben beïnvloed. De belangrijkste zijn: motivatie van de leerlingen, selectie van de strategieën, onduidelijkheid van de reflectie-opdrachten, rol van de docent en de lage betrouwbaarheid van de strategietoetsen.

In studie 2 (*hoofdstuk 5*) is één onderwijsprogramma (gebaseerd op de bewustmakingsmethode) op z'n effect getoetst. In dit programma zijn de tien interfererende factoren verwerkt. De motivatie van de leerlingen werd aangewakkerd door de testresultaten mee te laten tellen voor het rapport en de 17 lessen te geven onder schooltijd. Zoals eerder werd vermeld is in studie 2 wel aandacht besteed aan vragen stellen en het maken van een semantische kaart. Tevens werd de rol van de docent meer benadrukt, door aan te sturen op meer individuele interactie.

Aan studie 2 namen 12 experimentele en 119 controleleerlingen deel. De opzet van dit experiment was als volgt: 1) voormeting: beheersing leesstrategieën, leesvaardigheid Nederlands en Engels, woordenschat Nederlands en Engels, 2) trainingssessie 1 t/m 15: training vier strategieën bij Nederlands, 3) eerste nameting: beheersing leesstrategieën, leesvaardigheid Nederlands en Engels, 4) trainingssessie 16 en 17: training vier strategieën bij Engels, 5) tweede nameting: leesvaardigheid Engels.

De betrouwbaarheidscoëfficiënten van de nieuw ontwikkelde strategietoetsen 'vragen stellen' bleken bedroevend laag te zijn. De alpha's van de 'semantische kaart toetsen' waren aanzienlijk beter. De resultaten van deze studie waren bemoedigender dan die van studie 1. Het bleek dat vergeleken met de controlegroep de leerlingen uit de experimentele groep na de training beter waren geworden in het uitvoeren van drie strategieën: sleutelfragmenten onderstrepen, vragen stellen en een semantische kaart maken. Ook op de toets 'leesvaardigheid Nederlands' werd een significant effect gevonden. Daarbij dient wel vermeld dat de toets was aangepast. Leerlingen moesten eerst de sleutelfragmenten aangeven en pas daarna kregen ze de eigenlijke vragen bij de tekst. Op test 'leesvaardigheid Engels' werd geen effect gevonden. Dus er vond geen transfer plaats.

In *hoofdstuk 6* geeft Bimmel onder meer de implicaties voor de theorie over leesvaardigheidsonderwijs en de onderwijspraktijk aan en geeft hij suggesties voor verder onderzoek.

Commentaar

Het onderzoek geeft nogmaals aan dat het mogelijk is om bij leerlingen (meta)cognitieve strategieën te ontwikkelen. De onderzoeken die op dit gebied de laatste jaren zijn verricht, wijzen dit ook al uit. Al hebben die onderzoeken inderdaad zoals Bimmel aangeeft vaak betrekking op jongere kinderen en zwakke lezers (Paris, Cross & Lipson, 1984; Brown & Palincsar, 1989; Brand-Gruwel, 1995). Echter zijn doelgroep, leerlingen voortgezet onderwijs, is maar zelden onderzocht op strategiegebruik tijdens lezen en ook het nagaan of er mogelijk een transfereffect kan optreden naar strategiegebruik bij het lezen van teksten in een andere taal is een vernieuwende stap. Toch is het jammer dat dit proefschrift waarin op zich interessante vragen worden gesteld, niet altijd de gewenste resultaten oplevert.

Peter Bimmel heeft waarschijnlijk na het analyseren van de data van de eerste studie de nodige slapeloze nachten gehad. De resultaten waren ronduit bedroevend. Er werd zelfs geen effect gevonden op de strategietoetsen, terwijl dat toch de toetsen waren om na te gaan of de doelen van de programma's - het ontwikkelen van de strategieën bij de leerlingen - waren bereikt. Ik hoor een Amsterdamse collega nog zeggen: 'als je er maar tijd aan besteedt, het maakt niet uit hoe, dan heb je op dit soort toetsen altijd wel effect'. Nee dus, je moet wel degelijk goed onderwijs geven en voldoen aan bepaalde randvoorwaarden. Ook Bimmel kwam tot deze conclusie. De vraag is echter of hij misschien had kunnen voorkomen dat studie 1 zo is uitpakkt. Mijns inziens had dit gekund. In de literatuur komen toch wel belangrijke zaken naar voren die Bimmel misschien beter had moeten implementeren in zijn programma's.

Zo kun je je afvragen of de keuze van strategieën wel de juiste is geweest. Bimmel kiest twee strategieën die betrekking hebben op het voorspellen van de tekstinhoud. Uit onderzoek van onder meer Brown en Palincsar (1989) bleek al dat voorspellen wel bijdraagt tot een beter tekstbegrip, maar dat dit voor tekstbegrip niet het belangrijkste is. Bimmel geeft ook zelf op pagina 16 van zijn proefschrift aan dat het maken van een semantische kaart leidt tot een diepere verwerking van de tekst dan sleutelfragmenten onderstrepen. Toch kiest Bimmel hier niet voor. Ook de toepassen van de strategie 'vragen stellen' wordt als een van de karakteristieke kenmerken van een goede lezer gezien (Pearson & Fielding, 1991): Ook deze strategie wordt in de eerste studie niet onderzocht.

Verder blijkt uit literatuur van bijvoorbeeld Simons en Verschaffel (1992) dat het bewerkstelligen van transfer een moeizame kwestie is. Bimmel concludeert op pagina 142 dat er helaas geen spontane transfer optreedt naar leesvaardigheid Engels. Dit mag ook niet verwacht worden. Gedegen transfertraining is nodig, waaronder mijn inziens twee extra Engelse

lessen niet vallen. Misschien had de training in het lezen van Engelse en Nederlandse teksten meer geïntegreerd plaats moeten vinden, zodat leerlingen ervaren hoe dicht de taken en taakeisen bij elkaar liggen. Pas dan kun je misschien enig transfer verwachten. Jammer is ook dat in studie 2 de leerlingen bij de Engelse leesvaardigheidstest niet zoals bij de leesvaardigheidstest Nederlands eerst sleutelfragmenten moesten onderstrepen alvorens de vragen te maken. Misschien was dan toch eerder een transfereffect bereikt. Far-transfer wordt dan iets meer near-transfer.

Tot slot geeft Bimmel op pagina 56 van zijn proefschrift aan dat het waarschijnlijk het meest verstandig is een zo compleet mogelijk trainingsprogramma te kiezen dat aan zoveel mogelijk criteria en voorwaarden voldoet. Toch besluit hij om in studie 1 twee programma's te maken. Was de keuze voor één zeer compleet programma, waarin meer aandacht voor transfer, ontwikkelen van metacognitieve vaardigheden en misschien meer rolwisselend onderwijs, waarvan hij toch ook het belang aangeeft, niet een betere geweest?

Het programma van studie 2 heeft een beter effect dan de twee programma's van studie 1. De onderwijsdoelen worden deels bereikt. De leerlingen uit de experimentele groep zijn in vergelijking met de leerlingen uit de controlegroep na de training in de uitvoering van drie van de vier strategieën beter. Alleen op 'schanierwoorden' werd geen effect gevonden. Wel dient vermeld te worden dat de betrouwbaarheden van de toetsen 'vragen stellen' heel erg laag waren en er eigenlijk geen conclusie daaromtrent getrokken had mogen worden. Verder is het jammer dat een aantal meetinstrumenten zoals de 'vragenlijst: kennis van lezen' niet meer is meegenomen. Dit zou het effect van het programma en de validatie van de instrumenten hebben versterkt. Het niet afnemen van deze instrumenten brengt met zich mee dat in studie 2 niet meer is gekeken naar het effect van het programma op de metacognitieve kennis en vaardigheden.

Ondanks deze kritische kanttekeningen toch een woord van bewondering voor de wijze waarop het onderzoek is uitgevoerd. Uit ervaring weet ik dat onderzoek doen binnen het voortgezet onderwijs een moeizame kwestie is. De meeste docenten zijn niet of nauwelijks bereid om uren af te staan waarin bijvoorbeeld een experimenteel programma kan draaien en laat staan om veel toetsen af te nemen. Daarnaast is de kwestie van de motivatie van de leerlingen een hele belangrijke. In studie 2 is Bimmel erin geslaagd het onderzoek zo in het reguliere onderwijs in te bedden dat voor leerlingen de motivatie (al was die waarschijnlijk extrinsiek) groter was dan tijdens de eerste studie.

Saskia Brand-Gruwel
Open Universiteit Nederland / OTEC
Postbus 2960, 6401 DL Heerlen

Literatuur

- Brand-Gruwel, S. (1995). *Onderwijs in tekstbegrip*. Proefschrift: Kath. Universiteit Nijmegen. Ubbergen: Tandem Felix.
- Brown, A.L., & Palincsar, 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: Erlbaum.
- Paris, S.G., Cross, D.R., & Lipson, M.Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology*, 76, 1239-1252.
- Pearson, P.D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M.L. Kamil, P.B. Mosenthal, & P.D. Pearson (Eds.), *Handbook of reading research* (vol. 2, pp. 815-860). New York: Longman.
- Simons, P.R.J., & Verschaffel, L. (1992). Transfer: Onderzoek en onderwijs. *Tijdschrift voor Onderwijs Research*, 17, 3-6.

Verslag Algemene Ledenvergadering VOR d.d. 21 mei 1999

(12.30-13.30 uur; Katholieke Universiteit Nijmegen - tijdens ORD 1999)

Aanwezig:

Leden:

P. Berendsen, Th. Bergen, H. Biemans, G. ten Dam, J. van Driel, J. Gordijn, F. de Jong, G. Kanselaar, P. Kirschner, E. Koerts, R. Poell, F. Prins, E. Rutgers, C. Terlouw, J. Terwel

Bestuursleden:

J. van den Akker, A.B. Dijkstra, W. Jochems, W. Kuiper (penningmeester), M. Mulder (secretaris & notulist), J. Pieters, N. Verloop (voorzitter)

Bericht van verhindering:

F. Dochy, W. van Esch, D. Mansvelder-Longayroux, M. Peters-Sips, F. Riemersma, A. Swennen, J. Thijssen, C. van Vilsteren, F. van Vught

1. Opening en agendavaststelling

De voorzitter heet de aanwezigen hartelijk welkom, opent de vergadering en stelt de agenda vast.

2. Verslag Algemene Ledenvergadering 13 mei 1998

Het verslag van de algemene ledenvergadering van 13 mei 1998 wordt ongewijzigd goedgekeurd.

3. Mededelingen

De voorzitter deelt mede dat Dr. F. van Bussel is overleden, hetgeen een tragisch verlies is. Van Bussel is actief geweest binnen het bestuur als penningmeester van de VOR. De voorzitter heeft een In Memoriam geschreven in het TOR. Van Bussel heeft destijds de financiële administratie van de VOR uit het slop gehaald en daarmee veel werk verzet voor de vereniging.

De ledenlijsten van de VOR en de NVO zijn vergeleken en gebleken is dat er slechts enkele dubbelleden zijn. De NVO-leden wordt een dubbellidmaatschap van de VOR aangeboden tegen het gereduceerde tarief dat ook geldt voor VELON- en NVVO-leden.

De Divisie Curriculum zal in het jaar 2000 vanuit de VOR een symposium verzorgen op de AERA.

SUNCOO heeft verzocht na te gaan of een deel van de leden aansluiting kan krijgen bij de VOR. Gedacht wordt aan een nieuwe divisie op het gebied van ICT. De secretaris zal dit opmen en inbrengen in het Bestuur.

4. Jaarverslag 1998-1999

Het Jaarverslag 1998/1999 wordt goedgekeurd.

Naar aanleiding van het Jaarverslag merkt R. Poell op dat hij punt 9 van het Jaarverslag ondersteunt, hetgeen inhoudt dat de ORD-secretaris van het voorgaande jaar wordt toegevoegd aan de volgende ORD-organisatiecommissie. De voorzitter deelt mede dat het de bedoeling is steeds de voorgaande, huidige en volgende ORD-secretaris deel te laten uitmaken van de ORD-organisatiecommissie. Zo wordt de ervaringskennis beter overgedragen.

5. Financiën

5.1 Verslag kascontrolecommissie (Thijssen – Van Vilsteren)

De heren Thijssen en Van Vilsteren konden niet aanwezig zijn tijdens de vergadering, maar er is een schriftelijke reactie van beide heren. Zij hebben op enkele punten gewezen die de penningmeester mee zal nemen voor het volgend jaar. Het voorstel is de penningmeester te dechargeren, hetgeen wordt besloten.

5.2 Verlies- en Winstrekening VOR 1998

De penningmeester licht de Verlies- en Winstrekening VOR 1998 toe. Hij stelt daarbij dat de lasten hoger zijn dan de baten. Het resultaat is een verlies van een kleine 22.000 gulden. Er is echter een niet terechte winst over 1997, over dat jaar is een kleine 40.000 gulden positief resultaat geboekt. Het negatieve resultaat over 1998 wordt veroorzaakt door de extra lasten met betrekking tot met name TOR-4 uit 1997 (f 10.425,99) dat is geboekt ten laste van 1998, en enkele andere tijdschriften waar leden abonnementen op hebben (in totaal f 7.118,96). Het resultaat over beide jaren is 17.000 gulden positief. Er moeten nog twee nummers van het TOR over 1998 worden betaald. De rekeningen daarvan waren nog niet binnen (en de nummers nog niet verschenen). Deze lasten zijn wel meegenomen in de Verlies- en Winstrekening over 1998.

De voorzitter stelt voor de Verlies- en Winstrekening VOR 1998 goed te keuren, hetgeen de vergadering besluit.

Balans 1998 VOR

Ultimo 1998 was er een saldo van f 63.864,75. Ultimo 1997 was dat f 68.936,45. Er is per Divisie per lid 5 gulden toegevoegd aan de Divisiesaldi. Dat zijn geormerkte gelden voor de Divisies.

De VOR-conferentie over het Studiehuis was niet begroot en de VOR heeft een rekening van de organisatoren ontvangen van circa 12.000 gulden. Dit punt is besproken binnen de Faculteit Toegepaste Onderwijskunde (TO) en voorgesteld is dat TO 9.000 gulden voor haar rekening zal nemen, en dat de VOR het resterende bedrag betaalt. Dit voorstel wordt overgenomen door de ledenvergadering en goedgekeurd.

De voorzitter stelt voor de Balans 1998 VOR goed te keuren, hetgeen wordt besloten.

Begroting 2000 VOR

De penningmeester licht de Begroting 2000 toe, en merkt op dat er een positief resultaat van f 10.000 verwacht wordt. Voorgesteld wordt de contributie met 15 gulden per jaar te verhogen (exclusief aio's/oio's en studenten), met name om de kosten van het VOR-Jaarboek te dekken. De voorzitter stelt voor de Begroting 2000 VOR en de verhoging van de contributie goed te keuren, hetgeen wordt besloten.

6. Uitgeefbeleid TOR

De voorzitter deelt mede dat het TOR met een tekort aan kopij kampt. Hij somt daarvan de redenen op en wijst onder andere op de druk om internationaal te publiceren. De TOR-redactie heeft een notitie opgesteld waarin wordt voorgesteld het TOR in het Engels te gaan uitgeven. Het zou dan volgens de redactie aantrekkelijker worden om in het TOR te gaan publiceren. De redactie zou dan moeten worden uitgebreid met bekende buitenlanders.

Dit voorstel is besproken in het bestuur en daar zijn fors wat bedenkingen geuit. Er zijn diverse alternatieven gesuggereerd die overwogen moeten worden.

Het bestuur vindt het onderwerp te belangrijk om een besluit met betrekking tot het uitgeefbeleid voor te stellen. Het bestuur stelt daarentegen voor de leden over dit onderwerp te polsen aan de hand van een enquête.

Uit de vergadering komen enkele opmerkingen naar voren die kort worden besproken:

- Rekent het ICO een publicatie in een Engelstalig TOR als een internationale publicatie?
- Is het nodig om met een Engelstalig TOR een andere uitgever te kiezen?
- Wat gebeurt er als er te weinig (minder dan 40%) leden reageren op de enquête?
- Is overleg met ISED (t.b.v. erkenning van TOR) zinnig om de kopijstroom te verhogen?
- Kan een elektronisch TOR het probleem helpen oplossen?
- De vergadering gaat akkoord met het houden van een enquête over de vraag of het TOR in het Engels zou moeten worden uitgegeven.

7. VOR Website

De VOR Website is in de lucht. Er was wat aansporing nodig bij de Divisies om kopij van de Divisies binnen te krijgen. W. Jochems doet verslag van de ontwikkelingen. De voorzitter bedankt Jochems en stelt dat de Website verder beheerst moet worden opgebouwd.

8. VOR Jaarboek 1999

Het VOR Jaarboek 1999 is verschenen, met zeer veel mutaties in het ledenregister. De leden hebben dit jaar trefwoorden van hun expertise opgegeven. Deze zullen worden bekeken en gegroepeerd. Bij de volgende ronde wordt het register aangeboden aan de leden die dan trefwoorden kunnen aankruisen. Dat kan de zoekfunctie van het Jaarboek verder vergroten. De instellingen worden nog zeer verschillend gepresenteerd. Volgend jaar zal opnieuw een format worden voorgelegd aan de instellingen met het verzoek dat format te volgen. Het uitgeven van het Jaarboek moet het volgend jaar worden afgestemd met de Webmaster. Op termijn kan het VOR Jaarboek wellicht op de Website worden geplaatst.

9. Bestuursverkiezing

De voorzitter leidt dit punt in en J. Pieters verlaat de vergadering. Het bestuur draagt als voorzitter van de VOR ter benoeming voor: Prof. dr.J.M. Pieters, Universiteit Twente, voor de periode 1999-2002; Pieters houdt het voorzitterschap van de Divisie Leren & Instructie tijdelijk aan tot een opvolger is gevonden (voorzien in najaar 1999). Er hebben zich voor de vergadering geen tegenkandidaten gemeld, er zijn geen bezwaren tegen deze benoeming en de vergadering besluit Pieters te benoemen. Pieters keert terug in de vergadering en aanvaardt het voorzitterschap zoals voorgesteld en bedankt het bestuur voor het in hem gestelde vertrouwen.

Als secretaris wordt benoemd: Dr. H. Biemans, LSOP Politie Onderwijs- en Kenniscentrum / Wageningen Universiteit, voor de periode 1999 – 2002. En als lid van het bestuur: Prof. dr. M. Mulder, namens de Divisie BBV, Wageningen Universiteit/Universiteit Twente, voor de periode 1999 – 2002. Aio-lid van het bestuur is Drs. F.J. Prins (Universiteit Leiden).

10. Rondvraag

J. Pieters spreekt een woord van dank uit aan de scheidend voorzitter. N. Verloop bedankt het bestuur en de vereniging voor de plezierige periode van zijn voorzitterschap.

Sluiting

27e Onderwijs Research Dagen / ORD 2000

Extra thema: 'Leren en werken'

Plaats: Leiden

Datum: woensdag 24 mei t/m vrijdag 26 mei 2000

Informatie te verkrijgen bij de congressecretaris:

J. van Driel

ICLON, Universiteit Leiden

Postbus 9555

2300 RB Leiden

tel.: +31 (0)71 527 40 15

fax: +31 (0)71 527 53 42

e-mail: driel@iclon.leidenuniv.nl

website: www.fsw.leidenuniv.nl/www/w3_iclon/level_1/congressen/

Ouderlijk gezin en schoolsucces

Een verklaring met demografische, culturele en sociale aspecten

G. Kraaykamp

Vakgroep Sociologie van de Katholieke Universiteit Nijmegen.

Correspondentie naar Dr. G. Kraaykamp, Vakgroep Sociologie/ICS, KUN, Postbus 9104, 6500 HE, Nijmegen, e-mail: G.Kraaykamp@maw.kun.nl.

ABSTRACT

In this article, we analyse the Dutch VOCL-1989 cohort data on 19.524 students in secondary school, to explain differences in educational attainment. Our explanation, predominantly, deals with characteristics of the parental home. It is concluded that the socio-economic position of parents explains most of variance in school success. Next, of the demographic features growing up in a single parent family mother's age and birth order seems to be particularly important. Students raised by culturally active parents indeed reach for somewhat higher levels of education, and positive parent-child interaction on school related issues enhances educational attainment. Almost no differences are found between female and male students for the effects of these family characteristics.

1. ACHTERGROND EN VRAAGSTELLING

Zoals in vrijwel alle Westerse landen zijn in Nederland diploma's verkregen in het onderwijs zeer belangrijk bij het realiseren van levenskansen. Een goede opleiding is vooral bepalend als het gaat om het bereiken van een hoge maatschappelijke positie; personen die succesvol zijn in het onderwijs beschikken namelijk over relatief veel hulpbronnen en zijn als gevolg daarvan ook capabel aantrekkelijke maatschappelijke posities te bezetten (Coleman 1988; Rupp & Wesselingh 1989). In de huidige tijd waar door diploma-inflatie het relatieve verschil tussen hoge en lage diploma's alleen maar toeneemt (Wolbers, 1998), lijkt daarom voor individuen het behalen van een hoog opleidingsniveau nog belangrijker dan enige decennia geleden.

In deze situatie is de vraag relevant, welke factoren het behalen van een hoog opleidingsniveau bepalen? Sociaal-wetenschappelijk onderzoek oriënteerde zich bij het beantwoorden van deze vraag vaak op de sociaal-economische kenmerken van ouders, met als idee dat maatschappelijk succesvolle ouders er beter in slagen het schoolsucces van hun kinderen te beïnvloeden (Tesser 1986; Vrooman & Dronkers 1986; Bakker & Creemers 1994; Wolbers & De Graaf 1996). In dergelijk onderzoek staan meestal ouderlijke kenmerken, zoals het opleidingsniveau en de beroeps- en inkomenspositie, onder de aandacht. Ouders oefenen echter niet alleen direct door hun sociaal-economische kenmerken, invloed uit op het onderwijssucces van hun kinderen. Dergelijke achtergronden gaan vaak samen met meer inhoudelijke gezinsfactoren die een verklaring vormen voor schoolprestaties (Mercy & Steelman 1982; Meijnen 1985; Dronkers 1990).

In dit artikel worden drie verschillende aspecten van het ouderlijk gezin onderscheiden. Deze aspecten bieden als het ware een verklaring voor de samenhang tussen de ouderlijke sociaal-economische achtergronden en het succes van kinderen op school. Allereerst onder-

scheiden we een aantal sociaal-demografische kenmerken van het gezin, zoals het opgroeien in een één-oudergezin, het aantal broers en zusters en de leeftijd van de moeder. Het idee is grofweg dat minder interne binding door demografische oorzaken er toe kan leiden dat kinderen het minder goed doen op school. Een tweede traditie van onderzoek naar opleidingssucces stelt culturele hulpbronnen centraal. Het idee is dat een cultureel hoogwaardig thuis-klimaat een goede aansluiting biedt op de eisen die men in hogere schoolopleidingen aan leerlingen stelt. Hierdoor is succes voor kinderen met een dergelijke achtergrond meer waarschijnlijk. Een derde verklaring richt zich op de schoolbetrokkenheid van ouders. Belangstelling voor het kind en de school, en de aansporing tot schoolprestaties zijn daarbij belangrijke elementen. Ook deze sociaal-emotionele stimulering zou een aanzet kunnen zijn voor kinderen om succesvol te zijn.

In dit artikel gaan we na hoe deze kenmerken van het ouderlijk gezin het schoolsucces van kinderen beïnvloeden. Daarbij is van belang te onderkennen dat hoog opgeleide en welgestelde ouders meer en alternatieve wegen lijken te hebben om succes van hun kinderen in het onderwijs te bevorderen. Onze eerste, beschrijvende onderzoeksvraag luidt derhalve: *Hoe verschillen gezinnen van verschillende sociaal-economische achtergrond naar sociaal-demografische, sociaal-culturele en sociaal-emotionele omstandigheden?* En vervolgens luidt de tweede, verklarende vraag: *Wat is de relatieve invloed van deze sociaal-demografische, sociaal-culturele en sociaal-emotionele kenmerken van het gezin op het schoolsucces van kinderen?*

In onderwijskundig onderzoek is een maatschappelijke relevant thema dat van de verschillen tussen jongens en meisjes (De Jong 1985). Zo is er als gevolg van de campagne 'Kies Exact' veel aandacht voor verschillen waar het opleidingsrichting betreft (Bosker & Dekkers, 1994). Veel minder is er bekend over hoe invloeden van het ouderlijk gezin verschillen voor jongens en meisjes als het gaat om onderwijssucces. Onze derde onderzoeksvraag luidt daarom: *In hoeverre zijn er verschillen tussen jongens en meisjes als het gaat om de invloeden van sociaal-demografische, sociaal-culturele en sociaal-emotionele kenmerken van het ouderlijke gezin op schoolsucces?*

Op voorhand benadrukken we dat onze belangstelling in dit artikel uitgaat naar de 'totale' effecten van de kenmerken van het ouderlijk gezin op het presteren van kinderen in het voortgezet onderwijs. We zijn niet geïnteresseerd in de vraag in hoeverre eerdere prestaties of schoolkeuzes belangrijk zijn, zoals in onderzoek naar schoolloopbanen gebeurt (Tesser 1986; Bakker & Creemers 1994; Veenstra 1999). Als eerste heeft dit tot gevolg dat we geen schoolkenmerken in ons onderzoek betrekken. We gaan er vanuit dat als er in onze opzet schoolverschillen zijn, deze random over de leerlingen verdeeld zijn, of dat de ouders een keuze voor goed of slecht presterende scholen bepalen. Een tweede gevolg is dat we geen schooladvies, CITO-toetsscores of tussentijdse prestaties bestuderen. Alleen het uiteindelijk behaalde opleidingsniveau in het voortgezet onderwijs staat hier onder de aandacht.

2. THEORETISCHE VERKLARINGEN

2.1 Sociaal-demografische selectie

Om te beginnen wordt het succes van kinderen in het onderwijs beïnvloed door de samenstelling van het ouderlijk gezin. Ouders voorzien kinderen van hulpbronnen die belangrijk zijn voor het presteren en er zijn demografische omstandigheden waarin die overdracht wordt bemoeilijkt.

Als eerste sociaal-demografisch kenmerk is het opgroeien in een één-oudergezin relevant (Schneider & Coleman 1993). Aangenomen wordt dat het traditionele gezin, waarin twee ouders aanwezig zijn, een betere leefomgeving biedt om tot ontwikkeling te komen dan een één-oudergezin (Amato, 1993). In deze gedachtegang zijn ouders voor hun kinderen een belangrijke bron van ondersteuning, praktische hulp, informatie en (bege-)leiding. Daarnaast hebben de ouders een voorbeeldfunctie, waardoor kinderen sociale vaardigheden leren, zoals

onderhandelen, compromissen sluiten en samenwerken. Het ontbreken van één van de ouders, door scheiding of overlijden, heeft tot gevolg dat er minder, maar ook minder goed hulpbronnen worden overgedragen (Coleman 1988). Ook in financiële zin heeft het opgroeien in een één-ouder gezinssituatie consequenties. Alleenstaande ouders beschikken veelal over minder financiële mogelijkheden om hun kinderen maatschappelijk volwaardig te laten participeren (denk aan kosten voor tijdschriften, kranten, verenigingslidmaatschap, bibliotheekbezoek en school-gerelateerde activiteiten), met als gevolg dat schoolprestaties negatief beïnvloed kunnen worden. Empirisch onderzoek voor de Verenigde Staten wijst inderdaad uit dat het opgroeien in een één-oudergezin negatieve consequenties heeft voor de kansen op het al dan niet behalen van een 'high school' diploma en 'drop-out' kansen (Milne, Myers, Rosenthal & Ginsburg 1986; Astone & McLanahan 1991). Met name kinderen uit één-oudergezinnen ontstaan door een echtscheiding blijken het minder goed te doen (Astone & McLanahan 1991). Gegeven het feit dat ook in Nederland het aantal kinderen in één-oudergezinnen sterk toeneemt, lijkt het nagaan van deze relatie relevant (Veenstra 1999). De eerste hypothese luidt daarom: *Kinderen uit twee-ouder gezinnen zijn succesvoller in het voortgezet onderwijs dan kinderen uit één-ouder gezinnen.*

Een tweede demografisch aspect, waarvan we denken dat het relevant is voor schoolsucces, is gezinsgrootte, en wordt verwoord via de 'resource dilution' hypothese (Steelman & Powell 1989; Downey 1995). Deze hypothese komt er simpelweg op neer dat men er vanuit mag gaan dat de hulpbronnen van ouders niet onuitputtelijk zijn. De implicatie van deze gedachte is dat in grote gezinnen de aanwezige hoeveelheid hulpbronnen met meer broers en zussen moet worden gedeeld, waardoor er per individu minder overblijft dan in een kleiner gezin met een vergelijkbare hoeveelheid hulpbronnen (Downey 1995; Sui-Chu & Wilms 1996). Zo is er dan bijvoorbeeld per kind minder geld voor school-gerelateerde uitgaven en maatschappelijke participatie. Er is hoogstwaarschijnlijk ook minder persoonlijke aandacht van ouders in grote gezinnen voor individuele begeleiding en schoolvoortgang, en er zijn minder mogelijkheden om cultureel actief te zijn. Voor deze gedachte vindt men echter niet louter bevestigingen (Bakker & Cremers 1994; Van Eijck 1996; Teachman, Paasch & Carver 1996).

Een aanvullende hypothese richt zich daarnaast op de geboortevolgorde in gezinnen (Steelman & Powell 1989). Het idee is dat de oudste kinderen het langst profiteren van de situatie waarin men nog maar met weinig broers en zussen de ouderlijke aandacht en begeleiding hoeft te delen. De verwachting is dan dat deze kinderen in het onderwijs het meest succesvol zijn (Mercy & Steelman 1982; Van Eijck 1996). Hiertegen spreekt echter ook een door Steelman & Powell (1989) genoemde hypothese dat broers en zussen elkaar kunnen begeleiden en ondersteunen. Juist de jongere broers en zussen zouden op deze wijze profiteren van de verworvenheden (financieel, cultureel en schoolse kennis) van de oudere broers en zussen. Onze tweede hypothese is dat *kinderen uit kleine gezinnen succesvoller zijn in het voortgezet onderwijs dan kinderen uit grote gezinnen, waarbij de eerstgeborenen het beter doen dan de latergeborenen.*

Een derde demografisch kenmerk, relevant bij het presteren van kinderen op school is het leeftijdsverschil tussen moeder en kind (Mare & Tzeng 1989). Deze problematiek staat in de Amerikaanse literatuur ook wel als het 'teenage pregnancy deficit' onder de aandacht. De gedachte hierbij is dat het moederschap op jonge leeftijd vaak samengaat met slechte sociaal-economische omstandigheden van het gezin. Kinderen van oudere moeders kunnen daarentegen profiteren van betere economische omstandigheden in het gezin en een grotere emotionele volwassenheid van de ouders (Mare & Tzeng 1989). Echter, tegelijkertijd speelt het fenomeen dat als ouders het voortgezet onderwijs al lang geleden hebben verlaten (ongeacht het niveau), zij bij de directe ondersteuning van hun kinderen nog maar van beperkte betekenis zijn. Zowel als het gaat om interesse in schoolse en niet-schoolse zaken (begrip voor situatie en vriendenkring), als om het direct assisteren van kinderen bij schooltaken (helpen bij huiswerk) lijkt dit fenomeen relevant. De bovenstaande gedachtegang impliceert een krom-

lijng verband met schoolsucces. Onze derde hypothese luidt daarom als volgt: *Kinderen met een relatief oude moeder zijn succesvoller in het voortgezet onderwijs dan kinderen met een relatief jonge moeder. Echter, na een bepaalde leeftijd van de moeder wordt een nog oudere moeder een nadeel in plaats van een voordeel.*

2.2 Sociaal-culturele voordelen

Sociologisch onderzoek naar sociale en culturele reproductie probeert inzicht te krijgen in de rol van ouders bij het goed presteren van kinderen uit hogere sociale milieu's (DiMaggio 1982; DeGraaf 1987). Geïnspireerd op het werk van Bourdieu richt men zich in deze traditie op het cultureel kapitaal dat men meer in de maatschappelijk hogere kringen aantreft (Bourdieu & Passeron 1977). Cultureel kapitaal is op te vatten als een bekendheid met, en bedrevenheid in de dominante cultuur, zoals die op scholen zou worden gedoceerd en gepropageerd. Deze bekendheid met de juiste culturele codes uit zich in smaken, voorkeuren en gedragingen op allerlei terreinen die herkend en beloond worden in de hogere vormen van het onderwijs (Lamont & Lareau 1988; Lareau & Horvat 1999).

In het empirisch onderzoek wordt cultureel kapitaal meestal gerelateerd aan de ouderlijke participatie in elitaire cultuuruitingen, zoals het gaan naar het theater, concerten en musea (DeGraaf 1986; Niehof 1997). De veronderstelling daarbij is dat kinderen met zo'n culturele socialisatie in het ouderlijke gezin de schoolse omgeving als vertrouwd ervaren. Kinderen zonder een cultureel hoogwaardige thuisomgeving zouden de culturele codes missen die nodig zijn om het goed te doen op school. Meer specifiek worden daarvoor in de literatuur drie redenen genoemd (Lamont & Lareau 1988; Kalmijn & Kraaykamp 1996). Allereerst, zien scholieren zonder culturele socialisatie af van de hogere vormen van onderwijs door zelf-selectie; men ervaart het hoger onderwijs als een omgeving waar men zich niet thuis voelt (Brands 1992). Ten tweede, komen leerlingen met weinig cultureel kapitaal minder ver in het onderwijs omdat men niet de verwachte resultaten behaalt. De reden hiervoor is dat in het onderwijs de elitaire cultuur wordt gedoceerd waarvan men in de thuissituatie weinig heeft meegekregen. Als derde, kunnen leerlingen met een geringe hoeveelheid culturele hulpbronnen minder ver komen doordat docenten de kwaliteiten van dergelijke scholieren slechter herkennen. Dit zou dan veroorzaakt worden door een culturele kloof tussen docent en leerling (Lareau & Horvat 1999). Samenvattend, kan men stellen dat als culturele hulpbronnen relevant zijn men mag verwachten dat *kinderen met cultureel actieve ouders succesvoller zijn in het voortgezet onderwijs dan kinderen met cultureel inactieve ouders* (hypothese 4).

Een nuancering van bovenstaande cultureel kapitaal gedachte vindt men bij Farkas (1996), die stelt dat men het cultureel kapitaal van het ouderlijk gezin breder moet opvatten dan alleen de deelname aan elitaire uitgaanscultuur. Farkas (1996) veronderstelt dat op school naast non-cognitieve culturele codes ook cognitieve kwaliteiten worden beloond, en dat ook deze worden gesocialiseerd in het ouderlijk gezin. Meer specifiek houdt deze redenering in dat een omgeving, waar boeken aanwezig zijn, waar lezen een gewaardeerde activiteit is en waar men kinderen stimuleert om te lezen, positief uitpakt voor het opleidings- en leesniveau (Leseman 1994; Kraaykamp & Dijkstra 1999). Empirisch onderzoek van DeGraaf, DeGraaf en Kraaykamp (2000) voor Nederland wijst inderdaad uit dat het leesgedrag van ouders een belangrijke indicator voor het schoolsucces is naast het culturele uitgaan (zie ook Tesser & Mulder 1990). Dit wijst er op dat het culturele kapitaal, zeker niet alleen maar 'habitus' en 'dispositie' representeert. Ook een socialisatie in als cultureel te duiden cognitieve vaardigheden lijkt daarom van belang bij de onderwijsprestaties van kinderen. De rol van deze cognitieve aspecten in de ouderlijke opvoeding komt tot uitdrukking komt in onze vijfde hypothese: *Kinderen met lezende ouders zijn succesvoller in het voortgezet onderwijs dan kinderen met niet-lezende ouders.*

Dat culturele hulpbronnen verschillend kunnen uitpakken voor jongens en meisjes illustreert DiMaggio (1982) in zijn studie naar de gevolgen van (eigen) cultureel kapitaal voor schoolresultaten. Met name meisjes blijken in het onderwijs te profiteren van een cultureel hoog-

waardig thuis klimaat. Meer recent onderzoek van Aschaffenburg & Maas (1997) vindt voor deze veronderstelling echter geen ondersteuning. Cultureel kapitaal biedt in hun onderzoek jongens en meisjes evenveel voordelen.

2.3 Sociaal-emotionele ondersteuning

De directe interesse en aandacht van ouders voor het presteren van kinderen op school is ook een belangrijke indicator voor het uiteindelijke schoolsucces (Schneider & Coleman 1993). Coleman (1988) interpreteert de ouderlijke belangstelling voor school als een sociale hulpbron, die wordt opgebouwd in de relatie tussen ouders en kinderen. De intensiteit van de relatie van ouders met elkaar, en met hun kinderen resulteert in een binding van kinderen met het gezin waardoor normen en verwachtingen over het belang van scholing duidelijker en overtuigender worden overgedragen (Furstenberg & Hughes 1995). In een overzicht van de literatuur onderscheiden Sui-Chu & Willms (1996) grofweg twee uitwerkingen van deze ouderlijke betrokkenheid; enerzijds een benadering waarbij het thuis klimaat zich kenmerkt door een positieve betrokkenheid en een wederzijdse discussie over schoolonderwerpen, en anderzijds een benadering waarbij restrictieve en controlerende aspecten (huiswerk maken, bewaken voortgang) centraal staan.

Coleman (1988) benadrukt het feit dat de interactie tussen ouder en kind positief dient te zijn. Alleen als de ouder-kind interactie een positieve teneur kent, zou het voor ouders mogelijk zijn ook andersoortige hulpbronnen op hun kinderen over te dragen. Recent empirisch onderzoek van Muller (1995) en Sui-Chu & Willms (1996) wijst uit dat positieve aandacht in de thuisomgeving (discussie en belangstelling) leidt tot significante hogere resultaten bij de vakken wiskunde en taal. Onderzoek van Rumberger, Ghatak, Poulos, Ritter en Dornbusch (1990) en Teachman, Paasch en Carver (1996) toont aan dat ook de 'high school drop-out' kansen aanzienlijk lager zijn als er positieve ouderlijke aandacht is voor schoolprestaties en -ervaringen. Onze algemene verwachting is daarom dat *kinderen waarvan ouders een positieve betrokkenheid bij school hebben succesvoller zijn in het voortgezet onderwijs dan kinderen waarvan ouders geen positieve betrokkenheid bij school hebben* (hypothese 6). Sui-Chu & Willms (1996) wijzen in dit verband op het fenomeen dat meisjes meer ouderlijke aandacht ontvangen dan jongens. Onderzoek van Muller (1998) laat tevens zien dat bij meisjes het bepraten van schoolaangelegenheden door ouders ook nog eens positiever uitpakt dan bij jongens.

Ouders kunnen ook door een negatieve schoolbetrokkenheid van invloed zijn op schoolprestaties. Daarbij denken we dan specifiek aan de situatie waarin de ouder-kind-interactie de negatieve lading krijgt van controleren, aansporen en dwingen. Rumberger, Ghatak, Poulos, Ritter en Dornbusch (1990) betogen dat de gezinssituatie van 'drop-outs' zich kenmerkt door een streng autoritair gezinsregime. Muller (1995) wijst in dit verband op de negatieve effecten van huiswerkbegeleiding en -controle door ouders op wiskunde-testresultaten. Hierbij is echter relevant te onderkennen dat ouders wellicht pas tot huiswerkbegeleiding overgaan als hun kind slecht presteert. De causaliteit is in dat geval dus omgekeerd. Opmerkelijk is verder dat Amerikaans onderzoek uitwijst dat meisjes significant minder vaak negatieve aandacht voor schoolgedrag krijgen dan jongens (Sui-Chu & Willms 1996). Dit heeft tot gevolg dat we verwachten dat negatieve ouderlijke betrokkenheid bij jongens meer negatieve consequenties zal hebben dan bij meisjes. De zevende hypothese luidt: *Kinderen waarvan ouders geen negatieve betrokkenheid bij school hebben zijn succesvoller in het voortgezet onderwijs dan kinderen waarvan ouders negatieve betrokkenheid bij school hebben*.

Ter afsluiting van onze bespreking van theoretische achtergronden is het relevant te benadrukken dat de drie bovengenoemde onderzoekstradities een met elkaar in verband staande verklaring bieden voor de sociaal-economische invloeden op onderwijsresultaten van kinderen. Als gevolg daarvan is het van belang de verwachtingen ten aanzien van de bovenstaande gezinsaspecten simultaan in een multivariate analyse te toetsen.

3. DATA EN ONDERZOEKSOPZET

3.1 Het Voortgezet Onderwijs Cohort Leerlingen 1989 (VOCL)

Om na te gaan welke gezinsaspecten belangrijk zijn bij een verklaring van het succesvol zijn van kinderen in het voortgezet onderwijs, gebruiken we informatie uit het cohortonderzoek VOCL-1989 van het Centraal Bureau voor de Statistiek (CBS, 1991). Het primaire doel van het VOCL-onderzoek is te bestuderen welke relaties er bestaan tussen de sociale achtergrondkenmerken, de schoolsituatie en de schoolprestaties van leerlingen (CBS, 1991). In het VOCL is informatie beschikbaar over 19524 leerlingen die zich in 1989 in het eerste leerjaar van het voortgezet onderwijs bevonden. Deze leerlingen zijn afkomstig van 381 scholen (uit een aselechte steekproef van 826 scholen).¹ In 1989 is leerlingen uit het eerste leerjaar verzocht een oudervragenlijst mee naar huis te nemen en door de ouders te laten invullen. In deze oudervragenlijst is onder andere gevraagd naar demografische en culturele kenmerken van het ouderlijk gezin en naar de schoolbetrokkenheid van de ouders. De school is gevraagd informatie te verstrekken over de leerling zelf (leeftijd, geslacht, lager schooladvies, schoolsoort in het afgelopen jaar en testresultaten). Vervolgens is in elk schooljaar volgend op 1989/1990 aan het VOCL-bestand informatie toegevoegd over de voortgang van de leerlingen. Het gaat dan om examenresultaten, diplomabezit en (eventueel) schoolverlating. In dit artikel gebruiken we de beschikbare informatie tot en met september 1997. Onze analyses gaan derhalve over de situatie na acht leerjaren in het voortgezet onderwijs. Na een selectie op missende waarnemingen beschikken we over een onderzoeksbestand met 14864 respondenten. Leerlingen zonder enig diploma na 8 jaar (N=1787) zijn niet in onze analyses meegenomen, omdat het fenomeen van schooluitval een afzonderlijke benadering vraagt.

3.2 Een beschrijving van de gemeten kenmerken

Als afhankelijk kenmerk staat in dit artikel het hoogst behaalde opleidingsniveau onder de aandacht. In onze onderzoeksopzet is dit gemeten aan de hand van het hoogste diploma na acht jaar voortgezet onderwijs. Deze meting houdt in dat diploma's behaald in het hoger beroepsonderwijs en op de universiteit niet voorkomen. Na acht jaar had nog geen leerling een dergelijke opleiding afgerond. Schoolsucces kent derhalve vijf niveaus lopend van lbo (1), mavo (2), mbo (3), havo (4), tot vwo (5).²

De meting van de onafhankelijke kenmerken komt hieronder aan de orde. Indien mogelijk zijn de aspecten voor vader en moeder onafhankelijk geoperationaliseerd om vast te kunnen stellen welke ouder het meest invloedrijk is. Tevens wordt in Tabel 1 voor elke van deze verklarende kenmerken de bi-variate relatie met het hoogst behaalde diploma beschreven.

Allereerst zijn de sociaal-economische achtergronden van de ouders gemeten door het *opleidingsniveau van vader en moeder*. Het opleidingsniveau van beide ouders is in vijf categorieën beschikbaar, lopend van lager onderwijs (0), voortgezet onderwijs, laag niveau (1), voortgezet onderwijs, hoog niveau (2), hoger onderwijs, eerste fase (3) tot hoger onderwijs, tweede fase (4). Uit Tabel 1 blijkt overduidelijk de samenhang met het schoolsucces van de kinderen; ouders van kinderen met een lbo-diploma zijn gemiddeld duidelijk lager geschoold (score: 1.0 en 0.8), dan ouders van leerlingen met een vwo-diploma (score: 2.4 en 1.8). Ten tweede is de sociaal-economische achtergrond van leerlingen gemeten door te kijken naar de hoogst voorkomende beroepspositie in het ouderlijke gezin (CBS, 1991). Hierbij is een onderscheid gemaakt naar *lagere beroepen*, *middelbare beroepen* en *hogere beroepen* in loondienst, en *kleine zelfstandigen*. Het idee is dat met deze variabele enig inzicht wordt verkregen in de financiële mogelijkheden en beperkingen die men in het ouderlijk gezin ervaart. Uiteraard hadden we liever de beschikking gehad over het huishoudinkomen van het ouderlijk gezin. Wederom zijn de relaties met het schoolsucces van de kinderen overduidelijk (Tabel 1). Waar 6% van de leerlingen met een lbo-diploma een ouder heeft met een hoog beroep, heeft 37% van de geslaagde vwo-leerlingen dat. Opmerkelijk is het dat we kinderen met een mbo-diplo-

Tabel 1: Sociaal-economische, sociaal-demografische, sociaal-culturele en sociaal-emotionele aspecten van het ouderlijk gezin naar soort diploma, 8 jaar na instroom in het voortgezet onderwijs.

Onafhankelijke Kenmerken	gemiddelde score op onafhankelijke kenmerk naar het hoogst behaalde diploma					gem.	Pearson corr.
	lbo	mavo	mbo	havo	vwo		
Sociaal-economische kenmerken ouders							
Opleidingsniveau vader (0-4)	1.02	1.51	1.39	1.93	2.42	1.57	.37
Opleidingsniveau moeder (0-4)	0.76	1.11	1.06	1.44	1.84	1.18	.35
Sociaal milieu ouders							
hogere beroepen (ref.)	6,2%	14,8%	12,4%	22,6%	36,6%	16,6%	.24
zelfstandigen (0-1)	7,0%	9,5%	10,4%	8,5%	7,7%	8,7%	.01
midden beroepen (0-1)	14,0%	21,9%	20,4%	28,2%	28,8%	21,7%	.12
lagere beroepen (0-1)	72,8%	53,8%	56,9%	40,7%	26,9%	53,1%	-.29
Sociaal-demografische kenmerken ouders							
Gezin zonder vader (0-1)	6,9%	6,9%	4,1%	5,3%	4,1%	5,4%	-.05
Gezin zonder moeder (0-1)	1,1%	1,0%	0,6%	0,6%	0,4%	0,8%	-.03
Aantal kinderen (1-8)	2,59	2,58	2,56	2,56	2,57	2,57	-.01
Oudere broers/zussen							
geen (0-1)	42,4%	44,9%	44,3%	46,3%	50,1%	45,1%	.05
1-2 oudere broers/zussen (0-1)	52,7%	51,1%	51,9%	50,9%	46,8%	51,0%	-.03
meer dan 2 oudere broers/zussen (0-1)	5,0%	4,0%	3,9%	2,8%	3,1%	3,9%	-.04
Leeftijd moeder 1989	38,9	39,1	39,3	39,8	40,3	39,4	.11
Sociaal-culturele kenmerken ouders							
Cultureel uitgaan ouders (0-9)	3,38	4,41	4,24	5,27	5,98	4,49	.28
Leesaanbod ouderlijk huis (0-9)	3,72	4,55	4,21	5,24	6,08	4,60	.26
Lezen vader (0-7)	0,79	0,99	0,90	1,20	1,41	1,02	.13
Lezen moeder (0-7)	1,58	1,82	1,77	2,13	2,30	1,87	.12
Sociaal-emotionele kenmerken ouders							
Positieve aandacht vader (0-1)	53,8%	57,6%	59,0%	59,8%	59,5%	57,7%	.04
Positieve aandacht moeder (0-1)	68,8%	73,9%	73,6%	75,1%	75,9%	73,0%	.05
Negatieve aandacht vader (0-1)	44,2%	38,7%	34,3%	27,6%	17,9%	33,8%	-.18
Negatieve aandacht moeder (0-1)	49,5%	44,9%	38,3%	30,9%	19,1%	38,0%	-.20

Bron VOCL-89 (N=13077).

ma relatief vaak aantreffen in een gezin waarvan ouders als zelfstandige werkzaam zijn. Blijkbaar wordt een beroepsopleiding in deze gezinnen meer gestimuleerd en gewaardeerd.

Bij de sociaal-demografische aspecten is om te beginnen nagegaan of het een één-oudergezin betrof. Daarbij is verschil gemaakt tussen *gezinnen zonder vader* en *gezinnen zonder moeder*. Dit is gedaan om te kunnen vaststellen of het opgroeien met alleen een moeder of alleen een vader tot verschillende uitkomsten in het onderwijs leidt. Uit Tabel 1 blijkt inderdaad een negatieve samenhang van schoolsucces met het afkomstig zijn uit een één-oudergezin; van de geslaagde lbo-leerlingen is 8% afkomstig uit een gezin zonder vader en 2% uit een gezin zonder moeder, terwijl dat voor leerlingen met een vwo-diploma respectievelijk 5% en 1% is. De correlaties zijn echter niet erg hoog ($r=-.04$ en $r=-.03$). Het aantal kinderen in het ouderlijk gezin is continue gemeten, lopend van 1 kind tot 8 of meer kinderen (inclusief de leerling zelf). De samenhang met onderwijssucces is relatief gering ($r=-.01$). Om een (te hoge) associatie met het aantal kinderen te vermijden, is het aantal oudere broers/zussen gedummyficeerd in 3 variabelen, te weten geen oudere broers/zussen, 1 of 2 oudere broers/

zussen, en meer dan 2 oudere broers/zussen. Het opgroeien in een gezin zonder oudere broers of zussen blijkt positief uit te pakken voor het hoogst behaalde opleidingsniveau; van de leerlingen met een lbo-diploma als hoogste, heeft 42% geen oudere broer of zus terwijl bij de vwo-geslaagden dit voor 50% opgaat ($r=.05$). Tenslotte is de *leeftijd van moeder in 1989* als continue kenmerk opgenomen.³ Aangezien in het VOCL-bestand alle leerlingen van min of meer dezelfde leeftijd zijn (plusminus 12 jaar) wordt daarmee een predictor voor het leeftijdsverschil tussen de leerling en de moeder geïntroduceerd. De samenhang met het schoolsucces is vastgesteld in Tabel 1. Vooral leerlingen met relatief 'oude' moeders blijken het goed te doen in het voortgezet onderwijs. Van de leerlingen met een vwo-diploma is de moeder gemiddeld 40.3 jaar in 1989, tegen 38.9 jaar voor geslaagde lbo-ers. Ook de bi-variate associatie is relatief hoog ($r=-.11$ met moeder's geboortejaar).

Van de sociaal-culturele achtergrond is het *cultureel uitgaansgedrag van ouders* vastgesteld met drie vragen in de (ouder-)vragenlijst over het bezoeken van concerten, toneel of ballet, en kunstmusea- of tentoonstellingen (apart aan beide ouders gesteld). De antwoordmogelijkheden waren nooit, 1-2 keer per jaar, 3-6 keer per jaar en meer dan 6 keer per jaar. Aangezien partners culturele uitgaansactiviteiten meestal samen ondernemen is er voor de ouders één schaal geconstrueerd ($\alpha=.80$). De scores op de zes items zijn gestandaardiseerd en opgeteld (met een minimum van 2 geldige waarnemingen), en daarna in decielen verdeeld. Het resultaat is een schaal met 10 categorieën waarbij de ouders evenredig over de categorieën zijn verdeeld. Daarbij staat de categorie 0 voor zeer weinig cultureel, terwijl de categorie 9 de zeer culturele ouders representeert. Tabel 1 laat zien dat het behalen van een hoog schooldiploma vaak samengaat met cultureel actieve ouders ($r=.28$).

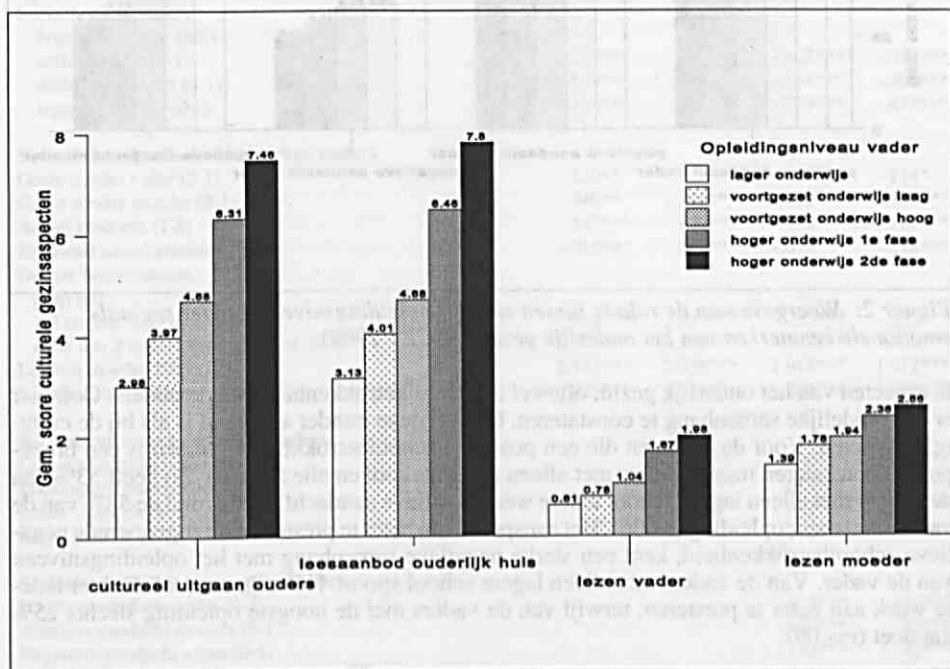
Om het *leesaanbod thuis* te meten, is om te beginnen het boekenbezit van de ouders genomen (12 categorieën lopend van 0 tot meer dan 1000 boeken). Het boekenkoopgedrag van de beide ouders is gemeten door na te gaan hoeveel boeken zij per maand aanschaffen (7 categorieën van 0 tot meer dan 5 boeken). Deze drie kenmerken van het leesaanbod thuis zijn gestandaardiseerd en opgeteld, en daarna in decielen onderverdeeld ($\alpha=.73$). De categorie 0 van de variabele leesaanbod thuis vertegenwoordigt de gezinnen waar slechts weinig boeken beschikbaar zijn en komen, terwijl de categorie 9 de ouders betreft waar de aanwezigheid van boeken een overduidelijk gegeven is. De relatie van het boekenaanbod met het hoogst behaalde diploma is relatief sterk ($r=.26$). Het daadwerkelijk *lezen* van beide ouders is vastgesteld aan de hand van de vragen over het aantal boeken dat men per maand leest (8 categorieën lopend van 0 boeken tot meer dan 6 boeken). Op basis van Tabel 1 is vast te stellen dat de samenhang voor moeder's en vader's lezen met het onderwijssucces van hun kinderen overeenkomstig is. Van de leerlingen met een lbo-diploma leest de moeder gemiddeld 1.58 boeken per maand, tegen 2.30 boeken bij de leerlingen met een vwo-diploma.

Ook bij de meting van sociaal-emotionele ondersteuning is er voor gekozen vaders en moeders apart te beschouwen. Voor beide zijn er twee aspecten vastgesteld, te weten de positieve en negatieve schoolbetrokkenheid. De variabele *positieve aandacht voor de leerling* is gemeten met drie items uit de vragenlijst, namelijk het praten met het kind over gebeurtenissen op school, praten met het kind over prestaties op school en het kind complimenten maken voor schoolprestaties. De antwoordmogelijkheden waren nooit, een of enkele keren per jaar, een of enkele keren per maand en een of enkele keren per week. Er is een 0/1-variabele geconstrueerd die aangeeft of een ouder iedere week aan het kind positieve aandacht schenkt of niet. Alleen ouders die bij alle drie de items aangaven dat te doen krijgen de score 1, anders krijgen ze de score 0; voor de vaders is dit in 57% van de gevallen zo en bij de moeders in 73%. De samenhang van deze positieve aandacht van ouders voor school met het uiteindelijk behaalde diploma is echter beperkt ($r=.05$ voor moeders en $r=.04$ voor vaders). De variabele *negatieve aandacht voor de leerling* is gemeten met een item dat aangeeft of men het kind aanspoort harder te werken. Wederom is een 0/1-variabele geconstrueerd die aangeeft of men dat iedere week doet; van de vaders is dat 35% en van de moeders 40%. Uit Tabel 1 blijkt dat deze negatieve ouderlijke belangstelling inderdaad minder goed uitpakt als het gaat om school-

succes; van de vwo-leerlingen blijkt zo'n 18% negatieve aandacht van hun ouders te krijgen, terwijl dat voor de mavo- en lbo-leerlingen ruim 40% is.

3.3 Sociaal-economische verschillen in het ouderlijk culturele gedrag en de ouderlijke schoolbetrokkenheid

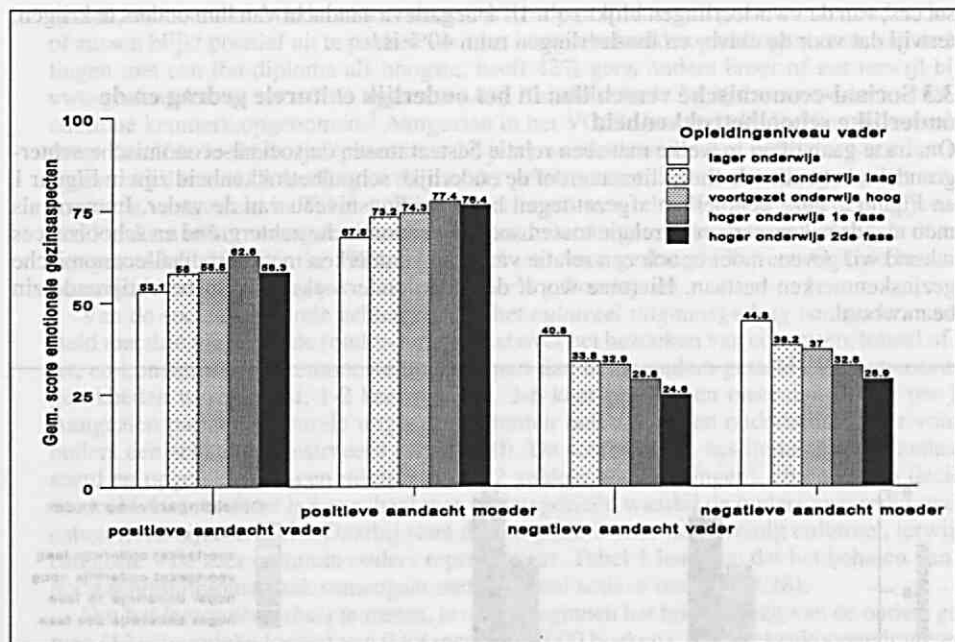
Om na te gaan of en in welke mate een relatie bestaat tussen de sociaal-economische achtergrond, en het culturele thuis klimaat en/of de ouderlijke schoolbetrokkenheid zijn in Figuur 1 en Figuur 2 deze kenmerken afgezet tegen het opleidingsniveau van de vader. Immers, als men met deze kenmerken de relatie tussen sociaal-economische achtergrond en schoolsucces inhoud wil geven, moet er ook een relatie van deze kenmerken met de sociaal-economische gezinskenmerken bestaan. Hiermee wordt de eerste onderzoeksvraag in beschrijvende zin beantwoord.



Figuur 1: Weergave van de relatie tussen vaders opleidingsniveau en vier sociaal-culturele kenmerken van het ouderlijk gezin (VOCL - 1989).

In Figuur 1 is voor de vier geselecteerde culturele aspecten van het ouderlijke gezin de relatie met vader's opleidingsniveau gepresenteerd. Met name voor het cultureel uitgaan en het boekenaanbod thuis is de samenhang hoog; vaders met alleen lager onderwijs scoren 2.98 op schaal voor het culturele uitgaan, tegen 7.46 voor de vaders met tweede fase hoger onderwijs ($r=.40$). Voor de schaal die het leesaanbod thuis indiceert is de samenhang zelfs nog iets groter ($r=.43$). Ook het daadwerkelijk lezen is gecorreleerd met vader's onderwijsniveau; laag opgeleide vaders lezen gemiddeld 0.6 boeken per maand, terwijl de hoog opgeleide vaders gemiddeld 1.98 boeken per maand lezen ($r=.27$). Hiermee is een samenhang tussen de sociaal-economische gezinsachtergrond en de intensiteit van de culturele participatie van ouders overduidelijk vastgesteld.

Figuur 2 toont de associatie tussen vader's opleidingsniveau en de vier sociaal-emotione-



Figuur 2: Weergave van de relatie tussen vaders' opleidingsniveau en vier sociaal-emotionele kenmerken van het ouderlijk gezin (VOCL - 1989).

le aspecten van het ouderlijk gezin, oftewel de schoolbetrokkenheid van de ouders. Ook hier is een duidelijke samenhang te constateren, hoewel deze minder sprekend is als bij de culturele aspecten. Voor de aspecten die een positieve schoolbetrokkenheid meten is een breekpunt te constateren tussen vaders met alleen lagere school en alle anderen. Zo geeft 53% van de vaders met alleen lagere school iedere week positieve aandacht, tegen rond de 58% van de vaders met meer opleiding ($r=.06$). Het aansporen om beter te presteren, de zogenoemde negatieve schoolbetrokkenheid, kent een sterke negatieve samenhang met het opleidingsniveau van de vader. Van de vaders met alleen lagere school spoort 41% zijn zoon of dochter iedere week aan beter te presteren, terwijl van de vaders met de hoogste opleiding slechts 25% dat doet ($r=-.09$).

4. ANALYSES

Om na te gaan welke de relatieve bijdrage is van elk van de hier gepresenteerde aspecten van het ouderlijk gezin is via een multi-variate regressie een vijftal modellen geschat. In elk model wordt een aantal aspecten toegevoegd aan het voorgaande model. Op deze wijze is vast te stellen, welke aspecten relevant zijn bij het verklaren van schoolsucces en in hoeverre effecten van bepaalde aspecten van het thuisklimaat worden geïnterpreteerd door andere gezinsaspecten. In Tabel 2 presenteren we de ongestandaardiseerde regressie-coëfficiënten. Deze zijn te interpreteren in termen van de schaal van onze afhankelijke variabele, het hoogst behaalde diploma na 8 jaar in het voortgezet onderwijs.

Opvallend in het eerste model van Tabel 2 is het substantiële effect van geslacht; het zijn de meisjes die na 8 jaar voortgezet onderwijs een hoger diploma bezitten dan de jongens. Veel cross-sectioneel onderzoek presenteert onderwijsachterstanden voor meisjes, maar blijkbaar zijn in de huidige cohorten deze achterstanden weggerukt en zelfs omgezet naar een voor-

Tabel 2. Regressie van hoogst behaalde diploma 8 jaar na instroom in het voortgezet onderwijs naar sociaal-economische, sociaal-demografische, sociaal-culturele en sociaal-emotionele aspecten van het ouderlijk gezin.

	ongestandaardiseerde regressie-coëfficiënten				
	Model I	Model II	Model III	Model IV	Model V
	b	b	b	b	b
Geslacht (0-1: leerlinge=1)	.368***	.373***	.367***	.318***	.239***
Sociaal-economische kenmerken ouders					
Opleidingsniveau vader (0-4)	.251***	.240***	.213***	.202***	.202***
Opleidingsniveau moeder (0-4)	.290***	.274***	.236***	.214***	.214***
Sociaal milieu ouders					
hogere beroepen (ref.)					
zelfstandigen (0-1)	-.214***	-.209***	-.170***	-.168***	-.242***
midden beroepen (0-1)	-.181***	-.164***	-.149***	-.154***	-.203***
lagere beroepen (0-1)	-.467***	-.425***	-.374***	-.379***	-.477***
Sociaal-demografische kenmerken ouders					
Gezin zonder vader (0-1)		-.130**	-.127**	-.113*	-.114*
Gezin zonder moeder (0-1)		-.340**	-.376**	-.351**	-.352**
Aantal kinderen (1-8)		.157***	.159***	.166***	.167***
Kwadraat aantal kinderen		-.019***	-.019***	-.018***	-.018***
Oudere broers/zussen					
geen (ref.)					
1-2 oudere broers/zussen (0-1)		-.236***	-.237***	-.235***	-.235***
meer dan 2 oudere broers/zussen (0-1)		-.313***	-.319***	-.342***	-.337***
Leeftijd moeder 1989 (/10)		2.143***	2.038***	1.918***	1.912***
Kwadraat leeftijd moeder 1989		-.229***	-.218***	-.201***	-.201***
Sociaal-culturele kenmerken ouders					
Cultureel uitgaan ouders (0-10)			.034***	.030***	.030***
Boekenbezit ouders (0-10)			.013***	.013***	.023***
Lezen vader (0-7)			-.006***	-.007***	-.007***
Lezen moeder (0-7)			.014***	.010***	.010***
Sociaal-emotionele kenmerken ouders					
Positieve aandacht vader (0-1)				.114***	.115***
Positieve aandacht moeder (0-1)				.130***	.129***
Negatieve aandacht vader (0-1)				-.133***	-.182***
Negatieve aandacht moeder (0-1)				-.391***	-.392***
Interacties met geslacht					
Leerlinge * hogere beroepen (ref.)					
Leerlinge * zelfstandigen					.150
Leerlinge * midden beroepen					.094
Leerlinge * lagere beroepen					.193**
Leerlinge * boekenbezit ouders					-.020**
Leerlinge * negatieve aandacht vader					.107**
Constante	2.200***	-.758***	-2.700***	-2.382***	-2.326***
Verklaarde variantie (adj.)	18.5%	21.5%	22.1%	24.7%	24.8%

Bron VOCL-89, (N=13077).

Significantie *p<.05; **p<.01; ***p<.001.

sprong. Het eerste model laat verder de effecten van de sociaal-economische gezinsachtergrond van een leerling zien. Het opleidingsniveau van zowel vader als moeder is relevant. De schaal van het ouderlijk opleidingsniveau loopt van 0 tot 4, hetgeen inhoudt dat beide effecten substantieel zijn. Een leerling(e) waarvan beide ouders het hoogste opleidingsniveau bezitten, heeft gemiddeld een 2.2 punten hoger diploma ($4 \cdot 251 + 4 \cdot 290$) dan een leerling(e) waarvan beide ouders alleen lagere school hebben gevolgd. Ook het beroepsniveau van de ouders blijkt relevant. Vooral leerlingen met ouders uit lage beroepen hebben een significant lager diploma 8 jaar na intrede in het voortgezet onderwijs ($b = -.467$).

Het tweede model introduceert de sociaal-demografische aspecten van de thuissituatie. Zoals verwacht blijkt het opgroeien in een één-oudergezin negatief uit te pakken voor het schoolsucces. Vooral kinderen die een moeder missen doen het slechter op de middelbare school ($b = -.340$). Eerder onderzoek naar gezinsomvang rapporteerde vaak negatieve effecten van het aantal broers en zussen. In het VOCL-89 cohort blijkt de situatie genuanceerder te liggen. Er is sprake van een kromlijng verband, vastgesteld via een kwadraatterm; meer broers/zussen werkt in eerste instantie positief uit op schoolsucces tot een zeker punt, waarna nog meer broers/zussen negatief uitpakt. Het omslagpunt ligt in Model II plusminus bij 3 ($= (.157/2 \cdot .019) - 1$) broers en zussen. Een tweede aspect van de gezinsomvang heeft betrekking op de geboortevolgorde. Uit onze analyses blijkt dat het hebben van oudere broers en zussen negatieve consequenties heeft voor de hoogte van het behaalde diploma. Anders gezegd, eerstgeborenen doen het in het onderwijs significant beter dan de latergeborenen. Een laatste opmerkelijk demografische aspect dat belangrijk blijkt, is de leeftijd van de moeder. Zoals voorspeld betreft het hier een kromlijng verband. Hoe ouder de moeder hoe beter de leerling(e) het doet op school tot een bepaald omslagpunt. Hier is dat 46.8 jaar bij intrede in het middelbaar onderwijs ($= 10 \cdot (2.143/2 \cdot .229)$). Blijkbaar geldt Mare's (1989) argument van een betere financiële basis van oudere ouders slechts tot op zekere hoogte. Daarna is de leeftijds kloof tussen de schoolgaande leerling en moeder zo groot dat het negatieve consequenties heeft.

De aspecten van het culturele thuisklimaat worden in Model III toegevoegd. In het bijzonder cultureel uitgaan blijkt relevant bij de verklaring van schoolsucces. Deze culturele hulpbronnen van de ouders betalen zich uit in het onderwijssucces van hun kinderen. Daarnaast blijkt ook het leesaanbod in het ouderlijk huis en het daadwerkelijk lezen van de moeder van belang in de verklaring van de hoogte van het diploma na 8 jaar voortgezet onderwijs. Echter, over het algemeen geldt dat de effecten van de ouderlijke culturele hulpbronnen betrekkelijk klein zijn in verhouding tot de sociaal-economische en demografische aspecten van het thuisklimaat. Daarbij komt dat zij slechts in zeer beperkte mate de effecten van sociaal-economische gezinsachtergrond intermedieëren; tussen Model II en Model III neemt vader's opleidingseffect af met slechts 13%, en moeder's opleidingseffect met 16%. Dat er dus sprake is van culturele reproductie, in termen van Bourdieu, is in deze sterk uitgedrukt.

Model IV brengt de vier aspecten van ouderlijke schoolbetrokkenheid in de analyse. Alle vier de kenmerken zijn significant. Zowel positieve aandacht van de vader als van de moeder heeft een positief effect op het schoolsucces van de leerling. Positieve schoolbetrokkenheid betaalt zich dus uit. Aan de andere kant werkt negatieve aandacht van ouders niet positief; vooral voor de moeder geldt dat voortdurende aansporing (eens per week) leidt tot minder presteren in het voortgezet onderwijs ($b = -.391$). Belangrijk is hierbij te vermelden dat het hier een meting in de eerste klas van het middelbaar onderwijs betreft en het diploma minstens vier jaar na deze vaststelling kon worden behaald. Hoewel hiermee niet alle causaliteitsproblemen worden omzeild, wordt wel aannemelijk dat de sociaal-emotionele betrokkenheid van ouders voorafgaat aan het presteren van hun kinderen en niet het gevolg is van te weinig schoolsucces.

Tenslotte wordt met Model V getracht een antwoord te geven op onze derde onderzoeksvraag naar de verschillen tussen jongens en meisjes als het gaat om de relatieve invloed van de demografische, culturele en emotionele aspecten van het thuisklimaat. De procedure die we hebben gehanteerd is de volgende. Allereerst zijn alle interacties van geslacht met de kenmerken van het thuisklimaat opgenomen. Vervolgens zijn één voor één de minst significan-

te interacties verwijderd, tot uiteindelijk het hier gepresenteerde model resulteert. Het over het hoofd zien van eventuele 'suppressor' effecten is door deze procedure tot een minimum beperkt. De gepresenteerde interactie-effecten in Model V geven inzicht in de relevante verschillen tussen jongens en meisjes. Over het algemeen geldt dat de verschillen voor jongens en meisjes betrekkelijk gering zijn; van de zestien getoetste interacties zijn er slechts 3 significant. Om te beginnen is opmerkelijk dat meisjes uit de lagere sociale milieu's het relatief beter doen dan jongens ($b=.193$). Gegeven de grootte van het hoofdeffect ($b=-.477$), houdt dit echter niet in dat een laag beroep van de ouders voor meisjes volledig wordt gecompenseerd. Verder blijken vooral jongens te profiteren van de aanwezigheid van boeken in de thuis-situatie. Anders gezegd, voor jongens is het effect significant ($b=.023$), terwijl het voor meisjes vrijwel nihil is ($b=.023-.020$). Het meest opmerkelijke verschil tussen jongens en meisjes is te constateren bij de ouderlijke betrokkenheid van de vader. Aansporingen van vaders om harder te werken op school blijken voor meisjes minder negatief uit te pakken dan voor jongens, gegeven het positieve interactie effect ($b=.107$). Blijkbaar zijn meisjes meer autoriteitsgevoelig waar het de vader betreft dan jongens.

5. CONCLUSIE EN DISCUSSIE

In dit artikel zijn we nagegaan welke aspecten van het ouderlijk gezin belangrijk zijn in de verklaring van het hoogst behaalde diploma 8 jaar na intrede in het voortgezet onderwijs. Daartoe zijn sociaal-demografische, sociaal-culturele en sociaal-emotionele aspecten onderscheiden. Om onze verwachtingen te toetsen is gebruik gemaakt van de VOCL-89 data ($N=19524$). Deze gegevens betreffen een cohort van leerlingen die allen in het najaar van 1989 in het eerste jaar van het voortgezet onderwijs verbleven. Via de scholen is informatie verkregen over het presteren van de leerlingen en via een ouder-enquête over de thuissituatie.

Onze analyses wijzen uit dat de sociaal-economische achtergrond van de ouders, gemeten met ouderlijk opleidings- en beroepsniveau, het belangrijkste is en blijft. Zelfs na introductie van alle andere relevante kenmerken wordt slechts 20% van het opleidingseffect van de vaders geïnterpreteerd en 26% van het opleidingseffect van de moeders. De effecten van het ouderlijk beroepsniveau veranderen zelfs in het geheel niet. Van de meer inhoudelijke aspecten van het ouderlijke thuisklimaat zijn de demografische kenmerken het meest relevant, waarbij opmerkelijk genoeg vooral de leeftijd van de moeder belangrijk is. Wellicht dat hierbij de financiële omstandigheden een belangrijke rol spelen. Ook bij het gevonden negatieve effect van het opgroeien in een één-oudergezin zou deze additionele verklaring kunnen opgaan. Verder is vastgesteld dat van een groot aantal broers en zussen niet de gedachte negatieve invloed op het schoolsucces uitgaat. De meest 'ideale' situatie is een gezinsomvang van 4 kinderen, waarbij zowel meer als minder kinderen negatiever uitpakt. Ook de geboortevolgorde is van belang. De eerstgeborenen doen het het best, wellicht omdat zij de ouderlijke hulpbronnen het meest intensief gebruiken, terwijl bij latere kinderen wordt vertrouwd op onderlinge ondersteuning.

Het meest verrassende resultaat uit onze analyses is misschien wel de geringe invloed die er van het ouderlijk cultureel kapitaal uitgaat op het bereikte middelbare schooldiploma. De invloeden van cultureel actieve ouders zijn er wel en zijn significant, maar spreken minder tot de verbeelding dan in eerder cross-sectioneel onderzoek naar culturele reproductie (De Graaf 1986). Ook de schoolbetrokkenheid van ouders blijkt relevant in de verklaring van schoolsucces, waarbij vooral het aansporen van ouders negatief uitpakt. Tenslotte is nagegaan of er verschillen zijn vast te stellen voor de invloeden van deze kenmerken voor jongens en meisjes. Deze differentiatie blijkt slechts in beperkte mate op te treden. Meisjes uit arbeidersgezinnen blijken het iets beter te doen dan hun mannelijke tegenpolen, terwijl ook voor meisjes het aansporen van vader iets minder negatief uitwerkt. Jongens, daarentegen, blijken meer te profiteren van het boekenaanbod thuis.

Wat moeten we nu concluderen uit deze analyses op het VOCL-89 cohort? Ten eerste zijn er de grote effecten van de sociaal-economische achtergrond. Deze zijn slechts in geringe mate te duiden in termen van een cultureel of emotioneel hoogwaardig thuis klimaat. Hoe moeten we die dan wel interpreteren? De overblijvende effecten van ouderlijke opleiding kunnen natuurlijk deels te verklaren zijn vanuit een genetische overerving van aanleg, maar tevens, en voor sociaal-wetenschappers aantrekkelijker, door ongemeten aspecten van het ouderlijk gezin (financiële positie, vroege socialisatie en aspiraties). Ten tweede lijkt de relevantie van cultureel kapitaal hier betrekkelijk klein. Wat kan daarvan de reden zijn? Het kan zo zijn dat in eerder cross-sectioneel onderzoek door de methode van respondent-rapportage en ver terug-grijpend retrospectief meten van cultureel kapitaal het effect enigszins is overschat. Nader onderzoek, waarin zowel ouderlijke rapportage als respondent-rapportage worden betrokken, moet hierover duidelijkheid verschaffen. Meer inhoudelijk, zou men het leesgedrag van de ouders genuanceerder wensen te meten, omdat het lezen van veel romantische boeken in de huidige opzet is gelijkgeschakeld met het lezen van veel literatuur. Een derde conclusie kan zijn dat de ouder-kind communicatie over schoolonderwerpen belangrijk is. Hierbij dienen we echter meer zekerheid te krijgen over de causale volgorde van het fenomeen. In een longitudinaal design moet aannemelijk worden gemaakt dat ouderlijke schoolbetrokkenheid daadwerkelijk vooraf gaat aan de schoolprestaties van hun kinderen. Als vierde conclusie is vast te stellen dat demografische aspecten van het ouderlijke gezin belangrijk zijn. Deze aspecten hebben eerder vooral in Amerikaans onderzoek onder de aandacht gestaan. Zeker gegeven het feit dat deze aspecten tijdens de schoolloopbaan kunnen veranderen (scheiding, overlijden, geboorte), hetgeen de verklarende kracht van deze aspecten nog zou vergroten, is aandacht voor demografische kenmerken in onderzoek naar schoolprestaties zeer gewenst. Tenslotte laten onze analyses duidelijk zien dat de verschillen tussen meisjes en jongens als het gaat om de hier gepresenteerde inhoudelijke kenmerken vrijwel afwezig zijn. Wel overduidelijk is dat meisjes het anno 1997 beter op school doen dan jongens, in ieder geval waar het gaat om het behalen van een hoog diploma.

NOTEN

1. Als belangrijkste reden voor scholen om te weigeren, worden de problemen samenhangend met fusies genoemd (CBS 1991).
2. Het middelbaar beroepsonderwijs is tussen mavo en havo geplaatst op basis van de bi-variate samenhang met het opleidingsniveau van de ouders en de sociaal-culturele en sociaal-emotionele kenmerken. Zoals uit Tabel 1 blijkt lijken leerlingen met een afgeronde middelbare beroepsopleiding meer op leerlingen met een mavo-diploma dan op leerlingen met een havo-diploma.
3. Om interpretatie te vergemakkelijken is in de multi-variate regressie-analyse de exacte leeftijd gedeeld door tien, wat betekent dat men de ongestandaardiseerde coëfficiënten in termen van 10-jaarsleeftijdsgroepen moet interpreteren..

LITERATUUR

- Aschaffenburg, K. & I. Maas. 1997. Cultural and Educational Careers. The Dynamics of Social Reproduction, *American Sociological Review*, 62:573-587.
- Amato, P.R. 1993. Children's Adjustment to Divorce: Theories, Hypotheses, and Empirical Support, *Journal of Marriage and the Family*, 55:23-38.
- Astone, N.M. & S.S. McLanahan. 1991. Family Structure, Parental Practices and High School Completion, *American Sociological Review*, 56:309-320.
- Bakker, B.F.M. & P.G.J. Cremers. 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.
- Bosker, R.J. & H.P.J.M. Dekkers. 1994. School- en sekseverschillen in vakkenkeuzen in het VWO, *Tijdschrift voor OnderwijsResearch*, 19:214-226.
- Bourdieu, P. 1984. *Distinction*, Harvard University Press: Cambridge Mass.
- Bourdieu, P. & J.C. Passeron. 1977. *Reproduction in Education, Society and Culture*, Sage: London.
- Brands, J. 1992. *Die hoeft nooit meer wat te leren*, SUN: Nijmegen.
- C.B.S. (Centraal Bureau voor de Statistiek). 1991. *Schoolloopbaan en achtergrond van leerlingen: deel I; cohort 1989*, Staatsuitgeverij: Den Haag.
- Coleman, J.S. 1988. Social Capital in the Creation of Human Capital, *American Journal of Sociology*, 94:S95-S120.
- DiMaggio, P. 1982. Cultural Capital and School Success, *American Sociological Review*, 47:189-201.
- Downey, D.B. 1995. When Bigger is not Better: Family Size, Parental Resources and Children's Educational Performance, *American Sociological Review*, 60:746-761.
- Dronkers, J. 1990. De ontwikkelingen in het schoolloopbaanonderzoek: een terugblik op een decennium, *Tijdschrift voor OnderwijsResearch*, 15:8-22.
- Eijck, K. van 1996. *Family and Opportunity*, Tilburg University Press: Tilburg.
- Farkas, G. 1996. *Human Capital or Cultural Capital? Ethnicity and Poverty Groups in an Urban School District*, Aldine de Gruyter: New York.
- Furstenberg, F.F. & M.E. Hughes. 1995. Social Capital and Successful Development among At-Risk Youth, *Journal of Marriage and the Family*, 57:580-592.
- Graaf, N.D., P.M. de Graaf & G. Kraaykamp. 2000. Parental Cultural Capital and Educational Attainment in the Netherlands: a Refinement of the Cultural Capital Perspective, *Sociology of Education*, 73: (forthcoming).
- Graaf, P.M. de. 1987. *De invloed van financiële en culturele hulpbronnen in onderwijsloopbanen*, ITS: Nijmegen.
- Sui-Chu, E. Ho & J.D. Willms. 1996. Effects of Parental Involvement on Eight-Grade Achievement, *Sociology of Education*, 69:126-141.
- Jong, U. de. 1985. Vrouwen en onderwijs, in: J.L. Peschar & A.A. Wesselingh (red.), *Onderwijssociologie*, Wolters-Noordhoff: Groningen, 391-416.
- Kalmijn, M. & G. Kraaykamp. 1996. Race, Cultural Capital and Schooling, *Sociology of Education*, 69:22-34.
- Kraaykamp, G. & K. Dijkstra. 1999. Preferences in Leisure Time Book Reading; a Study on the Social Differentiation in Book Reading for the Netherlands, *Poetics*, 26:203-234.
- Leseman, P. 1994. Socio-cultural Determinants of Literacy Development, in L. Verhoeven (ed.), *Functional Literacy*, John Benjamins Publishing Company: Amsterdam/Philadelphia, 163-184.
- Lamont, M. & A. Lareau. 1988. Cultural Capital: Allusions, Gaps and Glissandos in recent Theoretical Developments, *Sociological Theory*, 6:153-168.
- Lareau, A. & E. Horvat, 1999. Moments of Social Inclusion and Exclusion: Race, Class and Cultural Capital in Family-School Relationships, *Sociology of Education*, 72:37-53.
- Mare, R.D. & M-S Tzeng. 1989. Fathers' Ages and the Social Stratification of Sons, *American Journal of Sociology*, 95:108-131.
- Meijnen, G.W. 1985. Socialisering: taal, talent en geschiktheid, in: J.L. Peschar & A.A. Wesselingh (red.), *Onderwijssociologie*, Wolters-Noordhoff: Groningen, 161-184.
- Mercy, J.A. & L.C. Steelman. 1982. Familial Influence on Intellectual Attainment of Children, *American Sociological Review*, 47:532-542.
- Milne, A.N., D.E. Myers, A.S. Rosenthal & A. Ginsburg. 1986. Single Parents, Working Mothers, and the Educational Achievement of School Children, *Sociology of Education*, 59:125-139.
- Muller, C. 1995. Maternal Employment, Parental Involvement, and Mathematics Achievement among Adolescents, *Journal of Marriage and the Family*, 57:85-100.
- Niehof, J. 1997. *Resources and Social Reproduction*, Thesis Publishers: Amsterdam.

- Rumberger, R.R., R. Ghatak, G. Poulos, P.L. Ritter & S.M. Dornbusch. 1990. Family Influences on Dropout Behavior in one California High School, *Sociology of Education*, 63:283-299.
- Rupp J.C.C. & A.A. Wesselingh. 1989. Onderwijs en maatschappelijke ongelijkheid: een overzicht van 25 jaar theorievorming en onderzoek in Nederland, *Tijdschrift voor OnderwijsResearch*, 14:237-258.
- Schneider, B. & J. Coleman. 1993. *Parents, Their Children, and Schools*, Westview Press: Boulder.
- Steelman, L. & B. Powell. 1989. Acquiring Capital for College: the Constraints of Family Configuration, *American Sociological Review*, 54:844-855.
- Teachman J.D., K. Paasch & K. Carver. 1996. Social Capital and Dropping Out of School Early, *Journal of Marriage and the Family*, 58:773-783.
- Tesser, P. 1986. *Sociale herkomst en schoolloopbanen in het voortgezet onderwijs*, ITS: Nijmegen.
- Tesser, P. & Mulder, L. 1990. Cultureel kapitaal en schoolprestaties in het basisonderwijs, in: C. Klaassen & P. Jungbluth (red.), *OnderwijsResearchDagen 1990*, ITS: Nijmegen, 39-51.
- Veenstra, R. 1999. *Leerlingen-klassen-scholen. Prestaties en vorderingen van leerlingen in het voortgezet onderwijs*, Thela Thesis: Amsterdam.
- Vrooman, J. & J. Dronkers. 1986. Changing Educational Attainment Processes. Some Evidence from the Netherlands, *Sociology of Education*, 59:69-78.
- Wolbers, M. 1998. *Diploma-inflatie en verdringing op de arbeidsmarkt*, Thesis Publishers: Amsterdam.
- Wolbers, M. & P. de Graaf. 1996. Onderwijsexpansie of meritocratisering? Een verklaring voor de afgenomen samenhang tussen het ouderlijk sociaal milieu en het hoogst bereikte opleidingsniveau van kinderen, *Tijdschrift voor OnderwijsResearch*, 12:1-14.

Samenwerkend leren in computerondersteund rollenspel

G. A. Holsbrink-Engels

Faculteit der Toegepaste Onderwijskunde, Afdeling Instructietechnologie, Postbus 217,
7500 AE Enschede, e-mail: geralien@holsbrink.demon.nl

ABSTRACT

Collaborative learning is a recent theme in the field of educational technology. The acquisition of knowledge and skills is seen as a result of interactions between actors. In the domain of interpersonal skills training, collaborative learning exists for a long time but is not described so. Currently, role playing is seen as a complex learning environment for novices to develop interpersonal skills because of high cognitive load of solving social-communicative problems and the performance in a play. Simplification of the current collaborative learning environment entails the development of expertise. Another collaborative learning environment, computer-based role playing, is proposed. Two studies are reported with this new developed collaborative learning environment for solving social-communicative problems. The instructional strategies, the use of a conversational model and providing opportunities for reflection, have positive effects on and facilitate learning.

1. INLEIDING EN ACHTERGROND

Samenwerkend leren is op dit moment een van de kernthema's in de onderwijskunde (de Jong, 1999). Samenwerkend leren omvat die (onderwijs-) leersituaties waarin de lerenden in interactie met een of meerdere actoren onder gedeelde verantwoordelijkheid een leer- of probleemtaak uitvoeren met een gemeenschappelijk doel of product dat door iedere participant wordt nagestreefd (Erkens, 1997). De lerenden staan centraal en het constructief vergaren van kennis en vaardigheden wordt benadrukt.

In dit artikel wordt verslag gedaan van een tweetal onderzoeken naar het oplossen van sociaal-communicatieve problemen in een op samenwerkend leren gerichte leeromgeving. Is het mogelijk om gespreksvaardigheden te verwerven met behulp van computerondersteund netwerk dat het samenwerkend leren ondersteunt en zo ja hoe dan? Wat zijn de leereffecten van samenwerkend leren waarbij gebruik wordt gemaakt van computerondersteunde rollenspelen? Is de aan- of afwezigheid van een gespreksmodel van invloed op het effect van de leeromgeving? En leiden mogelijkheden tot reflectie, die in de leeromgeving zijn ingebouwd, tot grotere leereffecten? Om deze vragen te beantwoorden wordt in het eerste onderzoek het effect van samenwerkend leren in computerondersteund rollenspel op het verwerven van gespreksvaardigheden onderzocht. In het tweede onderzoek staat de effectiviteit van twee instructiestrategieën op het verwerven van gespreksvaardigheden centraal.

Uitgangspunt is dat voor de verwerving van gespreksvaardigheden, lerenden sociaal-communicatieve problemen moeten leren oplossen. Door verschillende sociaal-communicatieve problemen op te lossen, wordt kennis verworven omtrent de probleemsituatie en wordt de probleemoplossingsprocedure geleerd. De vaardigheid is geleerd wanneer de (her)geconstrueerde kennis kan worden toegepast.

Het trainen van gespreksvaardigheden is van belang voor studenten in verschillende disciplines ter voorbereiding op hun professionele rol. Naast expert kennis van het domein, gaat het erom dat toekomstige professionals effectief met anderen kunnen communiceren en samenwerken. Sinds de tachtiger jaren wordt het trainen van gespreksvaardigheden opgenomen als vast onderdeel van het curriculum van diverse (beroeps-)opleidingen. Verschillende vormen van samenwerkend leren worden ingezet om studenten te trainen omtrent gesprekken in organisaties.

Er wordt ingegaan op *wat* er onderwezen dient te worden om gespreksvaardigheden te verwerven: het oplossen van sociaal-communicatieve problemen. Verder wordt een beschrijving gegeven van *hoe* samenwerkend leren binnen dit domein gestalte kan krijgen. De ontwikkelde leeromgeving wordt beschreven. Daarna wordt gerapporteerd over een tweetal onderzoeken met de leeromgeving. De resultaten worden gepresenteerd. Tenslotte wordt in de discussie nader ingegaan op de implicaties voor onderzoek en onderwijs.

2. HET OPLOSSEN VAN SOCIAAL-COMMUNICATIEVE PROBLEMEN

Naar probleemoplossingsprocessen in andere domeinen dan gespreksvoering is uitvoerig onderzoek gedaan (onder andere, Anderson, 1982; 1983; 1987; 1993; Chase & Simon, 1973; De Groot, 1965; Elshout, 1987; Greeno, 1978; 1980; Newell & Simon, 1972; Schoenfeld, 1985; Vos, 1987a; 1987b; 1990). De cognitief psychologische benadering van probleemoplossen wordt bruikbaar geacht en als uitgangspunt gekozen voor de onderhavige onderzoeken. Terzijde wordt opgemerkt dat tot op heden weinig onderzoek is verricht naar sociaal-communicatieve probleemoplossingsprocessen.

Om gespreksvaardigheden te verwerven en expertise te ontwikkelen, dienen studenten sociaal-communicatieve problemen op te lossen. Om sociaal-communicatieve problemen te kunnen oplossen, moet kennis worden verworven omtrent de probleemsituatie (bijvoorbeeld kenmerken van een slecht-nieuwsgesprek opdat deze gespreksituatie later wordt herkend) en moet de probleemoplossingsprocedure worden geleerd. Deze probleemoplossingsprocedure bestaat uit een aantal stappen om een sociaal-communicatief probleem op te lossen.

De studenten verwerven een gespreksvaardigheid door samen actief de benodigde stappen te construeren om het doel te kunnen bereiken. De gegeven informatie omtrent de probleemsituatie en het oplossen van problemen leiden tot de (her-)constructie van sociaal-communicatieve kennis en probleemoplossingsprocedures. De gespreksvaardigheden worden verworven door verschillende sociaal-communicatieve problemen op te lossen waarbij het probleemoplossingsgedrag herhaaldelijk wordt geoefend.

Het probleemoplossingsproces kent de volgende onderdelen (Holsbrink – Engels, 1998): (a) de herkenning van het sociaal-communicatieve probleem, (b) het beschrijven en exploreren van het probleem, (c) het oplossen van het probleem en (d) evaluatie van de gekozen oplossing. Ondanks het feit dat elk sociaal-communicatief probleem unieke elementen in zich draagt, zijn ook overeenkomsten te onderscheiden.

Onderkenning van het sociaal-communicatieve probleem. De eerste fase van het probleemoplossingsproces is de bewustwording van het probleem. Er is alleen sprake van een probleem als het probleem wordt onderkend. Er loopt iets fout. Een obstakel dient te worden opgemerkt. De vage notie wordt gevoeld dat er iets aan gedaan moet worden. Vervolgens wordt de urgentie van het probleem afgewogen om te bepalen of het zinnig is om aandacht aan het probleem te schenken. Immers niet alle sociaal-communicatieve problemen zijn de moeite waard om de aandacht aan te schenken.

Het beschrijven en exploreren van het probleem. In deze fase staan centraal het exploreren en het doordringen tot de kern van het probleem. De kritische factoren van het probleem worden geïsoleerd en de barrières worden beschreven. Soms wordt een probleem hergedefinieerd op basis van de verzamelde informatie.

Oplossen van het probleem. Deze fase van het probleemoplossingsproces kent vier verschillende activiteiten: (a) doel formuleren, (b) opsommen van de meest relevante oplossingen, (c) keuze van de meest bevredigende oplossing en (d) uitvoeren van de stappen van de procedure. Het doel, kan zijn het oplossen van het probleem of het verwijderen van een obstakel, het bereiken van overeenstemming of voorkoming van een bepaalde activiteit.

Evaluatie van de gekozen oplossing. In de laatste fase is nadrukkelijk plaats voor evaluatie- en terugkoppelingsactiviteiten en zo nodig verbeteren van een oplossing waardoor het oplossen van sociaal-communicatieve problemen een cyclisch karakter krijgt of kan krijgen. Het probleemoplossingsproces wordt pas beëindigd als voor alle betrokkenen een bevredigend resultaat is bereikt.

Het probleemoplossingsproces van sociaal-communicatieve problemen kent een affectieve en een cognitieve of intellectuele inhoud. Affecties kunnen grote invloed hebben op de intellectuele inhoud van een gesprek. Het beschreven probleemoplossingsproces is geenszins een lineair proces. Regelmatig worden sociaal-communicatieve problemen in een andere volgorde opgelost als hierboven beschreven. Bijvoorbeeld, de oplossingen worden al gegenereerd alvorens de kern van het probleem is vastgesteld.

3. BESCHRIJVING VAN DE OP SAMENWERKEND LEREN GERICHTE LEEROMGEVING

3.1 Algemene typering

Nadat is nagedacht over *wat* er onderwezen dient te worden om gespreksvaardigheden te verwerven, kan de vraag beantwoord worden *hoe* samenwerkend leren binnen dit domein gestalte kan krijgen. In welke op samenwerkend leren gerichte leeromgeving kunnen studenten effectief gespreksvaardigheden verwerven? Hoe kan het oplossen van sociaal-communicatieve problemen geoefend worden en kan expertise worden ontwikkeld?

Samenwerkend leren

In het algemeen wordt van samenwerkend leren gesproken als studenten in kleine groepen samenwerken aan de uitvoering van een gemeenschappelijk gestelde leer- of probleemtaak. Een voorwaarde hierbij is dat de groepen zo klein zijn dat in principe elk groepslid kan participeren en invloed kan uitoefenen op de uitvoering van de taak (Cohen, 1994). Er is sprake van een positieve onderlinge afhankelijkheid tussen studenten tijdens het samenwerken om gezamenlijke doelen te bereiken (Johnson & Johnson, 1989). De samenwerking wordt beloond op basis van de kwaliteit of de kwantiteit van het produkt van de groep overeenkomstig een aantal vastgelegde maatstaven. Deze maatstaven geven de mate van beheersing weer. Het hebben van een gemeenschappelijk belang is kenmerkend voor samenwerkend leren.

Al sinds de tweede wereldoorlog (Lippitt, 1943; French, 1945; Argyris, 1951; Corsini, 1960) krijgt het samenwerkend leren in het domein van gespreksvoering gestalte in het rollenspel. Het rollenspel is *de* instructiemethode voor het verwerven van gespreksvaardigheden (Argyle, 1984). Het rollenspel voorziet in een op samenwerkend leren gerichte leeromgeving waarin de lerende vanuit verschillende rollen interacteert met een of meerdere actoren om een gegeven sociaal-communicatief probleem op te lossen. Het leerproces wordt verlevendigd en verrijkt door analoge situaties aan de werkelijkheid aan te bieden omdat een "echte" gesprekssituatie vaak niet beschikbaar, te duur of te veel risico's in zich draagt. Realistische sociale contexten worden aangeboden doordat essentiële delen van de gespreksvoerings-werkelijkheid worden gerepliceerd. Een case, een beschrijving van een situatie waarmee een professional in de praktijk wordt geconfronteerd, wordt gesimuleerd. De interactie in het rollenspel is doelgericht en is te typeren als een dynamisch en transactioneel proces. De studenten hebben geen volledige controle over de gesprekssituatie. De tegenspelers komen met onverwachte responsen. De gesprekssituatie verandert ten gevolge van de boodschappen die uitgewisseld worden.

Complexiteit van samenwerkend leren

Sinds het begin van de jaren negentig is in de onderwijskunde een ontwikkeling gaande naar onderzoek en onderwijs dat meer constructief, gesitueerd en samenwerkend van karakter is. Deze ontwikkeling is beloftevol. Toch kleven aan deze ontwikkeling voor het domein van de gespreksvoering, waar samenwerkend leren al lange tijd wordt toegepast maar niet zo wordt beschreven, enkele bezwaren. Het gangbare rollenspel wordt voor beginners als complex beschouwd. Het leren van gespreksvaardigheden wordt hierdoor negatief beïnvloed. De complexiteit van het rollenspel is gelegen in een tweetal factoren, namelijk (a) een hoge cognitieve belasting en (b) het niet kunnen bieden van voldoende oefening.

Hoge cognitieve belasting ontstaat omdat sociaal-communicatieve problemen slecht gedefinieerde problemen zijn en die brengen een hogere cognitieve belasting mee dan goed gedefinieerde problemen (o.a. Greeno, 1978; Simon, 1973; Voss, 1990). De definitieve formulering van het op te lossen sociaal-communicatieve probleem verandert gedurende het probleemoplossingsproces als nieuwe informatie wordt verkregen door de probleemoplosser. Er bestaat bovendien niet *een* juiste manier om informatie te verzamelen. Beslissingen moeten genomen worden zonder dat definitieve kennis beschikbaar is. Zekerheid over de juistheid van de beslissing kan niet gegeven worden. Meerdere oplossingen van het probleem zijn denkbaar. De taak is niet transparant. (Kass, Burke, Blevis & Williamson, 1993/1994).

Daarnaast leidt het direct uitvoeren van de probleemoplossingsstappen tijdens de communicatie en de uitvoering in een dialoog tot een hogere cognitieve belasting. De beginners moeten de probleemoplossingsstrategieën voor een bepaalde situatie onmiddellijk paraat hebben en direct toepassen.

Weinig oefening. Het rollenspel biedt beginners veel informatie en weinig oefening (Holsbrink - Engels, 1994). Beginners worden in het eerste rollenspel overweldigd met gegevens en zij ervaren hun eerste rollenspel als verwarrend (Kessels & Smit, 1990). Zij worden in het diepe gegooid met rol- en situatiebeschrijvingen. Zij moeten deze informatie ordenen, zich inleven in hun rol, en effectief optreden in het spel.

Tijdens trainingen is het vaak lastig- zeker bij grotere groepen- om iedereen aan de beurt te laten komen, laat staan twee of meer keer te kunnen oefenen. Beginners leren natuurlijk ook tijdens het observeren van andere rollenspelen maar dit leerproces verloopt toch anders dan "het zelf doen". Veel oefenen is een vereiste om tot een goede uitvoering van de vaardigheid te komen. Een enkel rollenspel leidt niet tot het verwerven van gespreksvaardigheden (Georges, 1989).

Op grond van de geschetste complexiteit wordt een vereenvoudiging van het samenwerkend leren voorgesteld voor het verwerven van gespreksvaardigheden.

3.2 Vereenvoudiging van het samenwerkend leren

Vanuit het perspectief van onderzoek en onderwijs dat meer constructief, gesitueerd en samenwerkend van karakter is, bestaat het doel van de op samenwerkend leren gerichte leeromgeving erin te beginnen met een vereenvoudigde, maar wel representatieve wijze van oplossen van sociaal-communicatieve problemen. In deze leeromgeving wordt het oplossen van sociaal-communicatieve problemen aangepast en vereenvoudigd opdat het leren wordt vergemakkelijkt en bevorderd.

Dit doel wordt in de eerste plaats bereikt door een tweetal instructiestrategieën als uitgangspunt te nemen, namelijk het *gespreksmodel* en door het bieden van *mogelijkheden tot reflectie* tijdens het spel. In de tweede plaats wordt een computernetwerk ingezet als het medium voor instructie.

Het gespreksmodel

Het gespreksmodel is een heuristiek die in de leeromgeving wordt aangeboden en is behulpzaam bij het oplossen van het sociaal-communicatieve probleem. Het is een leidraad, een beschrijving van een systematische probleemaanpak om te komen tot de oplossing van het sociaal-communicatieve probleem. De heuristiek specificiert het doel, de naam van de pro-

cedure en identificeert en sequentieert de oplossingsstappen (zie Tabel 1). Om het gespreksmodel voor de leeromgeving te kunnen vaststellen is gebruik gemaakt van een SPA-analyse (Mettes, Pilot & Roossink, 1981; Van Merriënboer, 1997). Deze taakanalyse-techniek is geschikt voor het beschrijven van een probleemoplossingsstrategie voor het oplossen van complexe problemen.

Uit onderwijskundig onderzoek in andere domeinen (o.a. Chi, Glaser & Farr, 1988; Mettes, Pilot & Roossink, 1981; Schoenfeld, 1979) is gebleken dat het raadzaam is om heuristische kennis te beschrijven en expliciet te onderwijzen.

Het bieden van reflectiemogelijkheden

Het bieden van reflectiemogelijkheden (zelfkritiek) geschiedt door (a) vermindering van de complexiteit van de probleemsituatie en door (b) het verminderen van de tijdsdruk die aanwezig is in werkelijke gesprekken.

De complexiteit van de probleemsituatie wordt gereduceerd door het weglaten van irrelevante en ook relevante kenmerken van een werkelijk gesprek. De werkelijkheidsgetrouwheid ("Fidelity") van de probleemoplossingssituatie representeert de authentieke probleemsituatie. Volgens Alessi (1988) bevordert een leersituatie met een maximale werkelijkheidsgetrouwheid van de authentieke situatie het leren niet. Een effectief instructieontwerp voor beginners begint met een lage werkelijkheidsgetrouwheid ("low fidelity"). Via enkele tussenstappen dient tegen het einde van het instructieproces een leersituatie met een hoge werkelijkheidsgetrouwheid ("high fidelity") aan te worden geboden. In het begin richt de aandacht zich op het leren van een of twee variabelen totdat, via stapsgewijze progressie, alle belangrijke variabelen onderwezen zijn.

De tijdsdruk wordt verminderd door het vrijmaken van tijd voor het schrijven en corrigeren van (gespreks-) boodschappen. Mogelijkheden voor reflectie worden bevorderd door meer tijd te schenken aan en te reflecteren op de inhoud van de te verzenden boodschap. In het gangbare rollenspel is geen tijd voor reflectie tijdens het gesprek omdat er binnen luttele seconden een reactie van de ander komt. Deze karakteristiek van het gangbare gesprek wordt door Winograd en Flores (1986) "Thrownness" genoemd. Het belang van het kunnen reflecteren in een leerproces is al lange tijd bekend. Recentelijk is er meer belangstelling voor het bieden van reflectiemogelijkheden tijdens het oefenen van vaardigheden. Dit proces van "reflectie-in-actie" (Schön, 1983, 1987) vormt een essentieel onderdeel in het leerproces (Boud & Walker, 1991).

Computernetwerk

De leeromgeving is ontwikkeld in HyperCard 2.1. Het omvat een netwerk voor communicatie tussen twee computers. De interactie tussen de spelers vindt plaats op twee met elkaar verbonden computers die fysiek gescheiden zijn opgesteld. Het proces van versturen en ontvangen van boodschappen geschiedt door het versturen van getypte teksten. De te versturen tekst van de ene speler verschijnt in het eigen tekstvenster en na verzending op het computerscherm van de andere speler in de daarvoor beschikbare tekstballon naast het hoofd van de fictieve gesprekspartner (op de foto op het computerscherm) en visa versa. Het verzenden van de boodschappen geschiedt door het klikken op een verzend-button. Synchrone (gelijktijdige) communicatie is hier een technische parameter alsmede een sociale regel. Degenen die spreekt verwacht dat degene die luistert wacht op zijn boodschap en de boodschap zal verwerken op het moment dat deze ontvangen is.

In de leeromgeving wordt het zien, horen en voelen van de ander onmogelijk gemaakt. Informatie omtrent fysieke verschijning, sekse, ras, stem en intonatie ontbreekt en indien gewenst kan anonimiteit worden gerealiseerd. De inhoud van het gesprek komt centraal te staan. De interacties over de inhoud beïnvloeden de cognitieve processen van de studenten voortdurend. Er is sprake van een collaboratieve situatie die volledig interactief is.

In de leeromgeving is het (zij het voor korte tijd) mogelijk om *tijdens het oefenen* van de

Tabel 1: Een gespreksmodel voor het slecht-nieuwsgesprek; zes stappen nader beschouwd aan de hand van een voorbeeld

Om het slechte nieuws mee te delen dient de protagonist:	Een voorbeeld
1. Klap uitdelen: <ul style="list-style-type: none"> • korte, betrokken inleiding • boodschap • STOP 	Peter: Brenda, je wilt me spreken? Brenda: Ja Peter, laten we hier gaan zitten. Eh, Peter, ik moet je iets vertellen wat je niet leuk zal vinden. Je aanvraag voor groepsmanager is afgewezen. (STOP) Peter: Wat? Brenda: De reden voor de afwijzing is dat de staf besloten heeft dat er iemand moet komen die erg goed geschoold is in de automatisering.
2. Klap opvangen: <ul style="list-style-type: none"> • luistervaardigheden, m.n. reflecties • herhalen slechte nieuws 	Peter: Momentje Brenda. Je wilt me vertellen dat ik geen geschikte groepsmanager ben, terwijl ik al tien jaar plaatsvervanger ben. Brenda: Peter, ik begrijp je verontwaardiging heel goed. Het besluit je af te wijzen is zeer overwogen genomen. Peter: Ja?
3. Argumenten noemen	Brenda: Maar, laat me uitleggen waarom dit besluit zo is genomen, wat de achtergrond is. Peter: Uhum. Brenda: Peter, het is van groot belang dat jouw afdeling geautomatiseerd wordt. Zolang we dit niet erkennen lopen we maandelijks achterstand op. We hebben een manager nodig met praktijkervaring, met veel ervaring op het gebied van automatisering. Nu, deze toegevoegde taak is belangrijker dan ancieniteit of bekendheid met de huidige functie. Er is gekozen voor Sylva van der Meer van de hoofdvesting.
4. Klap opvangen	Peter: Ja, maar dit kan je niet serieus menen! Het was zo goed als zeker dat ik de baan zou krijgen. Brenda: Ik begrijp je reactie heel goed omdat je een goede kandidaat voor de baan was. Het is waar. Ikzelf heb je voorgedragen, maar de staf heeft besloten iemand anders te nemen. Peter: Ja, en sorry dat ik het je lastig heb gemaakt Peter, dus pak je huidige baan op gemotiveerde wijze weer op. Nou, ik heb er genoeg van en ik doe niets meer. Zo! Brenda: Ja Peter, ik begrijp je reactie. Je bent boos en teleurgesteld. Je toekomst ligt op de helling. Je uit je twijfels over je motivatie om door te gaan. Peter: Ja, ja, ja, ben je verrast? Brenda: Maar ik hoop echt dat je niet gaat prullen of gefrustreerd of ongemotiveerd raakt. Peter, onze samenwerking is hecht en goed. Ik wil deze voortzetten ondanks deze tegenspoed.
5. Probleem oplossen	Brenda: Peter, het is een feit dat je geen groepsmanager wordt. Je houdt je huidige baan als vervanger. Dat is de nieuwe situatie voor jou, voor mij en voor Sylva. Peter: Welke Sylva? Brenda: Sylva van der Meer, de nieuwe groepsmanager. Het is voor ons alle drie een nieuwe situatie. Sylva van der Meer moet worden ingewerkt en daarvoor is jouw assis-

tentie essentieel. Kunnen we praten over hoe we dit gaan aanpakken?

Peter: Je vraagt teveel van mij, Brenda. Denk eens aan de reacties van mijn collega's. Nu zal mij steeds gezegd worden dat ik het niet geworden ben. Realiseer je dat?

6. Concrete afspraken maken

Brenda: Ja, je hebt gelijk. Dit is een belangrijk punt. Nu hebben we twee punten om over te praten; de inwerkperiode van Sylvia en de omgang met onze collega's. Ik wil hier aan het eind van de week over praten, vrijdag om 10.00 uur?

Peter: Uhm, goed.

vaardigheden te reflecteren op de inhoud van de boodschap waardoor de tijdsdruk wordt verminderd. Verder kunnen de studenten het gesprek terug lezen, de rolbeschrijving herlezen en reflectiebevorderende vragenlijsten invullen voor- en na het spelen van een rollenspel.

3.3 Inhoudsdomein

In het hierna te beschrijven onderzoek is het verwerven van de bekwaamheid in het voeren van een slecht-nieuwsgesprek gekozen als een representant van een van de te verwerven gespreksvaardigheden binnen het curriculum van diverse (beroeps-) opleidingen. Het is een tweegesprek tussen een brenger en een ontvanger van het slechte nieuws. Het gesprek kent twee doelen. Het eerste doel is het geven van informatie (slecht nieuws) aan de betrokkene. Het tweede doel is verandering en / of beïnvloeding van de ontvanger, namelijk het helpen bij het verwerken van het slechte nieuws en eventueel bij het vinden van oplossingen voor het ontstane probleem.

Een aanname is dat studenten eerst voorkennis over het slecht-nieuwsgesprek (wat is het?) dienen te verwerven alvorens zij de vaardigheid kunnen oefenen en verwerven in de op samenwerkend leren gerichte leeromgeving.

3.4 Leertaken en rollen in de leeromgeving

In de leeromgeving worden studenten in tweetallen geconfronteerd met het leren oplossen van sociaal-communicatieve problemen. De taken omvatten het afwisselend spelen van een tweetal rollen binnen gesimuleerde casussen. In de casussen wordt informatie gepresenteerd over op te lossen problemen in gespreksituaties. Het zijn open probleemsituaties waarin een dilemma wordt geschetst. De oplossingen worden niet gegeven.

De probleemoplossers zijn studenten die de protagonisten-rollen spelen. Zij hebben tot taak het sociaal-communicatieve probleem op te lossen. De "significante anderen" zijn studenten die de antagonist-rollen spelen. Zij verstrekken de protagonisten nieuwe informatie en zorgen voor (verbale) weerstand tegen geformuleerde ideeën en wensen. Deze rollen wisselen bij elk nieuw aangeboden probleem, dus per volgend rollenspel.

In de leeromgeving wordt het gespreksmodel aangeboden om behulpzaam te zijn bij het oplossen van sociaal-communicatieve problemen. Tijdens het formuleren van het gespreksmodel zijn een zestal acties geïdentificeerd voor de probleemoplossers. Tabel 1 beschrijft het gespreksmodel voor het slecht-nieuwsgesprek geïllustreerd door een voorbeeld.

De leeromgeving biedt vier verschillende sociaal-communicatieve oefenproblemen die opgelost dienen te worden in computerondersteunde rollenspelen. De moeilijkheidsgraad van de op te lossen problemen neemt toe door expliciet wijzigingen aan te brengen in (a) de bekendheid met de probleemsituatie en (b) het type slecht nieuws. Verondersteld wordt dat een probleem in een bekende situatie eenvoudiger oplosbaar zal zijn dan een probleem in een onbekende situatie. Daarnaast wordt verondersteld dat slecht nieuws veroorzaakt door een externe oorzaak eenvoudiger oplosbaar zal zijn dan als die veroorzaakt wordt door een intern (persoonlijke) disfunctioneren. Reorganisatie is een voorbeeld van een externe oorzaak, onvoldoende computerkennis en vaardigheden is een voorbeeld van een interne oorzaak. In dit laat-

ste geval voelt de ontvanger van het slechte nieuws zichzelf meer verantwoordelijk en wil zich rechtvaardigen. Dit leidt tot meer discussie, meer emoties en een grotere kans op conflicten. Tabel 2 geeft een overzicht van de aangeboden sociaal-communicatieve oefenproblemen. Het oplossen van een sociaal-communicatief probleem varieert tussen de 15 tot 45 minuten afhankelijk van de moeilijkheidsgraad en de gekozen probleemaanpak.

Tabel 2: Overzicht van de op te lossen sociaal-communicatieve oefenproblemen

Kern van het slechte nieuws	Reden(en)
1. geen reis naar Australië	loting bepaalt uitsluiting
2. verslag is afgekeurd	onvoldoende kwaliteit van verslaglegging
3. accepteren van een andere (lagere) positie	reorganisatie (extern)
4. accepteren van een andere (lagere) positie	persoonlijk disfunctioneren (intern)

3.5 Sequentiering van de inhoud en controle van de interacties

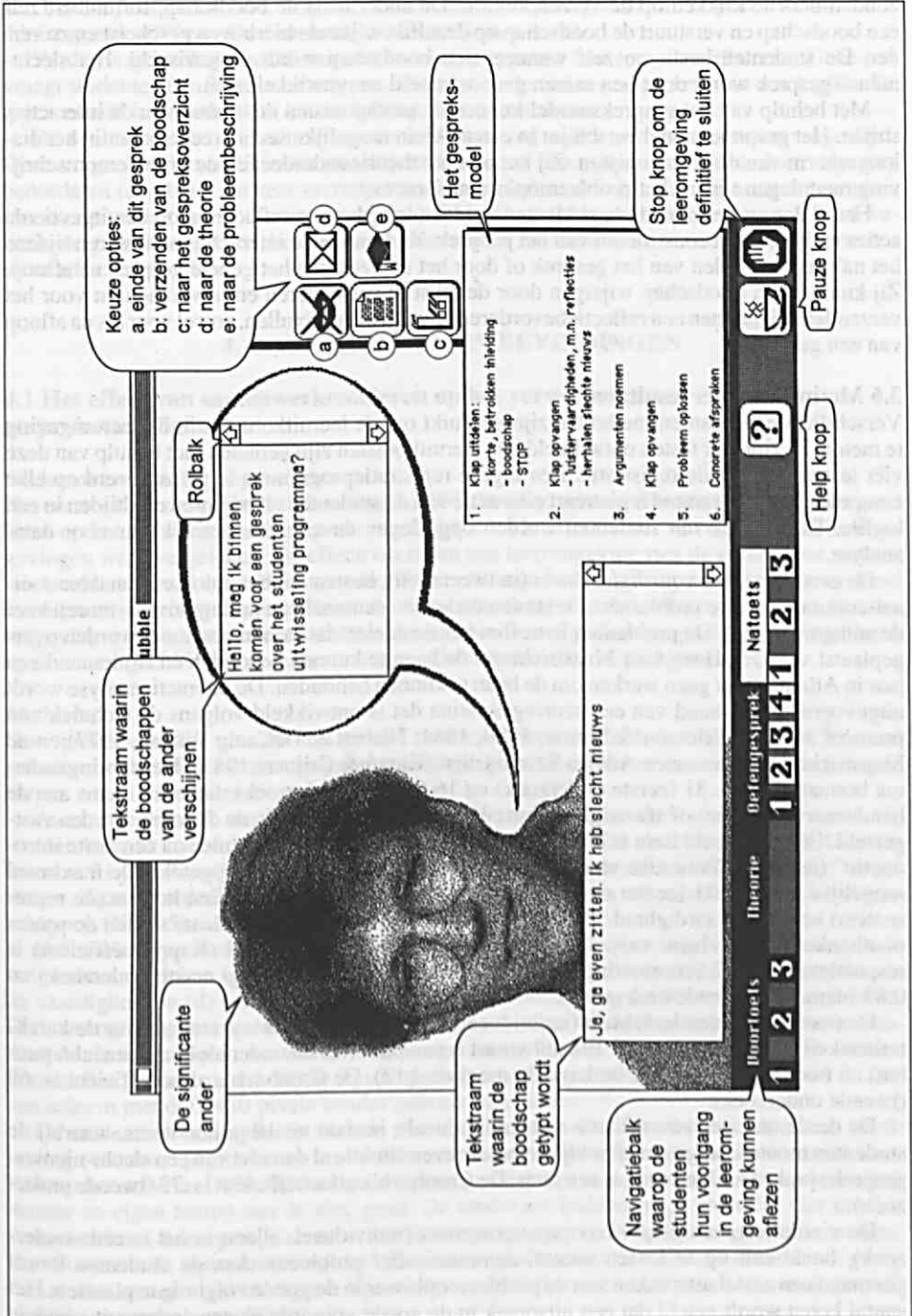
De aanname dat studenten eerst de vereiste voorkennis over het slecht-nieuwsgesprek dienen te verwerven alvorens zij de vaardigheid kunnen oefenen in de op samenwerkend leren gerichte leeromgeving is reeds ter sprake gekomen. Het samenwerkend leren vindt in de computerondersteunde rollenspelen plaats. De dialoogschermen vormen de kern van het samenwerkend leren. In de leeromgeving wordt daaromheen eerst de voorkennis aangeboden en wordt vanwege experimentele doeleinden een voor- en nameting verricht.

Door op een button van het openingsscherm te klikken, bereiken de studenten het algemene introductiescherm van de leeromgeving. Na de voormeting komen de studenten in het onderdeel "Theorie over het slecht-nieuwsgesprek". De gewenste voorkennis wordt hier gepresenteerd op pagina's met een fixeerde omvang. Het theorieonderdeel eindigt met een scherm waarop de studenten worden verzocht om de videopresentatie te gaan bekijken. Gedurende tien minuten wordt een slecht en goed voorbeeld van een slecht-nieuwsgesprek getoond op een televisiescherm met koptelefoon op en met handmatige bediening (alleen in het eerste onderzoek). In het tweede onderzoek vindt deze fase voor alle studenten plaats in een collegezaal. Onder leiding van de proefleiders bestuderen studenten de theorie uit een syllabus (8 pagina's) en bekijken zij samen de video. Daarna vond het samenwerkend leren plaats in de computerruimtes.

In de computerondersteunde rollenspelen wordt de taak vastgelegd door het op te lossen sociaal-communicatieve probleem te presenteren voorgaand aan het dialoogscherm. De probleemomschrijving bestaat uit (a) een introductiescherm met een foto van het gebouw waar het gesprek plaats vindt, (b) een tekstschermbeschrijving en (c) een tekstschermbeschrijving met een rolbeschrijving.

Het dialoogscherm legt de controle over de communicatie neer bij de studenten (die de prota- en antagonistrollen vervullen). Het dialoogscherm ondersteunt het samenwerkend leren. Voor een voorbeeld van een dialoogscherm, afkomstig uit het eerste rollenspel, wordt verwezen naar Figuur 1. Het samenwerkend leren start op het moment dat de protagonist de antagonist ontmoet. De antagonist verschijnt op het dialoogscherm door middel van een foto. De verschillende dialoogschermen in de leeromgeving kennen een consistent ontwerp. Per computerondersteund rollenspel wordt een andere foto gepresenteerd opdat verschillende slecht-nieuwsgesprekken niet telkens aan een zelfde persoon worden gekoppeld.

Het dialoogscherm biedt mogelijkheden voor opeenvolgende interacties en onbeperkte mogelijkheden tot het uitwisselen van boodschappen. De boodschappen verschijnen in de tekstballonnen naast het hoofd. De te verzenden boodschappen worden ingetypt in het daarvoor bestemde tekstveld, links onder en net boven de navigatiebalk. De ingetypte tekst verschijnt direct in dit raam. Wanneer de boodschap volledig is, kan de boodschap worden ver-



Figuur 1: Het dialoogscherm met het gespreksmodel voor het eerste oefengesprek.

zonden door te klikken op de verzendbutton. De ander leest de boodschap, formuleert zelf een boodschap en verstuurt de boodschap op dezelfde wijze als hierboven geschetst en zo verder. De studenten beslissen zelf wanneer een boodschap wordt uitgewisseld. Het slecht-nieuwsgesprek wordt door hen samen geconstrueerd en ontwikkelt zich.

Met behulp van het gespreksmodel kunnen de protagonisten de inhoud van de interacties sturen. Het gespreksmodel verschijnt in een zo klein mogelijke sectie, rechtsonder in het dialoogschermbaan van de protagonisten. Zij kunnen het theorieonderdeel en de probleemomschrijving raadplegen tijdens het probleemoplossingsproces.

Het dialoogschermbaan biedt de studenten voldoende tijd om te reflecteren op de uitgevoerde acties en het samen construeren van het gesprek. Zij kunnen de interacties analyseren tijdens het nalezen van delen van het gesprek of door het nalezen van het gehele gesprek na afloop. Zij kunnen een boodschap wijzigen door de tekst te verwijderen en te veranderen voor het verzenden. Zij kunnen een reflectiebevorderende vragenlijst invullen, zowel voor als na afloop van een gesprek.

3.6 Meting van leerresultaten

Verschillende assessmentmethoden zijn gebruikt om de leeruitkomsten in de leeromgeving te meten. Er zijn vier testen ontwikkeld. De leeruitkomsten zijn gemeten met behulp van deze vier testen met logfile registratie. Een logfile registratieprogramma is geïnstalleerd op elke computer. Dit programma registreert elke actie van de studenten en begin- en eindtijden in een logfile. De logfiles van studenten worden opgeslagen en zijn dus beschikbaar voor data-analyse.

De eerste test, een *vaardigheidstest* (in tweetallen), bestaat uit het oplossen van twee sociaal-communicatieve problemen. De studenten spelen eenmaal een protagonisten- en een keer de antagonistrol. De problemen betreffen het meedelen dat (a) iemand moet worden overgeplaatst van Den Haag naar Maastricht om de baan te kunnen behouden en (b) iemand een jaar in Atlanta moet gaan werken om de baan te kunnen behouden. De interactieanalyse wordt uitgevoerd aan de hand van een scoringsschema dat is ontwikkeld volgens de techniek van protocol analyse (Ericsson & Simon, 1980, 1984; Nisbett & DeCamp Wilson, 1977) en de Maastrichtse Anamnese en Advies Scoringslijst (Kraan & Crijnen, 1987). Het scoringsschema bestaat uit resp. 31 (eerste onderzoek) en 16 (tweede onderzoek) dichtome items aan de hand waarvan de aan- of afwezigheid van concrete boodschappen in de dialogen worden vastgesteld. Een voorbeeld item is bijvoorbeeld "deelt het slechte nieuws mee na een korte introductie" (ja / nee). Voor elke student worden de scores op de items opgeteld. De maximaal mogelijke score is 31 (eerste onderzoek) en 16 (tweede onderzoek). Een hoge score representeert een hoge vaardigheid. Twee onafhankelijk werkende beoordelaars stellen de scores onafhankelijk van elkaar vast. De interbeoordelaarsbetrouwbaarheid (Kappacoëfficiënt) is respectievelijk 0.47 (voormeting/ eerste onderzoek), 0.69 (nameting/ eerste onderzoek) en 0.83 (nameting/ tweede onderzoek).

De tweede test, een *kennistest* (individueel), omvat zes open essay-vragen over de karakteristieken van de vaardigheid. Elk antwoord is een punt (eerste onderzoek; maximaal 6 punten) en twee punten (tweede onderzoek; maximaal 12). De Cronbach's alfacoëfficiënt is .69 (tweede onderzoek).

De derde test, een *classificatie-test* (individueel), bestaat uit 10 ja-nee-items, waarbij de studenten moeten aangeven of er bij een beschreven situatie al dan niet van een slecht-nieuwsgesprek sprake is. Elke vraag is een punt. De Cronbach's alfacoëfficiënt is .73 (tweede onderzoek).

De vierde test, een *gesprekssequentieringstest* (individueel, alleen in het tweede onderzoek), bevat een op te lossen sociaal-communicatief probleem. Aan de studenten wordt gevraagd een zestal uitspraken van de probleemoplosser in de goede volgorde te plaatsen. Het aantal keren wordt geteld dat een uitspraak in de goede volgorde na een andere uitspraak is geplaatst. Elk item is een punt. Voorbeelden van de 15 gebruikte items zijn "Zijn de vier zin-

nen f-d-b-e juist geordend?; "Zijn de drie zinnen f-d-b juist geordend?" Zijn de twee zinnen f-d juist geordend?" etc. De Cronbach's alfacoëfficiënt is .79.

Verder is gebruik gemaakt van een bestaande typetest van een uitzendbureau. Deze test vraagt studenten een brief over te typen. De brief en de tijd is geregistreerd in een logfile. De typesnelheid is uitgedrukt in het aantal aanslagen per minuut.

De evaluatie van de kwaliteit van de leeromgeving wordt gemeten met een vragenlijst. Deze vragenlijst bevat 33 items waarbij studenten op een vijfpuntsschaal de leeromgeving beoordelen (een betekent zeer tevreden, twee is tevreden, drie is tevreden noch ontevreden, vier is ontevreden en vijf is zeer ontevreden). De studenten beoordelen drie aspecten van kwaliteit: de gebruikaspecten, het technisch ontwerp en het didactisch ontwerp. De Cronbach's alfacoëfficiënt is .88.

4. ONDERZOEKEN EN BEVINDINGEN

4.1 Het effect van samenwerkend leren op het verwerven van gespreksvaardigheden

Onderzoeksvragen en -hypothesen

Het eerste onderzoek (Holsbrink – Engels, 1997a) onderzocht de effecten van samenwerkend leren op het verwerven van gespreksvaardigheden. De effectiviteit van een tweetal leeromgevingen werd vergeleken. De effectiviteit van een leeromgeving met de component samenwerkend leren werd vergeleken met de effecten van een leeromgeving zonder samenwerkend leren (controlegroep). De hypothese die aan de basis lag van dit onderzoek was dat de leeromgeving met de component samenwerkend leren een significante, positieve invloed heeft op de vaardigheid in het oplossen van sociaal-communicatieve problemen.

Methode

Een "pretest-posttest plus control group" onderzoeksopzet (Cook & Campbell, 1979) werd gekozen om het effect van samenwerkend leren op de leeruitkomst te meten. Beide groepen kregen een voor-en nameting.

Aan het onderzoek werd deelgenomen door 41 studenten van de Universiteit Twente. Zij werden betaald voor hun deelname. De gemiddelde leeftijd van de 24 vrouwen en 17 mannen was 21 jaar (bereik 18 – 28 jaar). Zij werden random aan een van beide condities toegewezen. Zij studeerden Toegepaste Onderwijskunde (N = 26) en een van de technische wetenschappen (N = 15).

De instructieve materialen zijn hierboven geschetst. Beide leeromgevingen omvatten de volgende onderdelen: (a) voormeting, (b) bestuderen van theorie, (c) video-demonstratie van de vaardigheid en (d) nameting. Zij verschilden met betrekking tot het samenwerkend leren. De leeromgeving met de samenwerkend leren component kende daarnaast vier computerondersteunde rollenspelen die waren toegevoegd tussen (c) video en (d) nameting. Er werden gesprekken geconstrueerd. Beide leeromgevingen waren ontwikkeld met HyperCard 2.1, op een scherm met 640*480 pixels zonder gebruik van kleur.

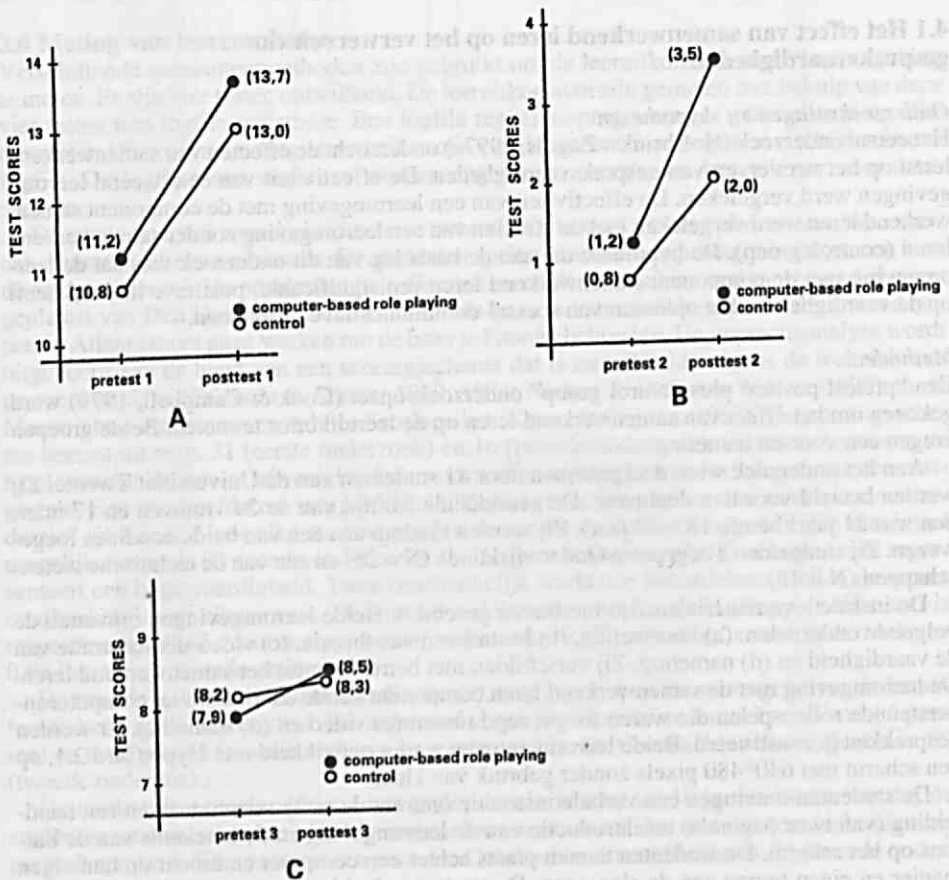
De studenten ontvingen een verbale instructie omtrent de uit te voeren taak en een handleiding (van twee pagina's) ter introductie van de leeromgeving en de betekenis van de buttons op het scherm. De studenten namen plaats achter een computer en mocht op hun eigen manier en eigen tempo aan de slag gaan. De studenten hadden respectievelijk vier uur (de samenwerkend leren groep) en twee uur (controle groep) om in de leeromgeving te werken. Na afloop van de bijeenkomst vulden alleen de studenten van de samenwerkend leren groep de kwaliteitsvragenlijst in. Aparte bijeenkomsten voorkwamen interacties tussen de experimentele en controle condities.

Resultaten

Vanwege computer registratie problemen waren er onvoldedige data-sets van drie proefpersonen. Dit resulteerde in een steekproef van 19 studenten in elke conditie.

Figuur 2 toonde de gemiddelde scores op de drie voor-en nametingen (de vierde test werd alleen in het tweede onderzoek gebruikt). De verschillen op de voormeting waren niet significant.

De Wilcoxon's Signed-Rank scores op de *vaardigheidstest* voor de experimentele en de controle groep waren $Z=-2.4$ and $Z=-2.6$, beide $p \leq .01$, en op de *kennis test* $Z=-3.6$ and $Z=-3.1$, beide $p \leq .01$. Beide groepen behaalden een significante hogere score op beide testen vergeleken met de voormeting. De experimentele groep toonde deze vooruitgang ($Z=-1.9$, $p \leq .05$) eveneens op de *classificatietest*, hetgeen niet van de controle groep aangetoond kon worden ($Z=-0.9$, $p \leq .35$).



Figuur 2: De gemiddelde scores op de drie voor- en nametingen (A = Vaardigheidstest, B = Kennis test en C = Classificatietest).

De verschillscores (scores nameting min scores voormeting) voor beide groepen werden geanalyseerd met behulp van de Mann-Whitney U-tests. De studenten van de samenwerkend leren groep presteerden beter op de kennis- en classificatietests. Het verschil op de *kennistest* is significant, $U(38)=94.5$, $p < 0.01$. De *classificatietest* toonde een trend in dezelfde richting, $U(38)=121$, $p < 0.10$. Het verschil op de *vaardigheidstest* is niet significant ($U(38)=165.5$, n.s.). De inhoud van de uitgewisselde boodschappen bevatte feitelijke informatie, reflecties, verschillende typen vragen, uitingen van emoties (veelal tussen haakjes) en grappes.

De studenten in de samenwerkend leren conditie waren tevreden over de computerondersteunde rollenspelen en zij beoordeelden de reflectiebevorderende vragenlijsten als het minst geapprecieerde onderdeel in de leeromgeving.

Discussie

In dit onderzoek kon een verschil in effect van samenwerkend leren tussen beide groepen worden aangetoond voor de kennistoets: na het doorlopen van de leeromgeving gingen beide groepen hierop vooruit, maar de vooruitgang was groter bij de samenwerkend leren-conditie. Op de vaardigheidstest en de classificatietest werden geen verschillen tussen beide condities gevonden. De belangrijkste bevinding is dat samenwerkend leren, hier in computerondersteunde rollenspelen, het verwerven van gespreksvaardigheden bevordert. Een intervenierende factor is hierbij de ongelijke hoeveelheid tijd die beide groepen hebben besteed aan het leren in de leeromgeving. In toekomstig onderzoek zal dit effect geminimaliseerd dienen te worden door het aanbieden van een "placebo samenwerkend leren"- onderdeel in de leeromgeving.

4.2 Het effect van het gespreksmodel en de mogelijkheden voor reflectie op het verwerven van gespreksvaardigheden

Onderzoeksvragen en -hypothesen

Ook is een effectstudie uitgevoerd naar het effect van het gespreksmodel en de mogelijkheden voor reflectie op het verwerven van gespreksvaardigheden (Holsbrink – Engels, 1997b). Dit onderzoek was een vervolg op het eerste onderzoek. In dit onderzoek lag de nadruk op twee instructiestrategieën die het samenwerkend leren ondersteunen. De effectiviteit van het gebruik van het gespreksmodel en de mogelijkheden voor reflectie in samenwerkend leren werden nader onderzocht. De volgende hypothesen werden geformuleerd. Er werd verwacht dat de aanwezigheid van een gespreksmodel een significante, positieve invloed heeft op het verwerven van de gespreksvaardigheid, op de kennisverwerving en inzicht in de vaardigheid (betere testresultaten). Bovendien werd verwacht dat de studenten in de condities met veel reflectiemogelijkheden significant beter de gespreksvaardigheid verwerven, meer kennis verwerven en meer inzicht in de vaardigheid verkrijgen (betere testresultaten) dan de studenten met weinig reflectiemogelijkheden. Er waren geen verwachtingen ten opzichte van de interactie tussen het gespreksmodel, de mogelijkheden voor reflectie en het geslacht van de subjecten.

Methode

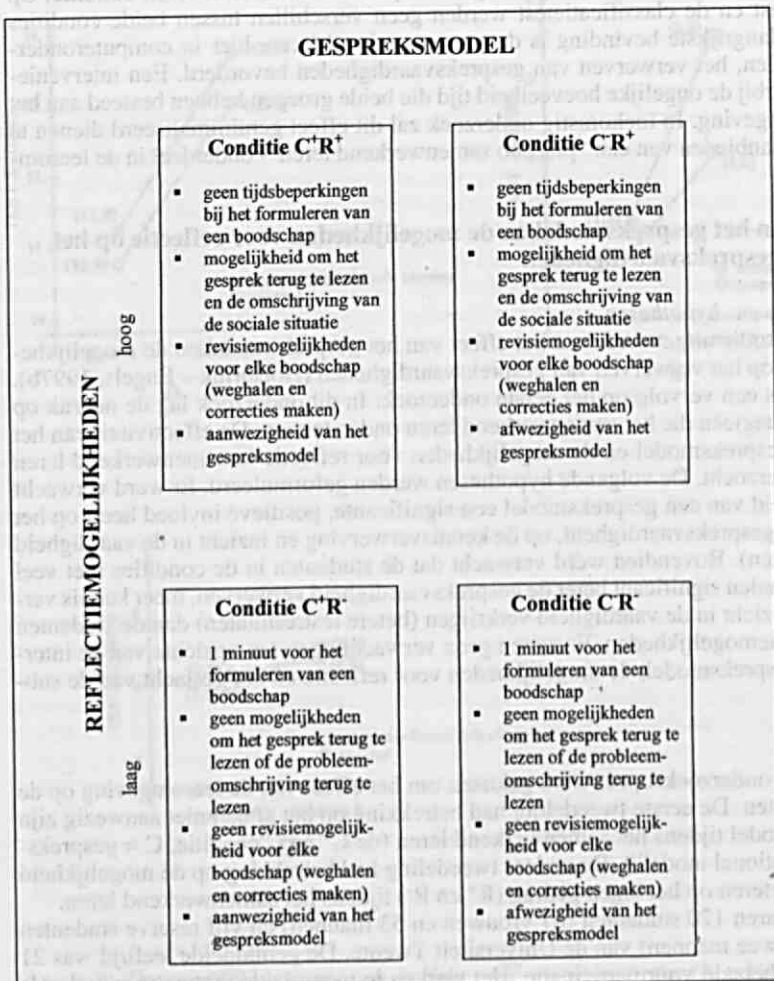
Een 2x2 factoriële onderzoeksopzet werd gekozen om het effect van de leeromgeving op de leeruitkomst te meten. De eerste tweedeling had betrekking op het al dan niet aanwezig zijn van het gespreksmodel tijdens het samenwerkend leren (de C⁺ en C⁻ conditie, C = gespreksmodel = "conversational model"). De andere tweedeling had betrekking op de mogelijkheid al dan niet te reflecteren op het eigen gedrag (R⁺ en R⁻) tijdens het samenwerkend leren.

De subjecten waren 120 studenten (67 vrouwen en 53 mannen) en vijf reserve studenten (drie vrouwen en twee mannen) van de Universiteit Twente. De gemiddelde leeftijd was 21 jaar en zij werden betaald voor participatie. Het werken in paren leidde tot extra uitval. Als een proefpersoon verstek liet gaan in een sessie, werd een vervanger gezocht om het even

aantal proefpersonen te verkrijgen dat nodig was om het samenwerkend leren te realiseren. Indien geen vervanger werd gevonden, werd vervolgens via loting bepaald wie van de wel aanwezige proefpersonen noodgedwongen niet kon deelnemen aan het experiment.

Er werd een gestratificeerde steekproef naar geslacht getrokken. Door aselekt uit de strata te trekken werd de steekproef in elk geval representatief voor de sekse van de populatie. De studenten werden vervolgens random toegewezen -in tweetallen- aan een van de vier condities ($R^+ C^+ = 30$ (17 vrouwen en 13 mannen), $R^+ C^- = 30$ (17 vrouwen en 13 mannen), $R^- C^+ = 30$ (17 vrouwen and 13 mannen), $R^- C^- = 30$ (16 vrouwen and 14 mannen)).

Alle vier leeromgevingen bestonden uit informatie over en videovoorbeelden van het slecht-nieuwsgesprek, samenwerkend leren in computerondersteunde rollenspelen en feedback. De leeromgevingen verschilden op het al dan niet aanwezig zijn van het gespreksmodel (de C^+ en C^- conditie) en op de mogelijkheid al dan niet te reflecteren op eigen gedrag (R^+ en R^-). Figuur 3 geeft een gedetailleerde beschrijving van de variatie in de vier condities.



Figuur 3: De variatie in de vier condities.

De studenten ontvingen een verbale instructie omtrent de uit te voeren taak en bestudeerden een tekst (drie pagina's) over het slecht-nieuwsgesprek. Deze tekst verschilde voor condities, afhankelijk van de aan- of afwezigheid van het gespreksmodel. Nadat alle studenten de tekst hadden bestudeerd (ongeveer 20 minuten) werd een video van een slecht-nieuwsgesprek bekeken waarin een toepassing van het gespreksmodel werd getoond. Voor de C-condities werd een placebo video getoond over het slecht-nieuws gesprek in het algemeen. Daarna namen de studenten plaats achter een computer en mochten op hun eigen manier en eigen tempo aan de slag gaan. Alle 12 bijeenkomsten van zes tot tien proefpersonen vonden plaats in een computerruimte onder leiding van een proefleider.

Resultaten

Vanwege experimentiële uitval (21 subjecten waarvan er vijf konden worden vervangen door reserve studenten) ontbreken er data van 16 subjecten. Dit resulteerde in een steekproef van 104 studenten (57 vrouwen en 47 mannen); 30 studenten in de R⁺C⁺- conditie (17 vrouwen en 13 mannen), 24 studenten in de R⁺C⁻-conditie (14 vrouwen en 10 mannen), 26 studenten in de R⁻C⁺- conditie (16 vrouwen en 10 mannen), and 24 studenten in de R⁻C⁻-conditie (10 vrouwen en 14 mannen).

De gemiddelde scores op de testen worden gepresenteerd in Figuur 4.

MANOVA toonde een significante interactie aan tussen het gespreksmodel en reflectiemogelijkheden, $F(4,91) = 2.69, p < .05$. Om de richting van de significante interactie vast te stellen werden aanvullende univariate analyses uitgevoerd. Univariate analyse toonde aan dat de interactie significant was voor de kennistest, $F(1,94) = 9.55, p < .01$. In de C⁺- conditie, verwierven studenten in combinatie met de R⁺- conditie meer kennis in de leeromgeving dan in combinatie met de R- conditie. In de C- conditie, werden geen aantoonbare verschillen in kennisverwerving gevonden tussen R⁺- and R- condities.

Scheffé multiple vergelijkingstest stelde vast dat studenten in de C⁺R⁺-groep ($M = 10.9$) en studenten in de C⁺R⁻-groep ($M = 9.8$) significant meer kennis verweven dan degenen in de C⁻R⁺- groep ($M = 5.7$) en dan degenen in de C⁻R- groep ($M = 6.5$). Er werden geen andere verschillen gevonden.

Ook leverde MANOVA een significant effect op voor het gespreksmodel, $F(4,91) = 29.38, p < .001$. Univariate analyse toonde aan dat het effect significant was voor twee testen. De vaardigheidstest, $F(1,94) = 8.79, p < .01$ en de kennistest, $F(1,94) = 115.28, p < .001$ leverden een significant verschil op. Studenten in de C⁺- groepen ($M = 9.8$) presteerden significant beter op de vaardigheidstest dan de C⁻- groepen ($M = 8.6$). Dezelfde resultaten werden aangetoond voor de kennistest, studenten in de C⁺- groepen ($M = 10.3$) presteerden significant beter dan studenten in de C⁻- groepen ($M = 6.1$).

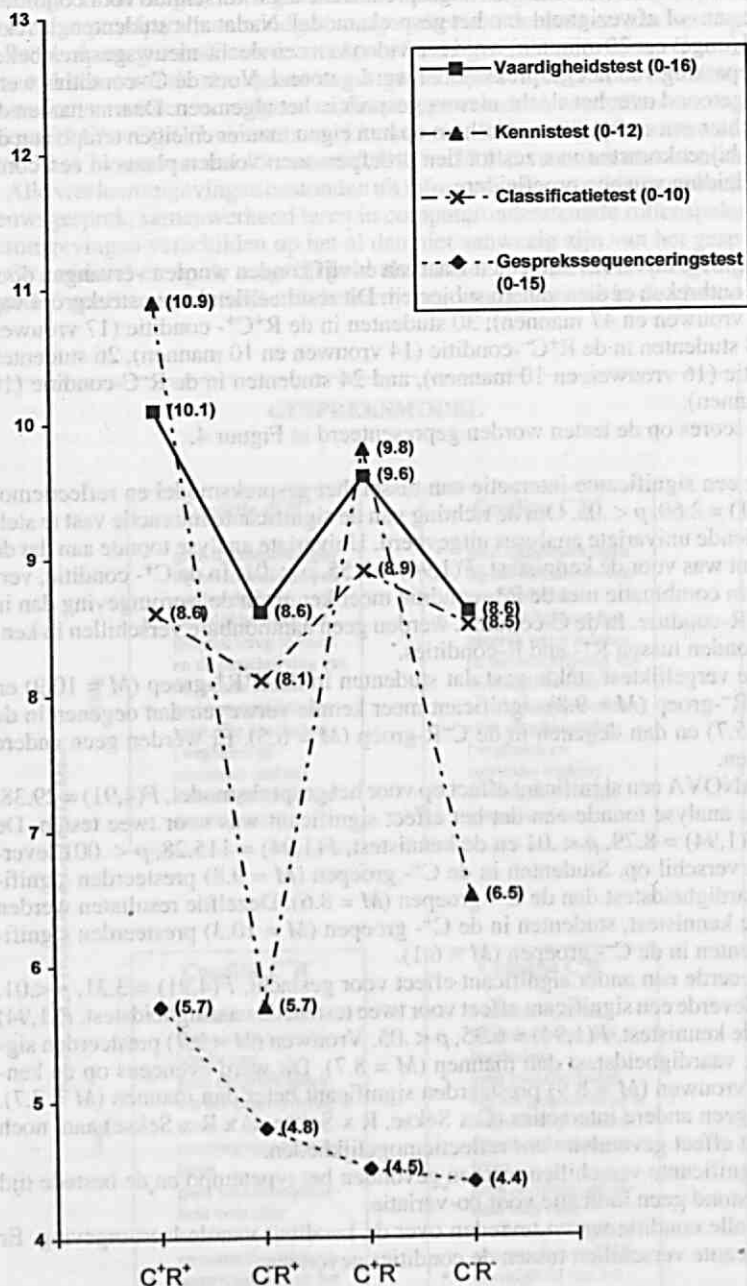
MANOVA indiceerde een ander significant effect voor geslacht, $F(4,91) = 3.31, p < .01$. Univariate analyse leverde een significant effect voor twee testen: de vaardigheidstest, $F(1,94) = 7.26, p < .01$, en de kennistest, $F(1,94) = 6.35, p < .05$. Vrouwen ($M = 9.7$) presteerden significant beter op de vaardigheidstest dan mannen ($M = 8.7$). Dit werd eveneens op de kennistest vastgesteld, vrouwen ($M = 8.9$) presteerden significant beter dan mannen ($M = 7.7$). MANOVA toonde geen andere interacties (C x Sekse, R x Sekse, C x R x Sekse) aan, noch werd een significant effect gevonden voor reflectiemogelijkheden.

Typetest. Geen significante verschillen werden gevonden het typetempo en de bestede tijd aan de testen. Er bestond geen indicatie voor co-variantie.

De studenten in alle conditie waren tevreden over de kwaliteit van de leeromgeving. Er werden geen significante verschillen tussen de condities gevonden.

Discussie

De effectiviteit van de twee instructiestrategieën die het samenwerkend leren ondersteunen, het gebruik van het gespreksmodel en de mogelijkheden voor reflectie, werden nader onder-



Figuur 4: De gemiddelde scores op de tests ter bepaling van de leeruitkomsten als een functie van de aan- of afwezigheid van het gespreksmodel en mogelijkheden voor reflectie.

zocht. De hypothese dat in de C⁺- en R⁺-condities sterkere leereffecten op het niveau van kennis en vaardigheden zouden worden gevonden, werd in het onderzoek duidelijk bevestigd. Daarbij bleek dat de aanwezigheid van het gespreksmodel de grootste invloed heeft op de leereffecten. Het gespreksmodel gaf bruikbare aanwijzingen bij het oplossen van sociaal-communicatieve problemen. Deze informatie was behulpzaam bij het leren uitvoeren van de vaardigheid.

Er waren geen verwachtingen uitgesproken ten opzichte van de interactie tussen het gespreksmodel, de mogelijkheden voor reflectie en geslacht van de subjecten. Deze verwachting kan worden bijgesteld. In het onderzoek werd aangetoond dat als het aanbieden van het gespreksmodel gepaard gaat met mogelijkheden voor reflectie de leereffecten groter zijn. Dit heeft consequentie voor de onderzochte instructiestrategieën binnen het samenwerkend leren. Het aanbieden van reflectiemogelijkheden is effectiever als dit geschiedt in combinatie met het aanbieden van een gespreksmodel.

Geslacht bleek eveneens van invloed op de leeruitkomst. Onderzoek heeft aangetoond dat mannen en vrouwen verschillen in hun communicatiestijlen (o.a. Coates, 1993; Cowan, Wilcox & Nykodym, 1990; Penly, Alexander, Jernigan & Henwood, 1991; Smith & DeWine, 1991; Tannen, 1994). Vrouwen zijn meer competent op bepaalde communicatie items (Luketich, Colliver & Galofré, 1992; Stillman, Regan, Swanson & Haley, 1990). Het geslacht van de studenten in het onderhavig onderzoek had invloed op de verwerving van de vaardigheid. Nader onderzoek werd nodig geacht om dit resultaat goed te kunnen interpreteren.

5. CONCLUSIE EN DISCUSSIE

De op samenwerkend leren gerichte leeromgeving blijkt te voldoen en is effectief voor het leren oplossen van sociaal-communicatieve problemen. Het blijkt mogelijk om gespreksvaardigheden te verwerven met behulp van een computerondersteund netwerk voor samenwerkend leren. Beide instructiestrategieën (de aanwezigheid van een gespreksmodel en mogelijkheden tot reflectie) die het samenwerkend leren ondersteunen, hebben een positieve invloed op het verwerven van gespreksvaardigheden en op het verwerven van kennis en inzicht. In de uitgevoerde empirische onderzoeken werd substantiële steun gevonden voor beide instructiestrategieën.

Er wordt voorgesteld om de onderzoeken te repliceren voor andere categorieën gespreksvaardigheden bijvoorbeeld onderhandelen, voor andere leersituaties bijvoorbeeld via afstandsleren (internettoepassing), voor andere doelgroepen en voor andere contexten bijvoorbeeld bedrijfsopleidingen. Een internettoepassing voor het verwerven van onderhandelingsvaardigheden is inmiddels binnen de afdeling ontwikkeld.

De beschreven leeromgeving is geschikt voor samenwerkend leren van twee partners. Samenwerkend leren vindt veelal in een grotere groep plaats. De implicatie voor de effectiviteit en efficiëntie van de beschreven instructiestrategieën in geval van meer dan twee partners is onbekend. Verder onderzoek zal hierover meer inzicht kunnen geven.

Over de centrale leerprocessen bij het verwerven van sociaal-communicatieve probleemoplossingsvaardigheden is nog weinig bekend. Meer onderzoek naar observerend leren (Bandura, 1969, 1972, 1977a, 1977b) dat al op jonge leeftijd begint, categorisatie en schemaverwerving lijkt op zijn plaats.

Het uitsluitend inzetten van deze op samenwerkend leren gerichte leeromgeving in gespreksstrainingen wordt ontraden. De leeromgeving kan omschreven worden als een alternatieve instructieve methode voor beginners, welke geïntegreerd zal dienen te worden met het gangbare rollenspel en rollenspel met acteurs (toename van de moeilijkheidsgraad). De bekwaamheden in gespreksvoering kunnen dan verworven worden. Vervolgonderzoek zal hierover meer duidelijkheid kunnen verschaffen.

Aangezien samenwerkend leren op verschillende wijzen plaats vindt, zal een verschei-

denheid aan meettechnieken en beoordelingsstrategieën ontwikkeld moeten worden om het samenwerkend leren te in kaart te brengen en te controleren.

Daarnaast zijn de onderzoeken en bevindingen van praktisch belang voor het onderwijs. Er is duidelijke steun gevonden voor de veronderstelling dat een heuristisch tijdens het samenwerkend leren "houvast biedt" aan de studenten. Het gespreksmodel helpt de studenten bij het doorlopen van de stappen van het probleemoplossingsproces. Door beginners veel te laten oefenen, dat wil zeggen veel sociaal-communicatieve problemen te laten oplossen, kunnen zij de gespreksvaardigheid verwerven.

Op samenwerkend leren gerichte leeromgevingen kunnen effectief zijn bij het verwerven van sociaal-communicatieve probleemoplossingsvaardigheden. Het leidt echter geen twijfel dat er nog veel werk verzet dient te worden om tot verantwoorde inzet van en gebruik van technologie bij dit type probleem oplossen en in dit type leeromgevingen te komen. Het gelijktijdig (synchroon) werken aan dezelfde taak (co-constructie van kennis en vaardigheden) door samenwerkingspartners lijkt hierbij cruciaal.

LITERATUUR

- Alessi, S. M. (1988). Fidelity in the Design of Instructional Simulations. *Journal of Computer-Based Instruction*, 15 (2), 40-47.
- Anderson, J. R. (1982). Acquisition of Cognitive Skill. *Psychological Review*, 89, 369-406.
- Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge: Harvard University Press.
- Anderson, J. R. (1987). Skill Acquisition: Compilation of Weak-Method Problem Solutions. *Psychological Review*, 94, 192-210.
- Anderson, J. R. (1993). Problem Solving and Learning. *American Psychologist*, 48 (1), 35-44.
- Argyle, M. (1984). Some New Developments in Social Skills Training. *Bulletin of the British Psychological Society*, 37, 405-410.
- Argyris, C. (1951). *Role Playing in Action*. (Bulletin No. 16). New York: Cornell University.
- Bandura, A. (1969). Social Learning Theory of Identificatory Processes. In A. Goslin (Ed.), *Handbook of Socialization Theory and Research*. Chicago: Rand McNally.
- Bandura, A. (1972). Modeling Theory: Some Traditions, Trends and Disputes. In R. Parke (Ed.), *Recent Trends in Social Learning Theory*. New York: Academic Press.
- Bandura, A. (1977a). A Self-Efficacy: Toward A Unifying Theory of Behavior Change. *Psychological Review*, 84, 191-215.
- Bandura, A. (1977b). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice Hall.
- Boud, D., & Walker, D. (1991). In the midst of experience: developing a model to aid learners and facilitators. *Paper presented at the National Conference of Experiential Learning*, University of Surrey, England, 16-18 July.
- Chase, W. G., & Simon, H. A. (1973). Preception in Chess. *Cognitive Psychology*, 4, 55-81.
- Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.) (1988). *The Nature of Expertise*. Hillsdale, NJ: Lawrence Erlbaum.
- Coates, J. (1993). *Women, men and language: A sociolinguistic account of differences in language*. London: Longman.
- Cohen, E. G. (1994). Restructuring the Classroom: Conditions for Productive Small Groups. *Review of Educational Research*, 64 (1), 1-35.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation: Design Analysis, Issues for Field Settings*. Chicago: Rand McNally Publishing Company.
- Corsini, R. J. (1960). Role Playing: Its Use in Industry. *Advanced Management*, 25 (2), 20-23.
- Cowan, K. M., Wilcox, J. R., & Nykodym, N. (1990). A comparative analysis of female communication style as a function of organizational level. *Communications*, 15 (3), 291-309.
- De Groot, A. D. (1965). *Thought and Choice in Chess*. The Hague: Mouton.
- Elshout, J. J. (1987). Problem-Solving and Education. In E. de Corte, H. Lodewijks, R. Parmentier, & P. Span (Eds.), *Learning and Instruction* (pp. 259-273). Oxford: Pergamon.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis. Verbal reports as data*. Cambridge, Mass.: The MIT Press.

- Erkens, (1997). *Coöperatief probleemoplossen met computers in de onderwijs: Het modelleren van coöperatieve dialogen voor de ontwikkeling van intelligente onderwijssystemen*. Utrecht: Universiteit Utrecht, Vakgroep Onderwijskunde.
- French, J. R. P. (1945). Role Playing as a Method of Training Foremen. *Sociometry*, 8, 410-422.
- Georges, J. C. (1989). The Hard Reality of Soft-skills Training. *Personnel Journal*, 4, 41-45.
- Greeno, J. G. (1978). A Study of Problem Solving. In R. Glaser (Ed.), *Advances in Instructional Psychology* (pp. 13-73). Hillsdale, NJ: Lawrence Erlbaum.
- Greeno, J. G. (1980). Trends in the Theory of Knowledge for Problem-solving. In D. T. Tuma, & F. Reif (Eds.), *Problem-solving and Education: Issues in Teaching and Research* (pp. 9-24). Hillsdale, NJ: Lawrence Erlbaum.
- Holsbrink - Engels, G.A. (1994). Visions of Dutch Corporate Trainers on Role Playing. In: Saunders, D., Percival, F. & Armstrong, R. (Eds.). *The 1994 Simulation and Gaming Yearbook*. London: Kogan Page Publishers Ltd.
- Holsbrink-Engels, G. A. (1997a). Computer-based Role-playing for Interpersonal Skills Training. *Simulation and Gaming*, 27 (4). 164-180.
- Holsbrink-Engels, G.A. (1997b). The effects of the use of a conversational model and opportunities for reflections in computer-based role playing. *Computers in Human Behavior*. USA: Pergamon/ Elsevier Science Ltd.
- Holsbrink-Engels, G.A. (1998). *Computer-based role playing for interpersonal skills training*. Enschede: University of Twente, Faculty of Educational Science and Technology, Department of Instructional Technology. [proefschrift]
- Johnson, D. W. & Johnson, R.T. (1989). *Cooperation and competition: Theory and Research*. Edina, MN: Interaction Book Company.
- Jong, A.J.M. de (1999). *Proef of de som* (oratie). Enschede: Universiteit Twente
- Kass, A., Burke, R., Blevis, E., & Williamson, M. (1993,1994). Constructing learning environments for complex social skills. *The Journal of Learning Sciences*, 3, 387-427.
- Kessels, J., & Smit, C. (1990). Rollenspel? Ik kijk wel [Role-play? I watch]. *Opleiding en Ontwikkeling*, 4, 9-15.
- Kraan, H. F., & Crijnen, A. A. M. (1987). *The Maastricht History-taking and Advice Checklist: Studies of Instrumental Utility*. Amsterdam: Lundbeck.
- Lippitt, R. (1943). The Psychodrama in Leadership Training. *Sociometry*, 6, 286-292.
- Luketich, G. F., Colliver, J. A., & Galofré, A. (1992). Two studies of the effects of gender within single case simulated by male and by a female standardized patient in a multiple-stations examination. In R. M. Harden, I. R. Hart & H. Mulholland. *Approaches to the assessment of clinical competence (part 1)*. Conference book, Centre for Medical Education, Dundee Scotland.
- Mettes, C. T. W., Pilot, A., & Roossink, H. J. (1981). Linking factual knowledge and procedural knowledge in solving science problems: A case study in a thermodynamics course. *Instructional Science*, 10, 333-361.
- Van Merriënboer, J. J. G. (1997). *Training Complex Cognitive Skills; A Four-Component Instructional Design Model for Technical Training*. Englewood Cliffs, NJ: Educational Technology Publications.
- Newell, A., & Simon, H. A. (1972). *Human Problem-solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, R. E., & DeCamp Wilson, T. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Penley, L. E., Alexander, E. R., Jernigan, I. E., & Henwood, C.I. (1991). Communication abilities of managers: The relationship to performance. *Journal of Management*, 17 (1), 57-76.
- Schoenfeld, A. H. (1979). Can heuristics be taught? In J. Lochhead, & J. Clement (Eds.), *Cognitive process instruction* (pp. 315-338). Philadelphia: Franklin Institute Press.
- Schoenfeld, A. H. (1985). *Mathematical Problem-solving*. New York: Academic Press.
- Schön, D. (1983). *The Reflective Practitioner: How Professionals Think In Action*. New York: Basic Books.
- Schön, D. (1987). *Educating the Reflective Practitioner: Toward a New Design For Teaching and Learning in the Professions*. San Francisco: Jossey Bass.
- Simon, H. A. (1973). The Structure of Ill-Structured Problems. *Artificial Intelligence*, 4, 181-201.
- Smith, G. L., & DeWine, S. (1991). Perceptions of subordinates and requests for support. Are males and females perceived differently when seeking help? group & *Organization Studies*, 16 (4), 408-427.
- Stillman, P. L., Regan, M. B., Swanson, D. B., & Haley, H. L. (1990). Does gender affect clinical skills as measured by multiple station examinations using standardized patients. In I. R. Hart & R. M. Harden, *More developments in assessing clinical competence*. Conference book. Montreal, Canada: Can-Heal Publications.
- Tannen, D. (1994). *Talking from 9 to 5*. New York: Morrow.

- Van Merriënboer, J. J. G. (1997). *Training Complex Cognitive Skills; A Four-Component Instructional Design Model for Technical Training*. Englewood Cliffs, NJ: Educational Technology Publications.
- Voss, J. F. (1987a). Learning and Transfer in Subject-Matter Learning: A Problem-Solving Model. *International Journal of Educational Research*, 11, 607-622.
- Voss, J. F. (1987b). Problem-solving and the Educational Process. In A. Lesgold, & R. Glaser (Eds.), *Foundations for a Psychology of Education* (pp. 251-294). Hillsdale NJ: Lawrence Erlbaum.
- Voss, J. F. (1990). Das Lösen Schlecht Strukturierter Probleme; Ein Überblick. *Unterrichtswissenschaft*, 18 (4), 313-337.
- Winograd, T., & Flores, F. (1986). *Understanding computers and cognition*. Norwood, NJ: Ablex Publishing Corporation.

Cultureel kapitaal en onderwijsprestaties: differentiëren naar etnische groep?

G.W.J.M. Driessen

ITS, Universiteit Nijmegen, e-mail: g.driessen@its.kun.nl

ABSTRACT

Some thirty years ago Bourdieu formulated his social reproduction theory to gain a better understanding of how social class influences the transmission of educational inequality. Cultural capital plays a central role in this theory. Surprisingly few studies have been carried out to test whether this theory is applicable to ethnic minorities as well. This paper presents the results of a study using data from the Dutch Primary Education cohort ('PRIMA'). Four ethnic groups are included: Dutch, Surinamese, Turkish, and Moroccan students. The main variables are language and mathematics test scores, and a number of cultural capital indicators, including financial resources, linguistic resources, parental reading behavior, and educational resources in the family. The results show that - with the exception of the language scores of the Surinamese students - within ethnic groups there are no mediating effects of cultural resources. The results suggest that in future research and in practice it is relevant to not treat ethnic minorities as one homogenous group, but to differentiate between the various groups.

1. INLEIDING

In loopbaan- en stratificatie-onderzoek zijn de laatste decennia tal van theorieën ontwikkeld die een verklaring proberen te bieden voor verschillen in schoolloopbanen (vgl. Roelandt, 1994). Niet overtuigd door de opvatting van een meritocratisch ingerichte maatschappij, zijn vanaf de jaren '70 een aantal reproductie-theorieën in zwang geraakt (Meesters, 1992). De bindende gedachte achter deze theorieën is dat er reproductie plaatsvindt van de bestaande maatschappelijke ongelijkheid, los van de intelligentie of capaciteiten van mensen. Bovendien, en dat is een wezenlijk element van deze theorieën, neemt het onderwijssysteem een sleutelpositie in bij dit reproductie-proces. Een van de belangrijkste exponenten van de reproductie-these is de Franse socioloog Bourdieu met zijn cultuur-politieke theorie (vgl. Pels, 1992). Centraal bij Bourdieu staat de culturele overdracht van sociale ongelijkheid. Volgens hem is er in de westerse samenleving, anders dan de meritocratische ideologie wil doen geloven, nog steeds sprake van een standenmaatschappij. Het onderwijs speelt een in Bourdieu's ogen verwerpelijke rol, doordat het de dominante statusgroepen de legitimering verleent voor sociale reproductie. In het onderwijs vindt een selectieproces plaats dat is gebaseerd op arbitraire criteria, namelijk op basis van verschillen in cultuurpatronen en levensstijlkenmerken. Belangrijk is dat deze patronen en kenmerken - 'het cultureel kapitaal' - binnen het eigen milieu van generatie op generatie worden overgedragen. Gedacht kan worden aan linguïstische en sociaal-culturele competenties en affiniteiten (bv. omgangsvormen, taalgebruik, interesses, smaak voor kunst en Cultuur, maatschappelijke en politieke opvattingen en voorkeuren). Deze 'habitus' verschilt van statusgroep tot statusgroep. Cultureel kapitaal kan niet op school verworven worden. Evenmin kan de school de verschillen in hoeveelheid kapitaal van

onderscheiden groepen van leerlingen wegwerken. Het tegendeel is het geval: door de wijze waarop het onderwijs is ingericht, het eenzijdige belang dat het onderwijs hecht aan een bepaalde cultuur (die van de dominante statusgroepen), het aansluiten bij die cultuur en het vooronderstellen van bepaalde codes en capaciteiten, leiden juist tot vergroting van de bestaande verschillen. Het cultureel kapitaal dat het onderwijs bij zijn leerlingen veronderstelt, kan vooral worden verworven tijdens de primaire socialisatie in het gezin, en dan met name via de opvoeding van ouders die zelf tot de dominante cultuur behoren. Hoe sterker de ouders zijn ingebed in de dominante cultuur en hoe meer hun opvoeding daar op gericht is, des te beter zouden hun kinderen zich thuis voelen op school en des te meer zouden ze kunnen profiteren van het onderwijs dat daar wordt aangeboden. Tussen thuis en school bestaat daarmee mogelijk een kloof waarvan de breedte wordt bepaald door het beschikken over cultureel kapitaal en het toebehoren tot de dominante statusgroepen. Omdat dus de dominante cultuur de kern vormt van het - verborgen - curriculum van het onderwijs (vgl. Jungbluth, 1985), is het voor kinderen uit de lagere strata (autochtone en allochtone arbeiderskinderen) moeilijk vat te krijgen op de aangeboden leerstof. Dit vormt daarmee de verklaring voor de lagere prestaties, de hogere drop-out en de lagere gerealiseerde onderwijsniveaus van deze groepen.

In Nederland zijn de 'theorieën' van Bourdieu door enkele onderzoekers empirisch getoetst. In sommige gevallen werd daarbij de meer algemene term 'kapitaal' vervangen door 'hulpbronnen'. De Graaf (1987) bijvoorbeeld, analyseerde met behulp van verschillende databestanden niet alleen de invloed van culturele hulpbronnen, maar ook die van financiële hulpbronnen (bv. inkomen). Hoewel de resultaten niet consistent waren voor de onderscheiden bestanden, luidde een conclusie dat de effecten van de financiële hulpbronnen bescheiden waren, maar dat leesgedrag van ouders en passieve cultuurparticipatie het hele effect van sociale achtergrond (i.c. opleidingsniveau ouders) op het bereikte opleidingsniveau van hun kinderen medieerde. Tesser & Mulder (1990) repliceerden voor een deel de studie van De Graaf. De bevindingen van De Graaf konden echter door hen niet bevestigd worden. Wel vonden ze dat aan het aspect 'boekenbezit' een speciale betekenis moet worden toegekend. Dit kan namelijk worden opgevat als een verschijningsvorm van een specifieke relatie tot taal. Mogelijk komt daardoor een vorm van participatie aan een 'cultuur van geletterdheid' (vgl. Leseman, 1989) tot uitdrukking. Van der Velden (1991) liet zien dat het bezit van cultureel kapitaal (o.m. leesgedrag, cultuur-pedagogisch niveau van de opvoeding, inzicht in het onderwijs) de effecten van het opleidings- en beroepsniveau van de ouders op de schoolloopbanen van hun kinderen vrijwel geheel verklaren. Echter, hoewel de onderzochte gezinskenmerken een deel van het statusseffect medieëren, dragen ze niet bij aan een betere voorspelling van de loopbaan. Ze geven daarentegen wel aan via welke processen de effecten hun beslag krijgen. Driessen (1993) onderzocht enkele aspecten van cultureel kapitaal (boekencultuur, cultuurparticipatie, onderwijsondersteunend gedrag van ouders en taalkeuze) bij een omvangrijke steekproef van voortgezet onderwijs-leerlingen. Hij kwam tot dezelfde conclusies als Van der Velden: de invloed van de gezinsstructurele indicatoren (sociaal milieu) verloopt via de gezinspedagogische indicatoren (cultureel kapitaal), maar deze laatste kenmerken voegen nagenoeg niets toe aan de verklaringskracht van de eerste kenmerken. Leseman, Sijsling, Jap-A-Joe & Sahin (1995) onderzochten effecten van gezinskenmerken op de prestaties van kinderen in het eerste jaar van het basisonderwijs. Hun analyses maakten duidelijk dat de effecten van het sociaal-etnische milieu op de woordenschat van de kinderen volledig worden gemedieerd door de thuistaal, de instructieve kwaliteit van de ouder-kindinteracties en de mate van sociaal-emotionele ondersteuning daarbij. Naast de thuistaal bleken met name de pedagogische denkbeelden en het cultureel-educatief kapitaal ('leefstijlkenmerken') binnen het gezin belangrijke schakels te zijn tussen de etnisch-culturele herkomst en de socialisatie van cognitieve vaardigheden. In een recente studie gingen Leseman & Van den Boom (1999) na wat de effecten zijn van moeder-kind interacties met betrekking tot het lezen van boeken en oplossen van problemen op de cognitieve ontwikkeling van kleuters uit verschillende etni-

sche groepen, te weten Nederlanders, Surinamers en Turken.¹ Aan dit onderzoek lag het biologische model van Bronfenbrenner en Ceci ten grondslag. Gebruik werd gemaakt van gegevens die waren verzameld via vraagesprekken, observaties en toetsafnames bij 130 gezinnen. De analyses lieten zien dat er grote verschillen waren tussen de etnische groepen wat betreft milieu, interacties en cognitieve vaardigheden. Maar niet alleen verschilden de gemiddelden, dat gold ook voor de samenhangen tussen de kenmerken. De werkingen van de moeder-kind interacties zijn onderzocht via multi-sample LISREL. Uit de resultaten van deze analyses bleek dat er ook hier grote verschillen waren in effecten tussen de etnische groepen. Ook na controle voor binnengroepseffecten van cognitieve vaardigheid op 3-jarige leeftijd, sociaal milieu en moeder-kind interacties, resteerde er een tussengroepseffect op de cognitieve vaardigheid op 4-jarige leeftijd. Op basis van de bevindingen uit deze studie kunnen we dus concluderen dat effecten die voor de hele (heterogene) steekproef gelden, niet noodzakelijkerwijs ook voor onderscheiden subgroepen van toepassing zijn.

Uit onderwijs-sociologisch onderzoek is inmiddels genoegzaam bekend dat de schoollopen van arbeiderskinderen vaak problematisch zijn. Vooral *allochtone* arbeiderskinderen vertonen flinke onderwijsachterstanden (Driessen & Tesser, 1998). Hoewel de beschrijving van deze achterstanden ver gevorderd is, blijft er wat de verklaring ervan betreft nog veel onduidelijkheid. Zoals we hierboven hebben aangeduid, is de reproductie-these van Bourdieu in Nederland inmiddels enkele malen getoetst, in hoofdzaak met behulp van gegevens over autochtone respondenten, dan wel met gegevens waarbij allochtone en autochtone respondenten niet worden onderscheiden. Voor een deel vond ze bevestiging, voor een deel niet. Tegelijkertijd is ze op onderdelen bekritiseerd, onder meer vanwege haar starheid, haar gerichtheid op de formele Cultuur, en op de empirische onderbouwing van haar vooronderstellingen. Daarnaast heeft ze ook nieuwe impulsen gekregen door de toepassing van meer geavanceerde analysetechnieken, de verlegging van het accent van participatie in de formele cultuur naar cultureel-pedagogische en cultureel-scholaire activiteiten binnen het gezin in algemene zin (vgl. Leseman, 1989; Meijnen, 1990; Van der Velden, 1991). Van recente datum is ook de hernieuwde aandacht voor diverse aspecten van 'geletterdheid' als vorm van cultureel kapitaal (zie hiervoor ook Carrington & Luke, 1997). Dit 'linguïstisch kapitaal' krijgt met name in combinatie met andere vormen van kapitaal haar betekenis, bijvoorbeeld in interacties binnen het gezin en op school (vgl. ook Leseman, Sijsling, Jap-A-Joe & Şahin, 1995; Leseman & Van den Boom, 1999). Dit gevoegd bij het toenemend inzicht in de cultuur-bepaalde opvoedingspatronen binnen allochtone gezinnen (vgl. Nijsten, 1998; Pels, 1998), maakt het aantrekkelijk en relevant de cultureel kapitaal-these nu ook voor allochtone leerlingen te toetsen. Dat is zoals gezegd weliswaar al een beperkt aantal keren gedaan (bv. Tesser & Mulder, 1990; Leseman, Sijsling, Jap-A-Joe & Şahin, 1995; Leseman & Van den Boom, 1999), maar die studies hadden meestal betrekking op slechts enkele honderden leerlingen uit een beperkte regio, dan wel er werd minder expliciet aandacht besteed aan de etnische factor en gebruik gemaakt van andere indicatoren voor cultureel kapitaal. In enkele van deze studies is dit kenmerk in een pad-analytische opzet geanalyseerd. In het ene geval is deze als enkele dummy-variabelen opgenomen, en in het andere geval als een interval-variabele. In elk van die studies blijkt etnische herkomst van belang te zijn; de ene keer direct en de andere keer indirect via kapitaal.

In Nederland is de aanwezigheid van etnische minderheden (hier: de combinatie van allochtoon en laag sociaal-economisch milieu) een fenomeen van betrekkelijk recente datum. Vandaar dat het zinvol is om na te gaan of en met welke resultaten de kapitaal-these in relatie tot minderheden is getoetst in landen met een langere migratie-traditie. We hebben daarom een literatuurstudie uitgevoerd naar onderwijs-sociologisch onderzoek op dit terrein in de Verenigde Staten. Bij die zoekactie hebben we de volgende combinaties van trefwoorden gehanteerd: (vormen van) kapitaal/hulpbronnen, sociaal milieu, etniciteit, onderwijs(prestaties). Om

maar meteen met de deur in huis te vallen: de resultaten vallen sterk tegen, niet alleen vanwege het geringe aantal relevante studies, maar ook vanwege de specifieke aandacht die 'race' en 'ethnicity' daarin toebedeeld krijgen en de consistentie in resultaten.² De studies vertegenwoordigen een breed scala aan thema's binnen uiteenlopende settings. We zijn daarbinnen weliswaar op het spoor gekomen van relevante aanzetten en nieuwe gezichtspunten, maar in zijn algemeenheid valt de bruikbaarheid tegen.

Als we de theoretische invalshoek bekijken, valt op dat in vergelijking tot de Nederlandse (en West-Europese) studies de Amerikaanse studies veel meer gericht zijn op het concept sociaal kapitaal vanuit de optiek van James Coleman (Coleman, 1988; Coleman & Hoffer, 1987). Colemans these van sociaal kapitaal heeft betrekking op de structuur van de sociale relaties tussen mensen. Het deel uitmaken van een gemeenschap waarin onderwijsprestaties gewaardeerd en gestimuleerd worden, kan op zich worden beschouwd als een vorm van kapitaal. De klassieke interpretatie van cultureel kapitaal à la Bourdieu (in termen van bv. theater-, concert- en museumbezoek: 'high-brow cultural participation') komt relatief weinig voor in de Amerikaanse studies (vgl. Aschaffenburg & Maas, 1997; Kalmijn & Kraaijkamp, 1996). Opvallend is ook dat cultureel/sociaal kapitaal specifieke invullingen krijgt: religieus kapitaal als de bindende factor binnen een (lokale) gemeenschap (Covrig, 1996), de culturele subgroep (van jongeren) als vorm van kapitaal (Sullivan, 1997), en ook etnisch kapitaal: het toebehoren tot een bepaalde etnische groep als waarde op zich (Bankston, Caldas & Zhou, 1997; Lareau & McNamara Horvat, 1999). Wat de uitkomsten van de onderzoeken betreft zijn de bevindingen - zoals al opgemerkt - niet consistent, in het ene geval heeft cultureel, respectievelijk sociaal kapitaal wel een mediërende functie, en in het andere geval niet. Hierbij aansluitend kan nog worden opgemerkt dat enkele studies wel de mobiliteitstheorie (cultureel kapitaal kan ook een weg tot mobiliteit voor kinderen uit de lagere klassen zijn, vgl. DiMaggio, 1982) bevestigen, maar niet Bourdieu's reproductietheorie (die strikt genomen betrekking heeft op de hogere klassen). Wat ten slotte de centrale invalshoek van onze speurtocht betreft, de betekenis van cultureel kapitaal voor allochtone groepen, stellen de bevindingen ook teleur. De verwachting was dat landen die een veel langere traditie met etnische minderheden kennen en waar de omvang van deze groepen ook veel groter is ook meer aandacht aan deze invalshoek zouden schenken. Wat de Verenigde Staten betreft is dit niet uitgekomen. In het algemeen wordt bij onderzoek op het gebied van cultureel kapitaal geen expliciet onderscheid gemaakt naar etniciteit of blijft de steekproef beperkt tot de groep blanken. De meeste van de in deze paragraaf aangehaalde studies vormen hierop weliswaar een uitzondering, maar dat wil niet zeggen dat etniciteit een centraal thema vormt of dat er expliciet aandacht aan wordt besteed. Dit neemt niet weg dat we enkele interessante noties tegen zijn gekomen. Relevant is bijvoorbeeld de conclusie van Lareau & McNamara Horvat (1999) dat er een verschil bestaat tussen het bezitten en het activeren van kapitaal, en dat kapitaal pas waarde krijgt in specifieke settings. Met name wijzen zij daarbij op het belang van de leerling-leerkracht interacties en op de rol die de leerkracht daarbij speelt als 'gatekeeper' (vgl. ook Farkas, Grobe, Sheenan & Shuan, 1990, en Kalmijn & Kraaijkamp, 1996). Noemenswaardig is ook dat in een aantal studies 'linguistic capital' een belangrijke rol speelt (bv. Stanton-Salazar & Dornbusch, 1995), hetgeen in een land met zoveel verschillende talen eigenlijk niet zo verwonderlijk is. Een aparte invalshoek is die van Ainsworth-Darnell & Downey (1998) die een koppeling leggen tussen cultureel kapitaal en John Ogbu's 'oppositional culture explanation' en 'resistance model' (vgl. ook Kalmijn & Kraaijkamp, 1996, en Roscigno & Ainsworth-Darnell, 1999). Van belang is verder de aandacht die uitgaat naar onderwijsondersteunende activiteiten door ouders, zoals het geven van huiswerkhulp en het bezoeken van ouderbijeenkomsten op school (Furstenberg & Hughes, 1995). Opmerkelijk is ten slotte het aantal keren dat gewezen wordt op het belang van de gezinssamenstelling, i.e. het aantal kinderen en eenouder- versus volledige gezinnen, als indicator van kapitaal (bv. Bianchi & Robinson, 1997; Downey, 1995; Runyan et al., 1998). De specifieke rol van etniciteit in dit alles is overigens niet altijd even duidelijk.

Opvallend in Amerikaans onderzoek is dat het doorgaans betrekking heeft op 'high school' studenten; slechts incidenteel richt men zich op andere leeftijdsgroepen, zoals 'high-risk preschool' kinderen (bv. Runyan et al., 1998). Vermeldenswaardig is ook dat veelal wat oudere data worden gebruikt (vgl. Furstenberg & Hughes, 1995; Teachman, Paasch & Carver, 1996a, 1996b; White & Kaufman, 1997). De manier waarop etniciteit in de analyses is meegenomen is tweërlei: enerzijds door de steekproef te beperken tot een bepaalde allochtone groep (bv. Roscigno & Ainsworth-Darnell, 1999) en anderzijds door de etnische herkomst in enkele dummy-variabelen uiteen te leggen. In het laatste geval worden deze dummies gebruikt in (normale) regressie-analyses dan wel logistische regressie-analyses (vgl. Ainsworth-Darnell & Downey, 1998; Aschaffenburg & Maas, 1997). De mogelijkheid om de analyses voor elke etnische groep apart te verrichten en de resultaten te vergelijken zijn we slechts een enkele keer tegengekomen, namelijk in het onderzoek van Okagaki & Frensch (1998). Zij deden onderzoek onder drie etnische categorieën in de Verenigde Staten (Aziaten, Latino's en Europeanen) naar de effecten van ouderlijke opvattingen en gedrag met betrekking tot opvoeding en onderwijs op de schoolprestaties van hun kinderen. Ze voerden daarvoor niet alleen regressie-analyses uit binnen elk van de onderscheiden groepen, maar ook over de groepen heen waarbij interacties tussen de predictoren en etnische groep werden ingevoerd. Hun conclusie luidde dat er grote verschillen zijn tussen etnische groepen in de samenhangen tussen opvattingen en gedrag en schoolprestaties, ook nadat gecontroleerd is voor sociaal milieu en inkomen.

Naar aanleiding van de bovenstaande Nederlandse en internationale bevindingen willen we met behulp van recente empirische gegevens de volgende onderzoeksvragen beantwoorden:

- In hoeverre beschikken onderscheiden groepen van allochtonen over culturele en financiële hulpbronnen? In hoeverre verschillen zij daarbij met autochtonen?
- Welke samenhangen bestaan er tussen de verschillende aspecten van cultureel en financieel kapitaal enerzijds en de onderwijsprestaties van basisschoolleerlingen anderzijds? Is er daarbij een verschil tussen allochtone en autochtone leerlingen?
- In hoeverre wordt het effect van sociaal milieu op onderwijsprestaties gemedieerd door de culturele en financiële hulpbronnen? Verschilt dit voor allochtone en autochtone leerlingen?

2. METHODE

2.1 Steekproef en variabelen

Voor het empirische deel van deze studie maken we gebruik van gegevens die in het schooljaar 1994/95 zijn verzameld in het kader van het cohortonderzoek Primair Onderwijs ('PRIMA'). Het betreft enerzijds de resultaten op een taal- en rekentoets die de leerlingen in groep 4 hebben gemaakt en anderzijds informatie over de leerlingen en hun thuissituatie die is verzameld met behulp van een schriftelijke vragenlijst voor de ouders. Opgemerkt moet worden dat de items in de oudervragenlijst niet specifiek zijn geformuleerd met het oog op het onderhavige onderwerp 'cultureel kapitaal'. Het betreft hier namelijk secundaire analyses op bestaande datasets (maar daarin wijkt dit onderzoek niet af van verreweg de meeste van de hierboven aangehaalde studies). Wat de respons op de oudervragenlijst aangaat, kan worden gemeld dat die voor de totale steekproef 67% bedroeg ten opzichte van de door de scholen opgegeven aantallen leerlingen die bij aanvang van het schooljaar in groep 4 zaten. Echter, omdat de eigenlijke dataverzameling enkele maanden later viel en er in de tussentijd nog verschillende mutaties in de leerlingpopulatie op de scholen plaatsvonden, is het aannemelijk dat de feitelijke respons hoger ligt. Dit neemt niet weg dat er een samenhang is tussen de sociaal-etnische achtergrond van de leerlingen en de respons op de vragenlijst (die in het Nederlands was gesteld, met een toelichting in het Turks en Arabisch). De respons onder de

laagopgeleide allochtonen bijvoorbeeld was 45%, terwijl die onder de hbo- en universitair-opgeleide ouders 82% was. Dit is overigens een relatief hoge respons; voor deelname aan meer-kwalitatief georiënteerd onderzoek worden beduidend lagere responspercentages gemeld (vgl. Leseman & Van den Boom, 1999). Omdat het in dit artikel niet gaat om het geven van een representatieve beschrijving, maar om het in kaart brengen van de systematische relaties tussen een aantal predictoren en afhankelijke variabelen, is dit minder problematisch (Roelandt, Martens & Veenman, 1990; Gijsberts, 1993). Voor een uitgebreide beschrijving van de opzet van PRIMA, de gehanteerde onderzoeksinstrumenten en constructie van variabelen verwijzen we naar Van Langen, Vierke & Robijns (1996), Driessen & Haanstra (1996) en Jungbluth, Van Langen, Peetsma & Vierke (1996).

Een centrale variabele in onze studie is de etnische herkomst van de leerlingen. Deze is bepaald op basis van het geboorteland van de ouders, waarbij alleen 'etnisch homogene' (inclusief eenouder) gezinnen zijn meegenomen. Vanwege getalsmatige overwegingen is een beperking aangebracht tot Nederlandse, Surinaamse en Antilliaanse, Turkse, en Marokkaanse gezinnen met respectievelijk 7531, 282, 515 en 415 cases. In deze studie zijn we op zoek naar verschillen die duiden op een *systematisch effect* tussen etnische groepen. Het gaat om de vraag of de gemiddelden, respectievelijk samenhangen voldoende verschillen tussen de groepen, onafhankelijk van het aantal eenheden waarop elk van de gemiddelden, respectievelijk samenhangen is gebaseerd. We zullen daarom gebruik maken van analyses waarbij de omvang van de groepen gelijk wordt gehouden. We hanteren daarvoor het principe van afwegen van de vier groepen, waarbij de kleinste groep, de Surinaamse gezinnen (N=282), een gewicht van 1 krijgt en de overige groepen een gewicht van 282/groepsaantal. We krijgen zo een totale gewogen N van $4 \cdot 282 = 1128$ leerlingen.³

Uit de literatuurstudie die we in de inleiding hebben beschreven, kwam een serie kenmerken naar voren die van empirisch belang (kunnen) zijn. De PRIMA-bestanden bevatten een aantal van deze kenmerken. Wat het kapitaal, c.q. de hulpbronnen betreft is een eerste indeling die van financiële en culturele hulpbronnen. Vervolgens kunnen de culturele hulpbronnen nog verder opgesplitst worden in linguïstische hulpbronnen, leesgedrag en pedagogisch gezinsklimaat. We veronderstellen dat deze 'kapitaal'-kenmerken een mediërende positie innemen bij de verklaring van de taal- en rekenprestaties door het sociale herkomstmilieu. We gaan er daarbij van uit dat er in deze samenhangen verschillen optreden tussen de onderscheiden etnische groepen. Hierna geven we allereerst een kort overzicht van de gebruikte instrumenten en de daaruit afgeleide variabelen; bij de beschrijving van de variabelen zelf in paragraaf 2.3 zullen we uitgebreider ingaan op de constructie ervan.

Toetsen

Er zijn twee toetsen afgenomen bij de leerlingen, een taal- en een rekentoets.

Oudervragenlijst

De vragen die de ouders hebben beantwoord handelen onder meer over het sociaal-etnische herkomstmilieu en over culturele, linguïstische en opvoedkundige aspecten van de gezinsituatie.

- Bij het *herkomstmilieu* gaat het om: etnische herkomst, opleiding vader en moeder en beroep vader.
- De *financiële hulpbronnen* worden geïndiceerd via het inkomen van vader en moeder.
- Binnen de *culturele hulpbronnen* kunnen drie elementen worden onderscheiden:

Linguïstische hulpbronnen. Indicators hiervoor zijn: taalkeuze (de spreektaal van kinderen in een viertal domeinen), taalattitude (de spreektaal van de ouders onderling in combinatie met de waarde die ouders hechten aan het leren van die taal door hun kinderen), taalbeheersing ouders (beheersing Nederlands wat betreft de vier taalmodaliteiten).

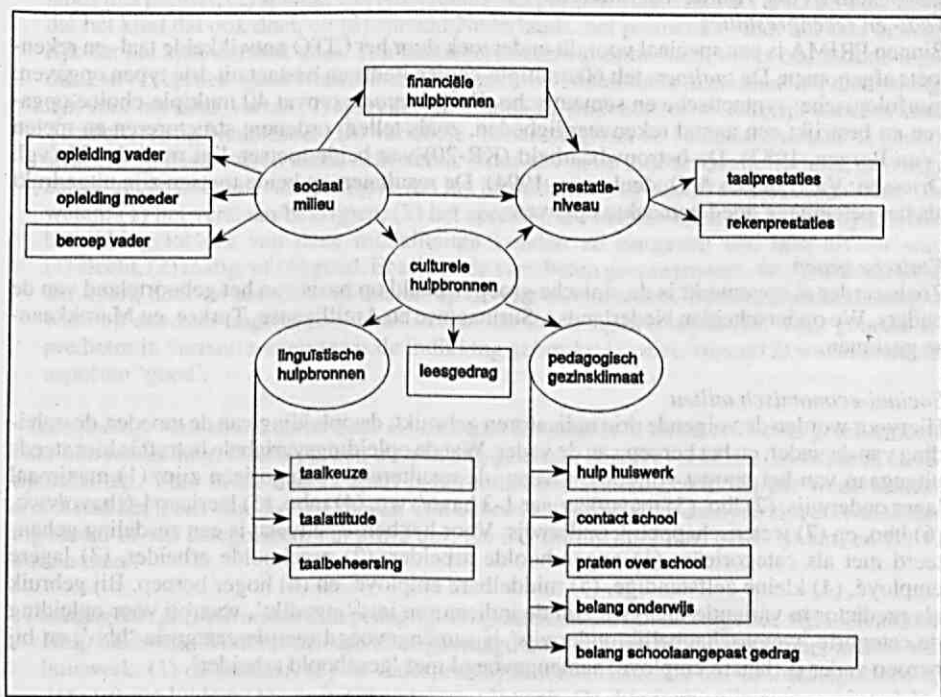
Leesgedrag van ouders. Dit wordt geïndiceerd met vragen naar de frequentie waarmee vader en moeder boeken, kranten en tijdschriften lezen.

Pedagogisch gezinsklimaat. Daaronder vallen met name aspecten die gerelateerd zijn aan het onderwijs. Allereerst drie aspecten die concreet met het onderwijs te maken hebben: hulp bij het maken van huiswerk (door vader, moeder, siblings), contact met school (frequentie waarmee ouders contact hebben met school, door de school georganiseerd dan wel op eigen initiatief), en het praten over school (frequentie waarmee ouders met het kind praten over dingen die op school gebeuren). Daarnaast twee aspecten die in attitudinale zin met het onderwijs te maken hebben: het belang van onderwijs (een serie items m.b.t. het belang dat ouders hechten aan onderwijs), en belang van schoolaangepast gedrag (een serie items m.b.t. het gedrag van de leerling t.o.v. school).

Om het geheel aan relaties wat inzichtelijker te maken hebben we de kenmerken en de veronderstelde relaties daartussen in Figuur 1 schematisch afgebeeld in de vorm van een basis-model.

2.2 Analyse-opzet

In dit onderzoek gaan we in op de vraag of er binnen verschillende etnische groepen een samenhang bestaat tussen het sociaal-economische herkomstmilieu en de onderwijsprestaties van kinderen. Uit eerder onderzoek is bekend dat er, over groepen heen, sprake is van een redelijke samenhang tussen sociaal milieu en prestaties (Driessen, 1997; Mulder, 1996). De vraag is nu of dit een samenhang is die ook op gezinsniveau, dus binnen de groepen, is terug te vinden, of dat de samenhang over groepen heen (deels) ontstaat door de samenhang tussen de gemiddelden op sociaal milieu en prestaties. In dat geval ontstaat de totale samenhang



Figuur 1: Basismodel voor de toetsing van de cultureel kapitaal-thesen.

doordat sommige groepen gemiddeld hoog scoren op zowel sociaal milieu als prestaties en andere groepen op beide gemiddeld laag scoren, terwijl er binnen die groepen geen samenhang hoeft te bestaan tussen sociaal milieu en prestaties.

Ook wordt wel verondersteld dat, in een totaalanalyse over groepen heen, culturele hulpbronnen een mediërend effect hebben op de samenhang tussen sociaal milieu en prestaties. Met andere woorden: onder constanthouding van hulpbronnen 'vervalt' de samenhang tussen sociaal milieu en prestaties. Er is dan geen directe werking van sociaal milieu op prestaties. Ook hier is het weer de vraag of deze mediëring binnen groepen is terug te vinden of dat de mediëring mede ontstaat door samenhangen tussen gemiddelden op groepsniveau. In dat geval zouden ook de samenhangen tussen milieu en hulpbronnen en tussen hulpbronnen en prestaties binnen groepen afwezig kunnen zijn, zodat dan van mediëring geen sprake kan zijn.

We willen dan ook eerst onderzoeken of de samenhangen van sociaal milieu en hulpbronnen met toetsprestaties voor onderscheiden etnische groepen verschillen. Het gaat daarbij dus om het interactieve effect van etnische groep en een predictor (sociaal milieu, resp. hulpbronnen) op toetsprestaties. We zullen inzicht proberen te krijgen in het bestaan van deze interactie-effecten via twee-weg variantie-analyse met herkomstland en steeds een van de predictoren als de twee factoren in de voorspelling van de taal- en rekenprestaties. Als er geen interactief effect bestaat, dan kunnen we besluiten dat de samenhang van sociaal milieu en hulpbronnen met toetsprestaties niet systematisch anders is in de verschillende groepen. In dat geval kunnen we de samenhang van een predictor met prestaties onder constanthouding van herkomstland, interpreteren als de gemiddelde binnengroepssamenhang tussen predictor en prestaties. We zullen vervolgens onder meer multipel regressies uitvoeren om inzicht te krijgen in de mate van mediërende werking van de hulpbronnen.

2.3 Beschrijving van de variabelen

Taal- en rekenprestaties

Binnen PRIMA is een speciaal voor dit onderzoek door het CITO ontwikkelde taal- en reken-toets afgenomen. De *taaltoets* telt 60 multiple-choice items en bestaat uit drie typen opgaven: morfologische, syntactische en semantische. De *rekentoets* omvat 40 multiple-choice opgaven en bestrijkt een aantal rekenvaardigheden, zoals tellen, ordenen, structureren en meten (Van Bergen, 1987). De betrouwbaarheid (KR-20) van beide toetsen ligt rond de .85 (vgl. Driessen, Van Langen & Oudenhoven, 1994). De resultaten op beide toetsen zijn uitgedrukt als het percentage goed gemaakte opgaven.

Etnische groep

Zoals eerder al opgemerkt is de etnische groep bepaald op basis van het geboorteland van de ouders. We onderscheiden Nederlandse, Surinaamse en Antilliaanse, Turkse, en Marokkaanse gezinnen.

Sociaal-economisch milieu

Hiervoor worden de volgende drie indicatoren gebruikt: de opleiding van de moeder, de opleiding van de vader, en het beroep van de vader. Wat de opleidingsvariabele betreft is hier steeds uitgegaan van het hoogst voltooide niveau; de resulterende categorieën zijn: (1) maximaal lager onderwijs, (2) lbo, (3) mavo/leerjaar 1-3 havo/vwo, (4) mbo, (5) leerjaar 4-6 havo/vwo, (6) hbo, en (7) wetenschappelijk onderwijs. Voor het beroepsniveau is een zesdeling gehanteerd met als categorieën: (1) ongeschoolde arbeider, (2) geschoolde arbeider, (3) lagere employé, (4) kleine zelfstandige, (5) middelbare employé, en (6) hoger beroep. Bij gebruik als predictor in variantie-analyses zijn de indicatoren iets 'ingedikt', waarbij voor opleiding de categorie 'wetenschappelijk onderwijs' is samengevoegd met de categorie 'hbo', en bij beroep vader is 'lagere employé' samengevoegd met 'geschoold arbeider'.

Hulpbronnen

Deze zijn op de eerste plaats onder te verdelen in financiële en culturele hulpbronnen, en de culturele hulpbronnen vervolgens weer in: linguïstische hulpbronnen, leesgedrag, en pedagogisch gezinsklimaat.

• *Financiële hulpbronnen.* Voor beide ouders is gevraagd naar het inkomen. Hun totale inkomen is vervolgens gedeeld door het aantal personen in het gezin. Het gemiddelde inkomen per gezinslid interpreteren we als het gemiddelde te besteden bruto-bedrag per gezinslid. Bij gebruik als predictor in variantie-analyses hebben we het inkomen ingedikt tot de volgende categorieën: (1) tot f 500, (2) f 500 - f 700, (3) f 700 - f 900, (4) f 900 - f 1100, en (5) f 1100 en hoger.

• *Linguïstische hulpbronnen.* Deze zijn vooral van belang in het geval van allochtone gezinnen, maar daarnaast ook in niet-standaard-Nederlands sprekende autochtone gezinnen (dialect, Fries). De volgende drie variabelen worden onderscheiden:

- *Taalkeuze.* Middels vier items is geïnformeerd naar het aantal domeinen waarin het kind Nederlands (vs. Nederlands dialect, Fries of een buitenlandse taal) spreekt, te weten: (1) met moeder, (2) met vader, (3) met broers/zussen, en (4) met vrienden/vriendinnen. De invalshoek is hier dus de taalsituatie van het kind. De score op deze variabele is berekend door het aantal malen dat Nederlands wordt gesproken te sommeren. Voor de variantie-analyses is deze variabele ingedikt tot: (1) 0-1, (2) 2-3, en (3) 4 domeinen.

- *Taalattitude.* In tegenstelling tot de variabele 'taalkeuze', is hier het vertrekpunt de taalsituatie van de ouders. Op basis van een combinatie van twee items hierover is een variabele geconstrueerd met de volgende categorieën onderscheiden: (1) spreekt geen Nederlands met partner, (2) spreekt wel Nederlands met partner, maar vindt het niet zeer belangrijk dat het kind dat ook doet, en (3) spreekt Nederlands met partner en vindt het zeer belangrijk dat het kind dat ook doet. Ten behoeve van de variantie-analyses is taalattitude ingedikt tot: (1) spreekt geen Nederlands of spreekt wel Nederlands maar vindt het niet belangrijk voor het kind, versus (2) spreekt Nederlands en vindt het zeer belangrijk voor het kind.

- *Taalbeheersing Nederlands.* De variabele 'beheersing van de Nederlandse taal' is geconstrueerd op basis van in totaal acht items. Voor zowel de moeder als de vader is gevraagd aan te geven hoe hun beheersing is van het Nederlands wat betreft de vier modaliteiten, te weten: (1) het verstaan/begrijpen, (2) het spreken, (3) het lezen, en (4) het schrijven. Met betrekking tot elk van deze modaliteiten konden ze aangeven wat hun niveau was: (1) slecht, (2) matig, of (3) goed. Er zijn twee variabelen geconstrueerd, voor vader en moeder apart, door de scores te middelen op de vier modaliteiten. Betrouwbaarheidsanalyse wees op een zeer hoge interne consistentie; de alfa's lagen rond de .95. Voor gebruik als predictor in variantie-analyses is de indikking gebruikt: (1) niet, versus (2) wel allé vier de aspecten 'goed'.

• *Leesgedrag.* Aan zowel de moeder als de vader is middels in totaal zes items geïnformeerd naar hun leesgedrag. Daarbij is een uitsplitsing gemaakt naar: (1) boeken, (2) kranten, en (3) tijdschriften. Voor beide ouders apart is het totaal aantal uren geteld dat per week besteed wordt aan het lezen van die drie media. Ten behoeve van variantie-analyse zijn de variabelen ingedeeld in vijf categorieën: (1) t/m 2 uur, (2) 3-5 uur, (3) 6-7 uur, (4) 8-10 uur, en (5) 11 uur en meer.

• *Pedagogisch gezinsklimaat.* Het pedagogisch gezinsklimaat telt de volgende vijf indicatoren:

- *Hulp huiswerk.* Middels drie items is gevraagd naar wie het kind helpt bij het maken van huiswerk: (1) de moeder, (2) de vader, en (3) broers/zussen. De antwoordcategorieën op deze vragen luiden: (1) nooit, (2) soms, en (3) vaak. De drie items zijn als aparte variabelen in de analyses opgenomen.

- *Contact met school.* Over de contacten met de school zijn twee vragen gesteld. De eerste betreft het bezoek van speciale bijeenkomsten (ouderavonden, rapportbesprekingen, etc.), met als categorieën: (1) (bijna) nooit, (2) meestal, en (3) altijd. De tweede vraag betreft het op eigen initiatief met de leerkracht gaan praten, met als categorieën: (1) nooit, (2) 1-2 keer per jaar, en (3) 3 of meer keer per jaar. Voor de variantie-analyses is 'bijna nooit' samengevoegd met 'meestal'.
- *Praten over school.* Dit betreft een vraag over de frequentie waarmee de ouders met het kind praten over dingen die op school gebeuren, met als categorieën: (1) minder dan 1 keer per week, (2) 1 keer per week, (3) een paar keer per week, en (4) (bijna) elke dag. Ten behoeve van de variantie-analyses is 'minder dan 1 keer per week' samengevoegd met '1 keer per week'.
- *Belang onderwijs.* Hierbij is middels een item gevraagd of de ouder het eens is met de uitspraak dat het kind zo lang mogelijk naar school zou moeten gaan, waarbij de antwoordcategorieën zijn: (1) helemaal niet mee eens, (2) niet mee eens, (3) niet/wel me eens, (4) wel mee eens, en (5) helemaal mee eens. Voor variantie-analyse is 'helemaal niet' samengevoegd met 'niet'.
- *Belang schoolaangepast gedrag.* Dit betreft een factorscore, die tot stand is gekomen door middeling van de scores op vijf items die het belang aangeven wat door de ouder wordt gehecht aan: 'gehoorzaam zijn', 'hard werken op school', 'netjes werken', 'zich gedragen zoals alle andere kinderen', 'goed luisteren naar wat volwassenen zeggen'. De antwoordmogelijkheden varieerden van (1) enigszins belangrijk, tot (5) zeer belangrijk. Betrouwbaarheidsanalyse leverde een alfa op van .76. Voor variantie-analyse is deze variabele ingedikt tot vijf categorieën, namelijk: (1) <2.6, (2) 2.6-2.8, (3) 2.8-3.2, (4) 3.2-3.4, en (5) >3.4.

Om een indruk te geven van de samenhangen tussen de verschillende indicatoren van hulpbronnen, hebben we ter informatie in Bijlage 1 de correlaties opgenomen. Tevens hebben we in die bijlage de gemiddelden en standaarddeviaties vermeld. In Bijlage 2 geven we de verdelingskenmerken van de milieu-indicatoren en de samenhangen die er tussen bestaan.

3. RESULTATEN

3.1 Verschillen in verdelingen tussen etnische groepen

We starten met het geven van een eerste beschrijving van de variabelen. Daartoe presenteren we in Tabel 1 de verdelingen van alle kenmerken over de etnische groepen. We maken hier gebruik van de ingedikte dan wel gedichotomiseerde predictoren; zie hiervoor paragraaf 2.3.

Tabel 1 laat zien dat er dikwijls sprake is van aanzienlijke verschillen tussen de vier groepen. Wat het opleidingsniveau betreft valt op hoe weinig Marokkaanse en Turkse ouders op mbo- of hoger niveau zitten; bij de Marokkanen is dat slechts 5 à 6%. Voor het beroepsniveau geldt een vergelijkbare situatie. Ook in dat opzicht staan de Marokkanen en Turken er zeer ongunstig voor. Dit keert terug als we naar het besteedbaar bedrag per gezinlid kijken. Bij de Marokkanen is dat minder dan de helft van het bedrag bij de Nederlanders. Dit komt enerzijds door het lage beroepsniveau en anderzijds door het grote aantal personen in het gezin (namelijk 6.2, tegen 4.4 bij de Nederlanders). In Turkse en Marokkaanse gezinnen wordt ook zeer weinig Nederlands gesproken, en bovendien wordt het ook niet zo belangrijk geacht dat het Nederlands geleerd wordt. De beheersing van het Nederlands is binnen de Turkse en Marokkaanse gezinnen ook vrij laag, en dan met name het niveau van de moeders. In Turkse, maar vooral in Marokkaanse gezinnen wordt door de ouders ook vrij weinig gelezen in hun vrije tijd. Ook hier zijn het met name de moeders die laag scoren. Ongetwijfeld heeft dit ook te maken met het feit dat het vooral de moeders zijn uit die etnische groepen die analfabeet zijn. Wat de hulp bij het maken van huiswerk betreft vallen enkele dingen op. Op de eerste plaats wordt

Tabel 1: Verschillen tussen etnische groepen (in % en gemiddelden).

	etnische groep				totaal	Eta	p
	Ned	Sur	Trk	Mar			
% mbo of hoger, moeder	44	21	8	5	20	.39	.000
% mbo of hoger, vader	53	28	18	6	27	.40	.000
% middelbaar/hoger beroep, vader	44	28	8	6	21	.40	.000
besteedbaar bedrag per gezinslid	870	720	660	400	663	.42	.000
% in vier domeinen Nederlands spreken	76	57	6	13	38	.61	.000
% belangrijk goed Nederlands spreken	66	28	5	3	25	.58	.000
% volledige beheersing Nederlands, moeder	88	80	17	16	50	.68	.000
% volledige beheersing Nederlands, vader	85	77	28	23	51	.57	.000
uren lezen per week, moeder	7.3	7.1	4.1	2.0	5.3	.38	.000
uren lezen per week, vader	6.5	7.2	5.6	3.2	5.6	.25	.000
% vaak huiswerkhulp van moeder	50	51	23	15	37	.33	.000
% vaak huiswerkhulp van vader	22	31	27	11	22	.17	.009
% soms/ vaak hulp van broer/zus	36	61	79	84	68	.41	.000
% altijd naar ouderbijeenkomsten	73	67	50	49	60	.21	.000
% meer dan 2 keer per jaar praten met leraar	26	38	42	31	34	.13	.001
% elke dag praten over school	82	74	51	51	66	.29	.000
% lang naar school belangrijk	29	63	62	68	55	.31	.000
% schoolaangepast gedrag belangrijk	35	66	73	74	62	.33	.000
taalprestaties	80	72	62	66	70	.49	.000
rekenprestaties	80	70	70	70	72	.25	.000

die vaker gegeven door de moeders dan door de vaders. Wellicht heeft dit te maken met het feit dat de vaders buitenshuis werken en de moeders thuis zijn. Het minst vaak wordt huiswerkhulp gegeven door de Marokkaanse ouders. Vervolgens valt op dat het in de allochtone gezinnen vooral de (oudere) broers en zussen zijn die helpen bij het maken van huiswerk. Binnen de Nederlandse gezinnen komt dat aanzienlijk minder vaak voor. Broers en zussen kunnen binnen allochtone gezinnen waarschijnlijk compensatie bieden voor het gegeven dat hun ouders zelf vaak nauwelijks opleiding hebben gehad, en ook het Nederlands niet goed beheersen (vgl. ook Crul, 2000). Het percentage Turkse en Marokkaanse ouders dat altijd naar ouderbijeenkomsten gaat ligt veel lager dan het percentage Nederlandse en Surinaamse en Antilliaanse ouders. Wat betreft het met de leerkracht gaan praten, zijn er nauwelijks verschillen tussen de vier etnische groepen. Die zijn er weer wel als het gaat om het elke dag over school praten. Dat gebeurt in Turkse en Marokkaanse gezinnen aanzienlijk minder vaak dan in de andere gezinnen. Opmerkelijk zijn de bevindingen met betrekking tot het belang dat wordt gehecht aan het lang naar school gaan. Tussen de drie allochtone groepen is er nauwelijks sprake van verschillen; het zijn echter juist de Nederlandse ouders die laag scoren op dit kenmerk (vgl. Ledoux, Deckers, De Bruijn & Voncken, 1992). Ook ten aanzien van het belang dat de ouders hechten aan schoolaangepast gedrag ('conformisme') zijn er niet zo'n grote verschillen tussen de drie allochtone groepen: deze vinden dat allemaal vrij belangrijk. Nederlandse ouders daarentegen vinden dat veel minder belangrijk (vgl. Nijsten, 1998; Pels, 1998). Samenvattend kan wat betreft de milieu- en kapitaal-variabelen worden geconcludeerd dat met name bij sociaal milieu, financiële en linguïstische hulpbronnen de Surinaamse gezinnen een tussenpositie innemen tussen enerzijds de Nederlandse en anderzijds de Turkse en Marokkaanse gezinnen. Dan de toetsprestaties. Voor de totale groep is het taalgemiddelde 70 met een standaarddeviatie van 14; het rekengemiddelde bedraagt 72 met een standaarddeviatie van 17. Maar er doen zich forse verschillen voor tussen de etnische groepen, met name ten aanzien van de taalprestaties. Gemiddeld genomen maken de Nederlandse kinderen 80% van de zowel de taal- als rekenopgaven goed. De taalprestaties van de drie allochtone groepen, en dan opnieuw van met name Turkse en Marokkaanse kinderen, liggen beduidend lager. Met

betrekking tot de rekenprestaties zijn er minder grote verschillen; binnen de drie allochtone groepen zijn er zelfs helemaal geen verschillen.

3.2 Verschillen in samenhangen tussen etnische groepen

In Tabel 2 presenteren we voor elk van de predictoren de samenhang met de taal- en rekenprestaties. Voor iedere categorie van de predictoren geven we de gemiddelde toetsprestatie, in de eerste regel voor taal en in de tweede regel voor rekenen. De cijfers 1 - 6 in de kop van de tabel duiden op de (wisselende aantallen) categorieën van de predictoren; welke labels daarbij horen staat in paragraaf 2.3 uitgelegd.

Uit Tabel 2 komt naar voren dat de *totale* samenhang van bijna alle predictoren met taal significant en relevant is, maar voor rekenen op zijn best net relevant (Eta zelden boven .15). De gemiddelde *binnengroepssamenhang* is echter zowel voor taal als voor rekenen nergens zelfs maar significant en in ieder geval niet relevant. Uitzondering vormt de samenhang tussen taalkeuze en taalprestaties, waarbij het effect net relevant is. Dat we de gemiddelde binnengroepssamenhang mogen interpreteren als 'de samenhang zoals die in elk van de groepen

Tabel 2: Totale en gemiddelde binnengroepssamenhangen tussen predictoren en toetsprestaties (eerste regel: taal; tweede regel: rekenen)

	totale samenhang								gemiddelde binnengroepssamenhang								interactie groep* predictor		
	1	2	3	4	5	6	Eta	p	1	2	3	4	5	6	Eta	p	totaal	Eta	p
opleiding moeder	65	72	73	76	75	80	.34	.000	69	70	71	72	72	74	.11	.031	70	.10	.606
	70	72	74	77	76	81	.18	.000	72	71	74	75	74	77	.09	.101	73	.10	.744
opleiding vader	66	71	70	74	74	78	.29	.000	69	69	69	70	72	73	.10	.124	70	.13	.118
	71	73	73	75	77	78	.14	.001	73	72	73	73	76	75	.06	.639	73	.12	.451
beroep vader	65	70	73	75	79		.32	.000	69	70	69	70	73		.10	.047	70	.11	.243
	71	73	73	77	79		.15	.000	73	73	71	74	76		.06	.468	73	.13	.141
besteding	66	70	73	74	74		.24	.000	69	70	71	71	72		.09	.093	70	.10	.354
	70	73	75	76	75		.14	.000	71	72	73	74	74		.06	.453	72	.11	.343
taalkeuze	64	67	76				.38	.000	67	70	72				.15	.000	70	.07	.379
	70	71	76				.18	.000	70	72	75				.09	.007	72	.06	.604
taalattitude	67	77					.32	.000	69	71					.07	.031	70	.02	.829
	71	77					.13	.000	72	73					.02	.436	72	.02	.898
taalbeheersing moeder	65	75					.35	.000	69	71					.07	.026	70	.09	.013
	70	75					.13	.000	72	73					.02	.498	73	.09	.037
taalbeheersing vader	65	74					.30	.000	69	70					.04	.204	70	.05	.312
	71	75					.11	.000	73	73					.00	.974	73	.07	.150
lezen moeder	66	70	74	74	76		.27	.000	69	70	72	72	73		.11	.019	70	.08	.749
	71	73	75	74	74		.10	.037	73	73	74	73	73		.02	.974	73	.11	.314
lezen vader	67	70	72	72	74		.18	.000	69	70	71	70	73		.11	.047	70	.09	.657
	72	73	75	75	75		.08	.157	73	72	74	74	75		.05	.748	73	.09	.798
huiswerkhulp moeder	66	70	73				.19	.000	70	70	71				.04	.617	70	.10	.307
	72	72	72				.00	.996	75	72	71				.07	.315	72	.06	.926
huiswerkhulp vader	69	70	69				.04	.644	71	69	69				.08	.328	70	.06	.917
	74	73	70				.07	.356	75	72	70				.11	.099	73	.06	.969
huiswerkhulp siblings	73	68	65				.23	.000	69	69	68				.04	.774	69	.15	.068
	74	74	69				.12	.045	72	74	71				.07	.363	73	.06	.948
contact school	67	72					.15	.000	69	71					.06	.069	70	.01	.985
	71	74					.07	.012	72	73					.04	.217	73	.05	.416
contact school zelf	70	71	69				.07	.050	70	70	70				.02	.866	70	.04	.935
	74	74	70				.10	.002	74	73	71				.08	.021	73	.04	.955
praten school	65	69	72				.20	.000	68	71	71				.07	.079	71	.05	.730
	69	74	74				.09	.013	71	75	73				.07	.097	73	.05	.853
belang school	73	76	71	68			.17	.000	70	71	70	70			.03	.850	70	.06	.858
	74	78	73	71			.11	.004	72	74	73	72			.03	.873	73	.09	.470
schoolaangepast gedrag	77	74	70	69	67		.24	.000	73	71	70	70	69		.09	.054	70	.07	.783
	77	76	72	72	71		.14	.000	75	74	72	72	72		.06	.408	72	.08	.856

bestaat' laat de sterkte van de interactie-effecten etnische groep*predictor op prestaties zien: geen van de interactie-effecten is significant. De samenhangen van predictoren met prestaties wijken dus tussen etnische groepen niet significant van elkaar af. Ze zijn namelijk (bijna) allemaal niet-systematisch afwijkend van nul; geen van de gemiddelde binnengroepssamenhangen is immers significant (behalve taalkeuze op taalprestaties).

Met name de gegevens met betrekking tot het sociaal milieu duiden er op dat het in eerder onderzoek veronderstelde 'totale effect' van sociaal milieu op prestaties mogelijk het gevolg is van een niveau-interpretatiefout. Door alleen naar de *totale* steekproef te kijken, wordt namelijk de indruk gewekt dat dit effect op gezinsniveau ligt, dat wil zeggen: gemiddeld binnen etnische groepen. Deze effecten bestaan echter slechts maar zeer gedeeltelijk op het niveau van het gezin en meer op het niveau van de etnische groepen. Nederlandse gezinnen hebben namelijk zowel *gemiddeld* een hoger sociaal milieu *alsook gemiddeld* een hoger prestatieniveau van de kinderen, met name wat betreft taal, dan de overige etnische groepen. Dit leidt tot de totale positieve samenhang tussen sociaal milieu en prestaties. Binnen de groepen bestaat deze samenhang echter nauwelijks. Hier dient echter het volgende bij opgemerkt te worden. Een probleem dat zich bij het bestuderen van samenhangen tussen genoemde variabelen kan voordoen betreft de kwestie van de 'restriction of range' (vgl. voor deze problematiek ook Driessen, 1995). Toch geldt voor het onderhavige onderzoek dat de verschillen in standaarddeviaties bij de milieu-indicatoren niet dramatisch zijn; zie hiervoor Bijlage 2. Bovendien blijkt dat de kleinste standaarddeviatie van 1.1 (opleiding vader Marokkaanse groep) nog op voldoende variatie duidt om samenhangen mogelijk te maken. Immers, 95% van de respondenten zit verspreid over $2 * 1.96 * 1.1 = 4.5$ punten bij een range van 1 - 7.

Bij de interpretatie van de mediërende werking van hulpbronnen doen zich soortgelijke problemen voor. Zeker als het gaat om zo sterk verschillende cultureel-etnische groepen als in dit artikel ligt het voor de hand dat er sprake zal zijn van een sterke contaminatie van hulpbronnen met etniciteit. Doordat de hulpbronnen binnen groepen niet samenhangen met prestaties, is het binnen etnische groepen, dat wil zeggen op gezinsniveau, dan ook zeer onwaarschijnlijk dat er sprake kan zijn van mediëring. Dan zou er binnen die groepen sprake moeten zijn van suppressor-effecten waardoor, ondanks de afwezigheid van een totale samenhang, er wel een relevant direct effect van hulpbronnen op prestaties zou zijn, hetgeen een voorwaarde voor mediëring is. Het feit dat een bepaalde predictor binnen groepen geen enkele variatie heeft, en de samenhang met prestaties dus per definitie nul is, is daarbij in methodologisch opzicht geen probleem. Als er in een dergelijke situatie wel een totale samenhang is, kan die alleen ontstaan door de ligging van de gemiddelden van predictor en prestaties. De *interpretatie* van die samenhang op gezinsniveau is dan een niveaufout. We zullen in de afsluitende paragraaf van dit artikel op deze problematiek terugkomen.

In het voorgaande hebben we gebruik gemaakt van variantie-analyses om de interactieve- en gemiddelde binnengroepssamenhangen te analyseren. Het betreft dus steeds analyses van de *ongestandaardiseerde* samenhangen. Dit is nodig omdat variantieverschillen tussen groepen in de modelvariabelen de directe vergelijkbaarheid van gestandaardiseerde samenhangen verstoren. Met ditzelfde doel hebben we ook nog multi-sample LISREL-analyses uitgevoerd, waarbij, na een 'overall' toets op het hele model, steeds elke afzonderlijke ongestandaardiseerde padcoëfficiënt is onderzocht op tussengroepsverschillen (zie: Driessen, Doesborgh & Claassen, 1999). Deze analyses leveren natuurlijk geen andere resultaten op dan de variantie-analyses wat betreft de interactieve werkingen van milieu en hulpbronnen op prestaties. Immers, als er geen totale interactie is (variantie-analyse), kan er ook geen lineaire interactie (verschillen in ongestandaardiseerde regressiecoëfficiënten) zijn. Alleen de interactieve werking van etnische groep en sociaal milieu op hulpbronnen zijn daar 'extra' onderzocht. Om een beter overzicht te krijgen in de *verklarende werking* van de afzonderlijke onderdelen van het model, en met name de tussengroepsverschillen hierin, zullen we hierna enkele regressie-analyses en tabellen met (partiële en multiële) correlaties presenteren.

3.3 De mediërende werking van hulpbronnen

Om meer inzicht te krijgen in de mediërende werking van de hulpbronnen hebben we multiple regressies uitgevoerd met sociaal milieu als predictor van de taal- en rekenprestaties. Als we daarna ook de hulpbronnen als predictoren toevoegen, kunnen we de zelfstandige werking van sociaal milieu berekenen, wat inzicht geeft in de mediërende werking van de hulpbronnen. Om de analyses overzichtelijk te houden, hebben we de drie indicatoren voor sociaal milieu samengenomen door middeling van de drie variabelen, waarbij voor 'opleiding' de categorie 'wetenschappelijk onderwijs' eerst is samengevoegd met 'hbo'. De gegevens in Bijlage 2 geven aan dat de constructie van een dergelijke milieuvariabele verantwoord is. Tabel 3 geeft de samenhangen tussen sociaal milieu en de taal- en de rekenprestaties, alsmede de samenhangen tussen de taal- en rekenprestaties.

Tabel 3: Samenhangen sociaal milieu met toetsprestaties en taal met rekenen (correlaties).

	taal	rekenen	taal*rek
totaal	.34	.17	.44
Nederlands	.27	.23	.41
Surinaams	.14	.05	.40
Turks	-.02	-.04	.37
Marokkaans	-.04	-.05	.40

Tabel 3 maakt duidelijk dat de correlaties tussen sociaal milieu en taalprestaties in alle groepen kleiner zijn dan de totaalcorrelatie. De totaalcorrelatie wordt dus voor een deel 'veroorzaakt' door de samenhang van de groepsgemiddelden. Dus: in Nederlandse gezinnen is zowel sprake van een relatief hoog sociaal milieu, alsook van een relatief hoge toetsscore van de kinderen, terwijl met name in Turkse en Marokkaanse gezinnen het omgekeerde geldt (zie ook Tabel 2). Voor de rekenprestaties geldt dit in mindere mate, met name door de kleinere verschillen tussen de groepen in gemiddelde rekenscore. Alleen binnen de Nederlandse gezinnen is er sprake van een net-relevante samenhang op gezins/leerlingniveau tussen sociaal milieu en toetsprestaties. De correlatie tussen de taal- en rekenprestaties is de enige samenhang, van alle samenhangen in dit onderzoek, die in alle groepen bijna even groot is als de totale correlatie en waar dus geen sprake is van een correlatie die gedeeltelijk ontstaat door samenhang van de groepsgemiddelden.

Om een eerste indruk te krijgen van de lineaire samenhangen tussen de variabelen, zullen we eerst de correlaties tussen sociaal milieu en hulpbronnen (Tabel 4) en tussen hulpbronnen en toetsscores (Tabel 5) presenteren. Aanvankelijk was het onze bedoeling de indicatoren voor de onderscheiden hulpbronnen samen te voegen tot latente constructen, dit overeenkomstig het model zoals afgebeeld in Figuur 1. Uit de correlaties tussen deze indicatoren bleek echter dat de betreffende samenhangen onvoldoende sterk waren, met name met betrekking tot het pedagogisch gezinsklimaat (zie hiervoor Bijlage 1). We hebben daarom besloten de indicatoren als afzonderlijke variabelen in de analyses op te nemen. Een voordeel hiervan is overigens dat een beter zich wordt verkregen op de werking van de afzonderlijke kapitaal-indicatoren.

In Tabel 4 zien we dat er redelijk hoge totaal-correlaties bestaan tussen sociaal milieu en met name financiële en linguïstische hulpbronnen. De correlaties met de indicatoren van pedagogisch gezinsklimaat zijn beduidend lager of zijn zelfs niet relevant. Ook kunnen we uit de tabel afleiden dat elke totaal-correlatie voor een deel ontstaan is uit de samenhang van de groepsgemiddelden op sociaal milieu en hulpbron. In bijna alle etnische groepen is de correlatie immers lager dan de totale correlatie. Met name binnen de Turkse en Marokkaanse gezinnen zijn de correlaties dikwijls nog beduidend lager dan binnen de Nederlandse en Surinaamse

Tabel 4: Samenhangen tussen sociaal milieu en hulpbronnen (correlaties).

	totaal	etnische groep			
		Ned	Sur	Trk	Mar
besteding	.46	.41	.39	.17	.19
taalkeuze	.33	.16	.12	.03	-.01
taalattitude	.47	.20	.23	.13	.08
taalbeheersing moeder	.50	.23	.31	.23	.34
taalbeheersing vader	.48	.29	.35	.36	.23
lezen moeder	.35	.22	.25	.08	.27
lezen vader	.28	.30	.26	.11	.17
huiswerkhulp moeder	.24	-.02	.00	.01	.29
huiswerkhulp vader	.19	-.02	-.06	.12	.37
huiswerkhulp siblings	-.29	-.10	.03	-.07	.07
contact school	.20	.10	.11	.12	.06
contact school zelf	.09	.09	.12	.09	.17
praten school	.22	.15	.12	-.02	.07
belang school	-.17	.00	-.02	-.13	-.00
schoolaangepast gedrag	-.28	-.25	-.12	-.10	.03
R	.66	.57	.54	.45	.42

gezinnen. Er zijn hier ook uitzonderingen: binnen Marokkaanse gezinnen wordt het leesgedrag van de moeder en het geven van huiswerkhulp door moeder en vader redelijk bepaald door het sociaal-milieuniveau van het gezin. Dit is binnen Turkse gezinnen niet het geval.

Uit Tabel 5 kan worden opgemaakt dat binnen de etnische groepen de correlaties tussen hulpbronnen en toetsscores vrijwel geheel afwezig zijn. Alleen binnen Surinaamse gezinnen is er nog een redelijke correlatie van taalkeuze (het aantal domeinen waarin het kind Nederlands spreekt), taalbeheersing moeder en huiswerkhulp door broers/zussen met taalprestaties. Dit ligt in de lijn van wat we al op basis van Tabel 2 concludeerden, namelijk dat er wel totaal-samenhangen zijn, met name alweer tussen financiële en linguïstische hulpbronnen en taalprestaties, maar dat de gemiddelde binnengroepssamenhangen vrijwel afwezig zijn.⁴

Tabel 5: Samenhangen tussen hulpbronnen en toetsscores (correlaties).

	totaal	taal				totaal	rekenen			
		Ned	Sur	Trk	Mar		Ned	Sur	Trk	Mar
besteding	.21	-.15	.13	.00	.00	.10	.10	.08	-.06	.00
taalkeuze	.35	.09	.22	.13	.08	.15	.04	.15	.10	-.02
taalattitude	.35	.09	.15	-.00	.01	.16	.04	.09	-.00	.00
taalbeheersing moeder	.34	.12	.18	.05	.02	.11	.05	.14	-.01	-.05
taalbeheersing vader	.29	.12	.14	.05	.00	.13	.07	.11	.02	-.01
lezen moeder	.22	.07	.11	.13	.00	.06	.05	-.00	.02	-.08
lezen vader	.13	.07	.11	.06	.04	.03	.05	-.02	.01	-.03
huiswerkhulp moeder	.19	-.08	.07	.05	.11	.00	-.11	-.09	-.00	-.02
huiswerkhulp vader	.00	-.07	-.01	-.17	-.00	-.06	-.08	-.15	-.16	-.04
huiswerkhulp siblings	-.22	-.06	-.24	.07	.03	-.10	-.01	-.00	-.00	-.02
contact school	.15	.05	.08	.03	.06	.07	.03	.14	.01	-.02
contact school zelf	-.03	-.01	-.04	.00	-.00	-.09	-.07	-.03	-.08	-.12
praten school	.19	.10	.07	.03	.06	.08	.06	.00	.00	.07
belang school	-.13	-.02	-.03	.05	-.07	-.08	-.04	-.03	.06	-.00
schoolaangepast gedrag	-.23	-.10	-.17	.00	-.06	-.13	-.07	-.02	-.04	-.05

Merk ook op dat de samenhangen, voor zover relevant, tussen huiswerkhulp geven door vader of broers/zussen (en bij Surinaamse gezinnen: het belang van schoolaangepast gedrag), met prestaties negatief zijn binnen Surinaamse en Turkse gezinnen. Men zou eventueel kunnen veronderstellen dat het geven van huiswerkhulp door vader of broers/zussen (in causale zin) leidt tot verslechtering van de prestaties, bijvoorbeeld door het geven van verkeerde instructie, maar de alternatieve verklaring ligt meer voor de hand, namelijk dat slechte toetsprestaties de aanleiding zijn voor vader of broers/zussen om huiswerkhulp te gaan geven. De positie van deze variabelen als predictoren van prestaties zou daarmee in twijfel kunnen worden getrokken.

Tot nu toe hebben we steeds gesproken over de *totale* samenhang van sociaal milieu en hulpbronnen met prestaties. Daarbij is geen rekening gehouden met de correlaties die tussen de hulpbronnen bestaan. In de hierna te presenteren regressie-analyses houden we wel rekening met deze samenhangen. Op deze manier kunnen de partiële regressiecoëfficiënten (ofwel zelfstandige of directe effecten) van sociaal milieu en hulpbronnen op toetscores bepaald worden. Tabel 6 geeft allereerst de partiële regressiecoëfficiënten, alsmede de totaal verklaarde variantie in de taal-, respectievelijk rekenprestaties door alle hulpbronnen samen. Ook de toename in verklaring door toevoeging van sociaal milieu is gegeven. Deze toename is direct gerelateerd aan de zelfstandige werking van sociaal milieu op toetsprestaties.

Ondanks de vrijwel algehele afwezigheid van sterke zelfstandige effecten van hulpbronnen op toetsprestaties, is er, met name voor de vier etnische groepen samen en binnen de Surinaamse groep in relatie tot de taalprestaties, sprake van nog redelijk hoge variantieverklaringen. Dit duidt er nogmaals (vgl. Bijlage 1) op dat de hulpbronnen over het algemeen niet sterk onderling gecorreleerd zijn. De variantieverklaring is dan ook opgebouwd uit vele kleine bijdragen van de hulpbronnen, in plaats van een verklaring die wordt bepaald door een beperkt

Tabel 6: *Partiële regressiecoëfficiënten en multiële correlaties van hulpbronnen en toetsprestaties*

	taal					rekenen				
	totaal	Ned	Sur	Trk	Mar	totaal	Ned	Sur	Trk	Mar
besteding	.04	.06	.10	.02	.01	.04	.03	.09	-.05	.04
taalkeuze	.17	.03	.19	.13	.08	.06	.02	.14	.10	-.01
taalattitude	.08	.03	.08	-.04	-.00	.06	.01	.06	-.01	.03
taalbeheersing moeder	.06	.05	.13	.01	.01	-.02	-.00	.14	-.02	-.02
taalbeheersing vader	.06	.02	-.00	.09	-.02	.02	.01	.04	.06	.03
lezen moeder	.04	.01	.04	.13	-.05	-.00	.03	-.01	.03	-.07
lezen vader	-.00	-.01	.02	-.01	.06	-.02	-.02	-.05	.00	.00
huiswerkhulp moeder	.07	-.05	.09	.05	.10	-.01	-.07	-.06	.04	.02
huiswerkhulp vader	-.08	-.02	.00	-.12	-.04	-.05	-.02	-.08	-.10	-.02
huiswerkhulp siblings	-.07	-.01	-.13	.04	.03	-.05	.02	.06	-.01	-.01
contact school	.05	.04	.06	-.00	.06	.05	.02	.13	.02	.00
contact school zelf	-.08	-.05	-.13	-.01	-.02	-.11	-.10	-.10	-.08	-.12
praten school	.04	.06	-.01	.01	.04	.03	.04	-.03	.01	.08
belang school	-.02	-.02	-.03	.05	-.05	-.02	-.04	-.07	.06	.01
schoolaangepast gedrag	-.09	-.04	-.11	-.02	-.07	-.06	-.01	-.02	-.06	-.05
verklaarde variantie door hulpbronnen (%)	24.0	6.7	14.9	5.6	3.2	6.6	4.2	8.9	3.9	3.0
sociaal milieu	.08	.20	.02	-.05	-.05	.07	.21	-.06	-.04	-.03
toevoeging verklaarde variantie door milieu (%)	.4	2.7	.0	.2	.2	.3	2.9	.3	.1	.1

aantal hulpbronnen met een sterke directe werking op prestaties. Binnen de etnische groepen is de toevoeging van sociaal milieu alleen binnen de Nederlandse gezinnen van enige relevantie. Alleen binnen de Nederlandse gezinnen is er dan ook sprake van een nog net-relevante directe werking van sociaal milieu op zowel de taal-, als rekenprestatie. Tabel 7 geeft een compacte samenvatting van de voorgaande gegevens.

Tabel 7: De mediërende werking van hulpbronnen.

	taal					rekenen				
	totale samenhang		zelfstandige verschil			totale samenhang		zelfstandige verschil		
	R	p	R	p	p	R	p	R	p	p
totaal	.33	.000	.08	.021	.000	.17	.000	.07	.057	.000
Nederlands	.27	.000	.20	.006	.963	.22	.000	.21	.005	.981
Surinaams	.14	.017	.02	.812	.001	.05	.403	-.06	.392	.044
Turks	-.02	.711	-.05	.457	.387	-.04	.470	-.04	.535	.772
Marokkaans	-.04	.523	-.05	.484	.884	-.05	.383	-.03	.657	.935

Het verschil tussen de totale samenhang en de zelfstandige samenhang geeft de sterkte van de mediërende werking van de hulpbronnen weer. We zien dat er sprake is van een redelijke mediërende werking van hulpbronnen op taalprestaties. De mediërende werking op de rekenprestaties is evenwel slechts zwak, hoewel significant door de hoge N. Ook zien we dat binnen de etnische groepen de totale en directe samenhang alleen binnen de Nederlandse gezinnen significant is. De mediërende werking van hulpbronnen is echter bijna steeds afwezig. De totale mediërende werking van hulpbronnen is immers opgebouwd uit alle indirecte werkingen van sociaal milieu op de toetsprestaties via de hulpbronnen. De indirecte werking van een bepaalde hulpbron is gelijk aan het product van de correlatie sociaal milieu*hulpbron (zie Tabel 4) en de directe werking hulpbron*toetsscore (zie Tabel 6). Door de vrijwel totale afwezigheid van directe werkingen van hulpbronnen op toetsprestaties, met name bij rekenen, is er dan ook binnen groepen vrijwel geen sprake van een mediërende werking van hulpbronnen. De significantie p-waarden (onder 'verschil') laten steeds niet-significante verschillen zien tussen de totale en zelfstandige samenhang. Een uitzondering vormen de Surinaamse gezinnen waar de totale samenhang met de taalprestaties van zwak (.14) wordt teruggebracht naar bijna nul (.02). De mediëring loopt daarbij met name via de hulpbronnen 'besteding', 'taalkeuze' en 'taalbeheersing moeder', die samen het grootste deel van de indirecte werking geven, namelijk: .39*.10+.12*.19+.31*.13=.10. Over de etnische groepen heen is er wel sprake van een mediërende werking van hulpbronnen, met name in de voorspelling van de taalprestaties, voor een deel tweegebracht door de hogere correlaties tussen sociaal milieu en hulpbronnen.

4. CONCLUSIES

Op basis van de hier gepresenteerde analyseresultaten kunnen we het volgende concluderen:

- Er zijn geen systematische verschillen tussen de etnische groepen wat betreft de totale en zelfstandige werking van sociaal milieu op toetsprestaties. De samenhangen wijken dus niet systematisch af van de zeer zwakke gemiddelde binnengroepssamenhang.
- De gemiddelde directe werkingen van sociaal milieu op de taal- en rekenprestaties zijn slechts zeer zwak.
- Binnen etnische groepen is er geen mediërende werking van hulpbronnen. Een uitzondering hierop is de Surinaamse groep wat betreft de voorspelling van de taalprestaties, waarbij de mediëring verloopt via het besteedbaar inkomen per persoon, de taalkeuze en de taalbeheersing van de moeder.

- Over groepen heen is er wel sprake van een mediërende werking van hulpbronnen, met name in de voorspelling van de taalprestaties, voor een deel teweegebracht door de hogere correlaties tussen sociaal milieu en hulpbronnen.
- De gegevens suggereren dat het in eerder onderzoek onder etnische groepen veronderstelde 'totale effect' van sociaal milieu op prestaties mogelijk gezien moet worden als een niveau-interpretatiefout. De veronderstelling is namelijk dat dit effect op gezins/leerlingniveau bestaat. In het onderhavige onderzoek is dat echter maar zeer gedeeltelijk het geval: het effect ligt meer op het niveau van de afzonderlijke etnische groepen. In het licht van deze constatering is het de vraag of niet ook de mediëring door culturele hulpbronnen op eenzelfde niveau-interpretatiefout berust.

Het hier beschreven onderzoek is vrij uniek voor de Nederlandse situatie. Immers, (groot-schalig, kwantitatief) onderzoek naar de mediërende effecten van cultureel kapitaal, waarbij expliciet rekening is gehouden met de etnische herkomst van de onderzochten is nog slechts sporadisch uitgevoerd. Onderzoek waarin de bevindingen van afzonderlijke analyses binnen onderscheiden etnische groepen zijn vergeleken is tot op heden nog slechts een enkele keer uitgevoerd. Dit maakt het ook moeilijk de huidige onderzoeksresultaten te vergelijken met eerdere bevindingen.

Laten we eerst beginnen met het noemen van enkele beperkingen van ons onderzoek. Een nadeel was dat het ging om secundaire analyses op bestaande databestanden met variabelen die niet alle specifiek voor dit onderwerp zijn verzameld. Van de andere kant moet dit ook niet overtrokken worden, immers, de meeste van de in de inleiding aangehaalde studies hebben met dat euvel te maken. Een voordeel is in ieder geval dat we de beschikking hadden over een grote dataset met recente gegevens. Als bezwaar zou verder kunnen worden ingebracht dat de geanalyseerde informatie louter is verkregen via schriftelijke vragenlijsten. In eventueel vervolgonderzoek zou - hoewel dat, zeker als het allochtonen betreft, geen sinecure en kostbaar is - aanvullend daarop ook informatie kunnen worden verzameld via bijvoorbeeld het gedurende langere tijd verrichten van observaties bij en het voeren van meerdere mondelinge vraaggesprekken met beide ouders - zowel de moeder als de vader. Daarbij zou meer aandacht moeten worden besteed aan concrete activiteiten die de ouders met hun kinderen ondernemen (vgl. bv. Leseman & Van den Boom, 1999).

De resultaten van de literatuurstudie lieten zien dat de onderzoeksbevindingen - in binnen- en buitenland - niet eenduidig waren: soms werd de kapitaal-these bevestigd, soms ook niet. Wat het huidige empirische onderzoek vooral heeft laten zien is dat - wanneer het gaat om de etnische factor - het relevant kan zijn de totale steekproef op te splitsen in substeekproeven van afzonderlijke etnische groepen. Bevindingen die mogelijk gelden voor de totale steekproef hoeven namelijk niet per se op te gaan voor elk van de etnische groepen afzonderlijk. Dit kan bijvoorbeeld veroorzaakt worden doordat zowel de verdelingen alsook de samenhangen van de te verklaren en de verklarende variabelen sterk variëren binnen de onderscheiden groepen (vgl. Leseman & Van den Boom, 1999; Okagaki & Frensch, 1998). De implicatie hiervan is uiteraard niet alleen theoretisch van aard. Ook voor de praktijk betekent deze bevinding namelijk dat aanpakken die geldig en relevant zijn voor de ene etnische groep dat niet per se hoeven te zijn voor andere groepen. In het uiterste geval kan het zelfs zo zijn dat sommige aanpakken in bepaalde groepen zelfs juist niet-gewenste effecten oproepen. Deze opmerking ligt in de lijn van die van Okagaki & Frensch (1998). Zoals in de inleiding van dit artikel al opgemerkt, vonden zij verschillen tussen etnische groepen in ouderlijke opvattingen en gedrag wat betreft opvoeding en onderwijs. Evenals in ons onderzoek waren deze voor de Europees-Amerikaanse groep het meest consistent gerelateerd aan de schoolprestaties. Volgens Okagaki & Frensch is dat niet zo verwonderlijk, aangezien de betreffende indicatoren waren gebaseerd op eerder onderzoek dat hoofdzakelijk onder Europees-Amerikaanse gezinnen is uitgevoerd. Zij opperen dat deze indicatoren daarom wellicht minder relevant zijn voor niet-Westerse groepen, of, in de Nederlandse situatie, voor etnische minderheidsgroepen. Het

geheel aan opvattingen en gedrag dat een faciliterende werking kan hebben met betrekking tot de schoolprestaties in een bepaalde etnische groep wordt waarschijnlijk beïnvloed door de sociale context van die groep. Dit houdt in - zoals we hierboven al opmerkten - dat we er niet van uit kunnen gaan dat wat werkt voor de ene groep noodzakelijkerwijze ook werkt voor andere groepen. Deze conclusie sluit ook aan bij de opmerking van Lareau & McNamara Horvat (1999, p. 50) dat 'Any form or type of capital derives value only in relation to the specific field of interaction.'. Daar kan de constatering van Roscigno & Ainsworth-Darnell (1999, p. 173) aan worden toegevoegd 'that differential racial returns are a consequence of the inapplicability of the status attainment model to nonwhites.'. Afsluitend willen we daarom opmerken dat de problematiek van contaminatie van specifieke, sterk cultureel bepaalde kenmerken met etnische herkomst - of in de terminologie van Allen & Yen (1979): effecten van 'group heterogeneity' - nader onderzoek vereist. Centraal zou daarbij de vraag moeten staan wat nu precies de betekenis is van cultureel kapitaal binnen elk van de etnische groepen en in hoeverre dit concept over groepen heen vergelijkbaar is - als het dat al is. Want dat is eigenlijk het fundamentele probleem dat deze studie oproept.

NOTEN

* Het onderzoek waarover hier wordt gerapporteerd, is in financiële zin gesteund door de Stichting voor de Gedragswetenschappen die wordt gesubsidieerd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Het betreft NWO-project # 575-53-013, toegekend aan dr. Geert W.J.M. Driessen van het ITS te Nijmegen. Dank gaat uit naar dr. Adrie Claassen en drs. Jan Doesborgh voor hun bijdrage aan de literatuurstudie, respectievelijk het empirische deel van het onderzoek. Dit artikel is gebaseerd op Driessen, Doesborgh & Claassen (1999).

1. We kregen pas de beschikking over deze onderzoeksresultaten na voltooiing van ons onderzoek.

2. Onlangs moesten ook Roscigno & Ainsworth-Darnell (1999) concluderen dat 'surprisingly little research has focused on if and how racial and ethnic groups differ with regard to these attributes in general, and cultural capital, in particular' (o.c.: p. 160).

3. Een nadeel van een zo hoge totale N is dat ook zeer kleine effecten nog significant zullen lijken op het niveau dat in statistische procedures minimaal wordt aangegeven, namelijk $p=.001$. Bij een totale $N=300$ wordt een maximale $p=.01$ aangehouden als grenswaarde van wat een net-significant effect aanduidt, en bij een totale $N=500$ een maximale $p=.001$. Een net-significante correlatie is dan in beide gevallen ongeveer gelijk aan $r=.15$. Bij $N=1128$ is een correlatie met een $p=.001$ niet groter dan $r=.10$, zodat significante niet-relevante effecten toch significant kunnen lijken. Daarom moet ook, tegelijk met de p -waarde voor significantie, gelet worden op de sterkte van de samenhang. Doorgaans wordt een (gestandaardiseerde) samenhang (hier meestal Pearson- r of variantie-analyse η^2) van $.15$ als minimum aangehouden; 2,25% van de variatie in de afhankelijke variabele wordt dan 'verklaard' vanuit de predictor. Met name bij interactie-effecten, waarbij meestal meer vrijheidsgraden betrokken zijn, zal een minimale $\eta^2=.20$ moeten worden aangehouden. Als we significante p -waarden laten zien voor samenhangen binnen elk van de groepen, dan is een effect met $p=.020$ net-significant, dat wil zeggen: $p=.020$ is het kritieke significantieniveau dat hoort bij $N=282$.

4. Het verschil tussen de η^2 voor de totaal-samenhang in Tabel 2 en de correlatie uit Tabel 5 geeft de mate van niet-lineariteit van de samenhang, waaruit we nogmaals kunnen zien dat de samenhangen sterk lineair zijn. In een enkel geval is η^2 iets kleiner dan Pearson- r (ofwel de nominaal-metrische, respectievelijk metrische samenhang), doordat in Tabel 2 ingedikte versies van de predictoren zijn gebruikt, terwijl in de correlaties in Tabel 5 de volledige variatie wordt gebruikt.

LITERATUUR

- Ainsworth-Darnell, J., & Downey, D. (1998). Assessing the oppositional culture explanation for racial/ethnic differences in school performance. *American Sociological Review*, 63, (4), 536-553.
- Allen, M., & Yen, W. (1979). *Introduction to measurement theory*. Monterey, LA: Brooks/Cole.
- Aschaffenburg, K., & Maas, I. (1997). Cultural and educational careers: The dynamics of social reproduction. *American Sociological Review*, 62, (4), 573-587.
- Bankston, C., Caldas, S., & Zhou, M. (1997). The academic achievement of Vietnamese American students: Ethnicity as social capital. *Sociological Focus*, 30, (1), 1-16.
- Bergen, J. van (1987). *Verantwoording constructie toetsen voor de evaluatie van het Onderwijsvoor-rangs-beleid*. Arnhem: Cito.
- Bianchi, S., & Robinson, J. (1997). What did you do today? Children's use of time, family composition, and the acquisition of social capital. *Journal of Marriage and the Family*, 59, (2), 332-344.
- Carrington, V., & Luke, A. (1997). Literacy and Bourdieu's sociological theory: A reframing. *Language and Education*, 11, (2), 96-112.
- Cohen, J., & Cohen, P. (1983). *Applied Multiple regression/correlation analysis for the social sciences*. Hillsdale, NJ/London: Lawrence Erlbaum Associates.
- Coleman, J. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 95-120.
- Coleman, J., & Hoffer, T. (1987). *Public and private high schools. The impact of communities*. New York: Basic Books.
- Covrig, D. (1996). Fundamentalists, social capital, and children's welfare: A place for religion in public education? *Journal of Educational Policy*, 12, (1/2), 53-60.
- Crul, M. (2000). *De sleutel tot succes. Over hulp, keuzes en kansen van Turkse en Marokkaanse jongeren van de tweede generatie*. Amsterdam: Het Spinhuis.
- DiMaggio, P. (1982). Cultural capital and school success: the impact of status culture participation on the grades of US high school students. *American Sociological Review*, 47, 189-210.
- Downey, D. (1995). When bigger is not better: Family size, parental resources, and children's educational performance. *American Sociological Review*, 60, (5), 746-761.
- Driessen, G. (1993). Achtergronden van onderwijsprestaties. De rol van gezinsstructurele en gezinspedagogische kenmerken. *Stimulans*, 11, (2), 14-17.
- Driessen, G. (1995). Het relatieve belang van sociaal milieu en etnische herkomst voor de verklaring van onderwijsachterstanden. Een overzicht van kwantitatief-empirische studies; bevindingen en problemen. *Tijdschrift voor Onderwijsresearch*, 20, (4), 341-362.
- Driessen, G., Doesborgh, J., & Claassen, A. (1999). *Cultureel kapitaal, etnische herkomst en onderwijsprestaties*. Nijmegen: ITS.
- Driessen, G., Langen, A. van, & Oudenhoven, D. (1994). *De toetsen voor het cohort Primaire Onderwijs. Verantwoording*. Nijmegen: ITS.
- Driessen, G., & Haanstra, F. (1996). *De oudervragenlijst basisonderwijs en speciaal onderwijs. Technische rapportage PRIMA-cohortonderzoek 1994/95*. Amsterdam/Nijmegen: SCO/ITS.
- Driessen, G., & Tesser, P. (1998). Alloctonen in het onderwijs. In R. Penninx, H. Münstermann & H. Entzinger (Eds.), *Etnische minderheden en de multiculturele samenleving* (pp. 330-391). Groningen: Wolters-Noordhoff.
- Farkas, G., Grobe, R., Sheenan, D., & Shuan, Y. (1990). Cultural resources and school success: Gender, ethnicity, and poverty groups within an urban school district. *American Sociological Review*, 55, (1), 127-142.
- Furstenberg, F., & Hughes, M. (1995). Social capital and successful development among at-risk youth. *Journal of Marriage and the Family*, 57, (3), 580-592.
- Gijsberts, M. (1993). *Culturele veranderingen: het wegen waard?* Rijswijk: SCP.
- Graaf, P. de (1987). *De invloed van financiële en culturele hulpbronnen in onderwijsloopbanen*. Nijmegen: ITS.
- Jungbluth, P. (1985). *Verborgene differentiatie. Leerlingbeeld en onderwijsaanbod op de basisschool*. Nijmegen: ITS.
- Jungbluth, P., Langen, A. van, Peetsma, T., & Vierke (1996). *Leerlinggegevens basisonderwijs en speciaal onderwijs. Technische rapportage PRIMA-cohortonderzoek 1994/95*. Amsterdam/Nijmegen: SCO-Kohnstamm Instituut/ITS.
- Kalmijn, M., & Kraaykamp, G. (1996). Race, cultural capital, and schooling: An analysis of trends in the United States. *Sociology of Education*, 69, (1), 22-34.
- Lareau, A., & McNamara Horvat, E. (1999). Moments of social inclusion and exclusion. Race, class, and cul-

- tural capital in family-school relationships. *Sociology of Education*, 72, (3), 37-53.
- Langen, A. van, Vierke, H., & Robijns, M. (1996). *Veldwerkverslag basisonderwijs en speciaal onderwijs. Technische rapportage PRIMA-cohortonderzoek 1994/95*. Amsterdam/Nijmegen: SCO-Kohnstamm Instituut/ITS.
- Ledoux, G., Deckers, P., Bruijn, E. de, & Voncken, E. (1992). *Met het oog op de toekomst. Ideeën over onderwijs en arbeid van ouders en kinderen uit de doelgroepen van het Onderwijsvoorrangsbeleid*. Amsterdam: SCO-Kohnstamm Instituut.
- Leseman, P. (1989). *Structurele en pedagogische determinanten van schoolloopbanen. Verslag van een longitudinaal onderzoek naar de invloed van het gezin op de schoolprestaties, uitgevoerd als onderdeel van de summatieve evaluatie van het project Onderwijs en Sociaal Milieu, aangevuld met een theoretische beschouwing over de relatie sociaal milieu, socialisatie, intelligentie en schoolprestaties*. Rotterdam: SAD.
- Leseman, P., & Boom, D. van den (1999). Effects of quantity and quality of home proximal processes on Dutch, Surinamese-Dutch and Turkish-Dutch pre-schoolers' cognitive development. *Infant and Child Development*, 8, (1), 19-38.
- Leseman, P., Sijtsling, F., Jap-A-Joe, S., & Sahin, S. (1995). Gezinsdeterminanten van de cognitieve ontwikkeling van vierjarige Nederlandse, Surinaamse en Turkse kleuters. *Pedagogische Studiën*, 72, (3), 186-205.
- Meesters, M. (1992). *Loopbanen in het onderwijs en op de arbeidsmarkt. Verticale en horizontale differentiatie in het voortgezet onderwijs: oorzaken en gevolgen voor de arbeidsmarktpositie van Nederlandse jongeren*. Nijmegen: ITS.
- Meijnen, G. (1990). Hulpbronnen voor schoolcarrières. Een poging tot synthese. *Sociologische Gids*, 37, (5), 288-303.
- Mulder, L. (1996). *Meer voorrang, minder achterstand? Het Onderwijsvoorrangsbeleid getoetst*. Nijmegen: ITS.
- Nijsten, C. (1998). *Opvoeding in Turkse gezinnen in Nederland*. Assen: Van Gorcum.
- Okagaki, L., & Frensch, P. (1998). Parenting and children's school achievement: A multiethnic perspective. *American Educational Research Journal*, 35, (1), 123-144.
- Pels, D. (1992). Pierre Bourdieu, opstellen over smaak, habitus en het veldbegrip. Amsterdam: Van Gennep.
- Pels, T. (1998). *Opvoeding in Marokkaanse gezinnen in Nederland. De creatie van een nieuw bestaan*. Assen: Van Gorcum.
- Roelandt, T. (1994). *Verscheidenheid in ongelijkheid. Een studie naar etnische stratificatie en onderklassenvorming in de Nederlandse samenleving*. Amsterdam: Thesis Publishers.
- Roelandt, T., Martens, R., & Veenman, J. (1990). Achterstand van allochtonen in het onderwijs: sociaal milieu en migratie-achtergronden. *Mens en Maatschappij*, 65, (2), 103-125.
- Roscigno, V., & Ainsworth-Darnell, J. (1999). Race, cultural capital, and educational resources: Persistent inequalities and achievement returns. *Sociology of Education*, 72, (3), 158-178.
- Runyan, D., Hunter, W., Socolar, R., Amaya-Jackson, L., English, D., Landsverk, J., Dubowitz, H., Browne, D., Bangdiwala, S., & Mathew, R. (1998). Children who prosper in unfavorable environments: The relationship to social capital. *Pediatrics*, 101, (1), 12-18.
- Stanton-Salazar, R., & Dornbusch, S. (1995). Social capital and the reproduction of inequality: information networks among Mexican-origin High School students. *Sociology of Education*, 68, (2), 116-135.
- Sullivan, L. (1997). Hip-hop nation: The underdeveloped social capital of black urban America. *National Civic Review*, 86, (3), 235-243.
- Teachman, J., Paasch, K., & Carver, K. (1996a). Social capital and dropping out of school early. *Journal of Marriage and the Family*, 58, (3), 773-783.
- Teachman, J., Paasch, K., & Carver, K. (1996b). Social capital and the generation of human capital. *Social Forces*, 75, (4), 1343-1359.
- Tesser, P., & Mulder, L. (1990). Cultureel kapitaal en schoolprestaties in het basisonderwijs. In: C. Klaassen & P. Jungbluth (Eds.), *Onderwijs Research Dagen 1990. Onderwijs en samenleving* (pp. 39-51). Nijmegen: ITS.
- Velden, R. van der (1991). *Sociale herkomst en schoolsucces. Het effect van culturele en sociale hulpbronnen op de schoolloopbaan*. Groningen: RION.
- White, M., & Kaufman, G. (1997). Language usage, social capital, and school completion among immigrants and native-born ethnic groups. *Social Science Quarterly*, 78, (2), 385-398.

BIJLAGEN

	gem	sd	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 besteding	663	395	1.00													
2 taalkeuze	2.5	1.4	.16	1.00												
3 taalattitude	.6	.9	.30	.52	1.00											
4 taalbeheersing moeder	2.5	.7	.34	.42	.47	1.00										
5 taalbeheersing vader	2.6	.6	.31	.37	.41	.48	1.00									
6 lezen moeder	5.3	5.6	.22	.23	.30	.37	.27	1.00								
7 lezen vader	5.6	5.8	.19	.13	.24	.16	.26	.59	1.00							
8 huiswerkhulp moeder	2.2	.7	.17	.20	.28	.45	.24	.28	.13	1.00						
9 huiswerkhulp vader	1.9	.7	.18	.06	.12	.15	.31	.12	.18	.47	1.00					
10 huiswerkhulp siblings	1.9	.8	-.27	-.18	-.30	-.32	-.34	-.22	-.13	-.05	.05	1.00				
11 contact school	2.5	.6	.16	.20	.18	.26	.25	.17	.13	.16	.12	-.06	1.00			
12 contact school zelf	2.1	.7	.09	.06	.07	.12	.11	.17	.13	.19	.17	.01	.22	1.00		
13 praten school	3.5	.9	.13	.19	.24	.32	.23	.22	.15	.24	.07	-.14	.18	.07	1.00	
14 belang school	4.3	1.0	-.10	-.11	-.12	-.13	-.12	-.07	.00	-.10	-.05	.17	-.03	.02	-.08	1.00
15 schoolaangepast gedrag	3.2	.5	-.16	-.18	-.21	-.17	-.16	-.14	-.06	-.06	-.01	.18	.00	.10	-.05	.31

Bijlage 1: Gemiddelden, standaarddeviaties en correlaties tussen de hulpbronnen.

	correlaties			alfa's	gemiddelden (standaarddeviaties)		
	oplm*oplv	oplm*berv	oplv*berv		oplm	oplv	berv
totaal	.56	.55	.72	.82	2.3 (1.5)	2.5 (1.8)	2.6 (1.7)
Nederlands	.53	.48	.72	.80	3.4 (1.6)	3.7 (1.8)	3.8 (1.8)
Surinaams	.37	.34	.72	.73	2.6 (1.3)	2.8 (1.6)	3.1 (1.7)
Turks	.30	.26	.50	.62	1.6 (1.2)	2.2 (1.5)	2.0 (1.3)
Marokkaans	.33	.32	.53	.66	1.4 (1.1)	1.5 (1.1)	1.6 (1.2)

Bijlage 2: Correlaties, betrouwbaarheid, gemiddelden en standaarddeviaties milieu-indicatoren (opleiding moeder, opleiding vader, beroep vader).

Effecten van curriculumkenmerken op studievoortgang en uitval van studenten techniek

M. van der Hulst

Katholieke Universiteit Nijmegen, Sectie Arbeids- en Organisationspsychologie, Postbus 9104, 6500 HE Nijmegen

E. Jansen

Rijksuniversiteit Groningen, GION, Afdeling COWOG, Postbus 335, 9700 AH Groningen

ABSTRACT

Procrastination and time investment are important issues in the study of students' performance and study progress. A large proportion of research on these issues concentrates on individual differences between students in personality and time management skills. In this article, it is hypothesised that efficient curriculum organisation can reduce procrastination and increase study progress. A multilevel analysis that was carried out on data concerning about 1600 engineering students who studied in twelve different study programmes provides evidence that curriculum characteristics do indeed affect study progress. The spread of study activities over the year, instruction characteristics and examination characteristics were found to have effects on study progress, but not on drop-out. Thus, institutes in higher education can affect their students' study progress by means of efficient curriculum organisation.

1. INLEIDING

De studievoortgang van studenten in het hoger onderwijs is in belangrijke mate afhankelijk van de inzet van studenten in termen van tijdsinvestering. Aangezien het aantal uren contactonderwijs vaak laag is en het bezoeken van de colleges vrijblijvend, hangt er veel af van de discipline van studenten om te komen tot regelmatige (zelf)studie (zie bijvoorbeeld Mee-
rum Terwogt-Kouwenhoven, 1990). Uitstelgedrag speelt daarbij een belangrijke rol. Veel onderzoek naar studievoortgang is daarom gericht op individuele verschillen tussen studenten in uitstelgedrag en tijdsinvestering in de studie.

In het onderzoek naar uitstelgedrag kunnen we drie perspectieven onderscheiden, namelijk het persoonlijkheidsperspectief, het vaardigheidsperspectief en het motivatieperspectief. Onderzoek naar de relatie tussen persoonlijkheidskenmerken en uitstelgedrag heeft laten zien dat sommige studenten inderdaad meer neigen tot uitstelgedrag dan anderen. Uitstelgedrag hangt samen met faalangst en gebrek aan zelfvertrouwen en meer in het algemeen met een gebrek aan zelfdiscipline (Schouwenburg, 1992, 1994; Johnson & Bloom, 1995). Het persoonlijkheidsperspectief leidt tot een tamelijk pessimistische visie op de mogelijkheden studievoortgang te verbeteren. Sommige studenten zouden een ongeschikte persoonlijkheid hebben om hun studie vlot te doorlopen en persoonlijkheden laten zich nu eenmaal moeilijk veranderen. Een veel optimistischer visie is het verklaren van studievoortgang vanuit de vaardigheden van de student. Vanuit het vaardigheidsperspectief wordt vooral onderzoek gedaan naar planningsvaardigheden in verband met prestaties en stress. In sommige onderzoeken is een verband gevonden tussen planningsvaardigheden en behaalde cijfers (Macan e.a., 1990),

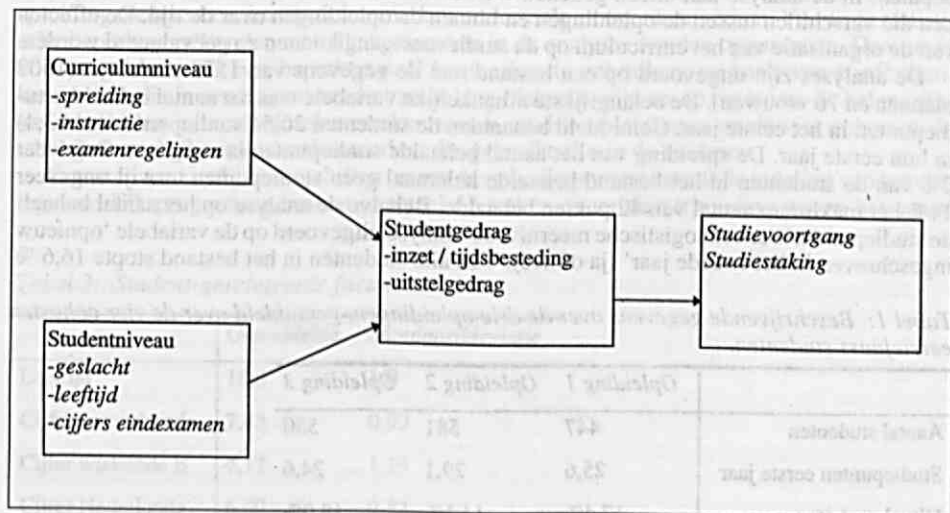
maar in andere niet (Nonis e.a. 1998). Meerum Terwogt-Kouwenhoven (1990) vond een verband tussen planningsvaardigheden en studievertraging. Matig vertraagde studenten blijken problemen te hebben met het omgaan met de tijd, ernstig vertraagden hebben vooral cognitieve problemen. Een implicatie van het vaardigheidsperspectief is dat het aanleren van planningsvaardigheden de studievoortgang kan bevorderen, maar de effectiviteit van dit soort vaardigheidstrainingen staat geenszins vast (Macan, 1994). Het derde onderzoeksperspectief is gericht op de relatie tussen motivatie en uitstelgedrag. Het uitstellen van activiteiten voor de studie hangt samen met een gebrek aan intrinsieke motivatie (b.v. Senécal e.a., 1995; Milgram e.a., 1995). Uitstelgedrag hangt dus niet alleen samen met algemene persoonlijkheidskenmerken en vaardigheden, maar ook met de specifieke taak of activiteit die uitgevoerd moet worden. Studenten zullen met name de taken uitstellen die ze niet interessant of niet plezierig vinden.

De tijdsinvestering in de studie hangt niet alleen af van de persoonlijkheid, vaardigheden en motivatie van de student, maar ook van de studeerbaarheid van het programma. In de afgelopen 15 jaar is er in Nederland een aantal onderzoeken gedaan naar de effecten van de organisatie van het studieprogramma op de studievoortgang van studenten. Het centrale uitgangspunt bij dit soort onderzoek is dat een efficiënte organisatie van het studieprogramma uitstelgedrag zou kunnen verminderen. Tenminste drie soorten programmakenmerken lijken daarbij van belang te zijn, te weten spreiding van onderwijs en tentamens, de hoeveelheid en de soort instructietijd, en formele examen- en compensatieregelingen. De spreiding van onderwijs en tentamens is van belang omdat studenten de voorbereiding van tentamens zo lang mogelijk uitstellen, zoals bijvoorbeeld bleek uit het onderzoek van Van der Drift en Vos (1987). Om studenten regelmatig te laten werken moet er volgens hen een constante druk zijn, maar concurrentie tussen vakken en piekbelastingen moeten daarbij voorkomen worden. Ook in andere artikelen (b.v. Vos, 1992; Bax, 1992; Jansen, 1996, 1997) wordt erop gewezen dat spreiding van tentamens over het jaar, herkansingen buiten de reguliere tentamenperiodes en een beperking van het aantal vakken dat parallel geroosterd is de studievoortgang bevordert. Wanneer we het studieprogramma nog wat preciezer bekijken blijkt de hoeveelheid en de soort instructietijd een belangrijke determinant te zijn van de tijdsinvestering in de studie. Uit het onderzoek van Van der Drift en Vos (1987) is gebleken dat er een kromlijng verband is tussen het aantal contacturen en de tijd die studenten in hun studie investeren. Naarmate het aantal contacturen toeneemt, neemt het aantal uren besteed aan zelfstudie in eerste instantie toe, maar na verloop van tijd weer af (zie ook Gijsselaers & Schmidt, 1993). De totaal geïnvesteerde tijd is gebonden aan een maximum en kan dan niet meer verhoogd worden door het aantal contacturen te verhogen. Ook het type contacturen lijkt van belang te zijn, waarbij een onderscheid gemaakt kan worden tussen passieve contacturen in de vorm van hoorcolleges en actieve contacturen in de vorm van werkcolleges, practica en projecten. Bij de actieve contacturen is er sprake van terugkoppeling over het begrip en de beheersing van de stof. Onderzoek aan de TU Delft heeft laten zien dat regelmatige tussentijdse terugkoppeling door middel van computertoetsen een positief effect kan hebben op de studievoortgang (Dijkman, 1991; Wilders e.a., 1996). Tenslotte zijn de formele examenkenmerken, met name compensatieregelingen, van belang. Hoe meer tentamens men moet halen om te slagen voor de propedeuse, hoe meer zakkansen, maar dit effect kan weer teniet gedaan worden door uitgebreide compensatiemogelijkheden. De invoering van compensatieregelingen zou de studieduur kunnen bekorten (zie ook Cohen-Schotanus, 1995). Er zijn dus veel aanwijzingen dat kenmerken van het studieprogramma de studievoortgang kunnen beïnvloeden.

Tot nu toe hebben we steeds gesproken over studievoortgang, maar een ander aspect van studiesucces waar veel onderzoek naar is gedaan is studiestaking. In waarschijnlijk het eerste grootschalige Nederlandse onderzoek naar studiesucces ('Mislukking en vertraging van de studie', 1959) werd uiteengezet dat de factoren die studievertraging veroorzaken niet hetzelfde hoeven te zijn als de factoren die tot studiestaking leiden. Dit onderzoek, dat evenals het hieronder te bespreken onderzoek werd uitgevoerd bij de TU Delft (toen nog de Techni-

sche Hogeschool Delft), concentreerde zich vooral op individuele factoren die studievertraging en studiestaking kunnen verklaren. In het rapport werd daarnaast de vraag gesteld of het onderwijs zelf ook niet een oorzaak van studievertraging en studiestaking zou kunnen zijn, aangezien een aanzienlijk deel van de studenten ontevreden was over de didactiek. Bovendien werd er door de onderzoekers ook een grote organisatorische ingewikkeldheid, ondoorzichtigheid en onpersoonlijkheid van het studieprogramma geconstateerd. In het huidige onderzoek zijn we vooral geïnteresseerd in de effecten van curriculum-gerelateerde factoren op studievoortgang en studiestaking. Een verbetering van de studeerbaarheid zou kunnen samengaan met een betere studievoortgang maar ook met een verminderde selectiviteit van de propedeuse, met het risico dat studenten pas in een latere fase van de studie uitvallen. Een uitgebreid onderzoek naar de effecten van student-gerelateerde en curriculum-gerelateerde factoren op studiestaking is gedaan door Prins (1997, 1998), die concludeerde dat studiebegeleiding, meer contacttijden, minder parallelle vakken, studentgerichtheid van het onderwijs en een uitstel voorkomend tentamenbeleid tot minder uitval of een vervroeging van het uitvalmoment leiden. Dit suggereert dat de factoren die leiden tot een snellere studievoortgang juist een positieve uitwerking hebben op studiestaking.

Uit verschillende onderzoeken (bijvoorbeeld de Boom e.a., 1998) blijkt dat studenten technisch relatief veel tijd investeren in hun studie maar niettemin weinig punten halen in vergelijking met andere studenten. De gemiddelde studieduur van studenten techniek is dan ook tamelijk lang, en techniek is de enige sector waarin de studieduur de afgelopen jaren niet is afgenomen (KUO, 1998). We zullen hier een onderzoek bespreken dat is uitgevoerd aan de TU Delft. Deze universiteit heeft de afgelopen jaren aan aantal maatregelen genomen om de studeerbaarheid te bevorderen. De belangrijkste veranderingen zijn de invoering van een vijfjarig programma, het toevoegen van vaardigheids cursussen, en de invoering van projectonderwijs en werkgroepen om de integratie van de stof die in verschillende theoretische vakken is aangeboden te bevorderen. Bij verschillende opleidingen zijn deze veranderingen in het curriculum op verschillende manieren uitgevoerd. Het onderzoek richtte zich op de effecten van curriculumkenmerken op studievoortgang en studiestaking. Door de samenhang tussen veranderingen in het curriculum en studievoortgang te bestuderen kunnen de effecten van ingevoerde maatregelen worden geëvalueerd. Studiestaking werd in het onderzoek betrokken omdat het een beeld geeft van eventuele veranderingen in de selectiviteit van de opleidingen.



Figuur 1: Analysemodel.

Het analysemodel in Figuur 1 geeft een overzicht van de onderzoeksvariabelen. De invloed van curriculumkenmerken en studentkenmerken wordt verondersteld te lopen via studentgedrag (tijdsbesteding aan de studie en uitstelgedrag). In het onderzoek wordt gekeken naar de verbanden tussen de cursief gedrukte variabelen.

2. METHODE

Het onderzoek werd uitgevoerd bij vier cohorten studenten (beginjaar 1994-1995 tot en met beginjaar 1997-1998) bij drie opleidingen. Vanwege de geneste datastructuur is ervoor gekozen een meerniveau-analyse (zie b.v. Bryk & Raudenbush, 1992; Sijnders en Bosker, 1994) uit te voeren. Er waren drie niveaus, namelijk studenten op het laagste niveau, cohorten op het middelste niveau en opleidingen op het hoogste niveau. Er is echter voor gekozen om een model met slechts twee niveaus te formuleren omdat de steekproefgroottes van de cohorten en de opleidingen erg klein waren en er geen verklarende variabelen waren op het opleidingsniveau. Aangezien het studieprogramma van elke opleiding van jaar tot jaar (dus per cohort) verschilde zijn er in totaal twaalf curricula in de analyses opgenomen. De meerniveau-analyses die hier gerapporteerd worden hebben dus betrekking op studenten op het laagste niveau en curriculum op het tweede niveau.

Om een homogene onderzoekspopulatie te krijgen werden alleen gegevens van studenten die direct aansluitend op het behalen van het VWO-diploma aan de opleiding begonnen in het onderzoek betrokken. Tabel 1 geeft een overzicht van aantallen studenten en de spreiding van de afhankelijke variabelen over de opleidingen. De analyses die hier besproken zullen worden hebben betrekking op de steekproef als geheel en niet op de afzonderlijke opleidingen.

In de meerniveau-analyse is studievoortgang in verband gebracht met studentkenmerken en curriculumkenmerken. Meerniveau-analyse is een regressietechniek die het mogelijk maakt te bepalen welke factoren in het onderwijsprogramma de studievoortgang beïnvloeden terwijl er gecorrigeerd wordt voor verschillen in studentkenmerken. De procedure is gebaseerd op een stapsgewijze opname van verschillende factoren in een model dat gebaseerd is op lineaire regressie. Bij een meerniveau-analyse is het mogelijk een schatting te maken van de mate waarin factoren op individueel niveau en factoren op curriculumniveau de studievoortgang bepalen. In de analyse kan alleen gekeken worden naar de effecten van curriculumkenmerken die verschillen tussen de opleidingen en binnen de opleidingen over de tijd. De effecten van de organisatie van het curriculum op de studievoortgang kunnen zo geëvalueerd worden.

De analyses zijn uitgevoerd op een bestand met de gegevens van 1578 studenten (1502 mannen en 76 vrouwen). De belangrijkste afhankelijke variabele was het aantal behaalde studiepunten in het eerste jaar. Gemiddeld behaalden de studenten 26,54 studiepunten (sd 13,4) in hun eerste jaar. De spreiding van het aantal behaalde studiepunten is erg groot. Ongeveer 2% van de studenten in het bestand behaalde helemaal geen studiepunten terwijl ongeveer 14% het maximum aantal van 42 punten behaalde. Behalve de analyse op het aantal behaalde studiepunten is er een logistische meerniveau-analyse uitgevoerd op de variabele 'opnieuw ingeschreven in het tweede jaar' (ja of nee). Van alle studenten in het bestand stopte 16,6%

Tabel 1: Beschrijvende gegevens over de drie opleidingen, gemiddeld over de vier cohorten eerstejaars studenten.

	Opleiding 1	Opleiding 2	Opleiding 3
Aantal studenten	447	581	550
Studiepunten eerste jaar	25,6	29,1	24,6
Uitval na 1 jaar	17,4%	14,1%	18,0%

na (of binnen) 1 jaar met de studie. De resultaten van deze analyse zullen hier kort aangehaald worden.

De onafhankelijke variabelen waren student-gerelateerde factoren enerzijds en curriculum-gerelateerde factoren anderzijds. Deze verklarende variabelen zijn stapsgewijs in de analyse opgenomen. Allereerst is het zogenaamde 'lege model' geanalyseerd, dat wil zeggen een model zonder verklarende variabelen. Vervolgende zijn alle student-gerelateerde variabelen toegevoegd. Aangezien de steekproefgrootte op het curriculumniveau vrij klein was, was het statistisch niet verantwoord om alle predictoren op curriculumniveau tegelijk in een model op te nemen (dit zou leiden tot een lage power en toevalseffecten). Daarom zijn de verklarende variabelen op het curriculumniveau in drie aparte modellen opgenomen. In het eerste model werden de effecten van spreiding van studielast bestudeerd, in het tweede model ging het om instructievariabelen, en in het derde model ging het om examenkenmerken. De curriculumvariabelen die in hetzelfde model werden opgenomen waren onderling niet gecorreleerd. De gegevens zijn verwerkt met MLwiN, een programma voor meerniveau-analyse. Van afzonderlijke predictorvariabelen is met een Wald-test (t-toets met aangepaste vrijheidsgraden, tweezijdig, $(\alpha = 5\%)$) getoetst of de coëfficiënt significant was. Bovendien is getoetst of het toevoegen van een groep variabelen een significante verbetering van het model opleverde met een deviance-test (Chi-kwadraat, $(\alpha = 5\%)$).

3. STUDENT-GERELATEERDE FACTOREN

Voor alle studenten waren vijf gegevens bekend, namelijk het geslacht, de leeftijd bij de aanvang van de studie en de cijfers op het VWO-examen voor natuurkunde, wiskunde B en Nederlands. Al deze gegevens zijn in de regressie-analyses als predictoren opgenomen. Tabel 2 laat de gemiddelden en de spreiding zien voor de student-gerelateerde factoren.

4. CURRICULUM-GERELATEERDE FACTOREN

Van elk van de twaalf curricula is aan de hand van studeerhandleidingen een groot aantal kenmerken in kaart gebracht. Deze kenmerken zijn onderverdeeld in drie categorieën, namelijk spreidingskenmerken, instructiekenmerken en formele examenkenmerken. Vanwege statistische beperkingen konden niet alle kenmerken in de analyses worden opgenomen. Daarom is er voor gekozen om een beperkt aantal kenmerken te gebruiken, op basis van curriculumkenmerken die in eerder onderzoek van belang bleken te zijn en op basis van de belangrijkste verschillen in curriculumkenmerken tussen de drie opleidingen en de vier cohorten. In tabel 3 staan de 8 curriculumkenmerken die gebruikt zijn in de analyses.

Er waren aanzienlijke verschillen tussen de opleidingen wat betreft de indeling van het collegejaar. Elk programma beslaat in totaal 42 weken, maar de verhouding tussen het aantal

Tabel 2: Student-gerelateerde factoren.

	Gemiddelde	Standaarddeviatie
Leeftijd	18,3	1,0
Cijfer natuurkunde	7,43	0,99
Cijfer wiskunde B	7,12	1,19
Cijfer Nederlands	6,60	0,82

Tabel 3: Curriculum-gerelateerde factoren.

		Minimum	Maximum
Spreiding	Aantal onderwijsweken	25	33
	Aantal onderwijsvrije weken	0	4
	Gemiddeld aantal vakken parallel	3,75	5
Instructie	Aantal uren hoorcollege	246	469
	Aantal uren werkcollege	12	270
	Aantal theoretische vakken	8	11
Examen	Aantal tentamens	14	19
	Aantal eindcijfers	8	16

onderwijsweken, het aantal tentamenweken en het aantal onderwijsvrije weken (of 'witte weken', voorafgaand aan de tentamenperiode) verschilt per opleiding. Aan de hand van de studeerhandleidingen is het ingeroosterde aantal uren hoorcollege en werkcollege berekend. Theoretische vakken zijn vakken met hoorcollege die worden afgesloten door een tentamen. Bij de vaststelling van het aantal theoretische vakken zijn vakken die over verschillende periodes doorlopen gerekend als één vak omdat ze over één onderwerp gaan. Het aantal theoretische vakken zegt dus iets over de samenhang van het programma, waarbij een groot aantal theoretische vakken op versnippering duidt. Het aantal eindcijfers hangt af van het aantal tentamens en de compensatiemogelijkheden tussen die tentamens. Wanneer alle tentamencijfers apart gerekend worden zijn er evenveel eindcijfers als tentamens. Wanneer de cijfers voor een aantal deeltentamens (bijvoorbeeld voor vakken die twee of meer periodes doorlopen) gemiddeld worden tot één eindcijfer is het aantal eindcijfers kleiner dan het aantal tentamens.

5. RESULTATEN

Het behaalde aantal studiepunten in het eerste jaar bleek samen te hangen met alle student-gerelateerde factoren. Vrouwen halen in het eerste jaar van hun studie meer punten dan mannen, jonge studenten halen meer punten dan oudere studenten en naarmate men hogere cijfers heeft behaald op de VWO-vakken natuurkunde, wiskunde B en Nederlands haalt men in het eerste studiejaar ook meer punten.

In tabel 4 is de geschatte bijdrage van de student-gerelateerde en de curriculum-gerelateerde factoren aan de studievoortgang weergegeven. Deze tabel moet als volgt gelezen worden. In de eerste kolom staat de schatting van het behaalde aantal studiepunten wanneer we nog niets van de student afweten (het gemiddelde), namelijk 26,54 punten. De variantie op niveau 1 (5,68) gedeeld door de variantie op niveau 1 en 2 samen (5,68 + 173,4) geeft de proportie variantie die toegeschreven kan worden aan het curriculum. Deze berekening laat zien dat het curriculum 3,2% van de variantie in het aantal behaalde studiepunten verklaart. De invloed van het curriculum op de studievoortgang lijkt dus vrij beperkt te zijn. In kolom 2 staat de schatting van het aantal behaalde punten als we alle student-gerelateerde factoren gebruiken als voorspeller. Het verwachte aantal studiepunten is dan de basisscore plus de bijdragen van de verschillende factoren. Geslacht was een dummyvariabele waarbij mannen de code 0 kregen en vrouwen de code 1. We zien dat de bijdrage van geslacht 2,8 punten is, dat wil zeggen dat een vrouw volgens de schatting 2,8 punten meer haalt dan een man met dezelf-

Tabel 4: De bijdrage van student-gerelateerde en curriculum-gerelateerde factoren aan het aantal studiepunten behaald in het eerste jaar.

<i>t</i>	Model 1	Model 2	+ Spreiding	+ Instructie	+ Examen
Basis	26,54	3,60	18,79	19,33	16,22
Geslacht (man = 0)		2,84	2,77	2,77	2,83
Leeftijd - 16		-1,70	-1,69	-1,69	-1,68
Cijfer na - 4		3,85	3,89	3,88	3,88
Cijfer wi B - 4		3,61	3,58	3,58	3,61
Cijfer ne - 4		0,86	0,89	0,88	0,87
Onderwijsweken			* 0,03		
Onderwijsvrije weken			1,28		
Vakken parallel			-4,09		
Uren hoorcollege / 100				* 0,76	
Uren werkcollege / 100				* -1,18	
Aantal theor. vakken				-1,80	
Aantal tentamens					* -0,28
Aantal eindcijfers					-0,74
Variantie niveau 1	173,4	108,1	108	108	108
Variantie niveau 2	5,68	5,39	1,22	1,19	1,78
Deviance	12633,2	11891,6	11878,1	11877,9	11881
Chi ²		+ 741,6	** 13,5	** 13,7	** 10,6
d.f.		5	3	3	2
p		p < 0,001	p < 0,005	p < 0,005	p < 0,005

de leeftijd en dezelfde vwo-cijfers. De minimale leeftijd bij aanvang van de studie was zestien jaar en het minimale vwo-cijfer was een vier. Het onderste gedeelte van deze schalen levert in dit geval dus geen informatie. Om te voorkomen dat de regressiecoëfficiënten van leeftijd en vwo-cijfers erg klein zouden worden zijn deze scores getransformeerd door de minimum waarde van de variabele er vanaf te trekken. Tabel 4 laat zien dat elk jaar dat een student ouder is dan zestien betekent dat er (gemiddeld) 1,7 studiepunten minder gehaald worden. Hogere vwo-cijfers zijn geassocieerd met een toename van het aantal behaalde punten. De meerniveau-analyse maakt het mogelijk een rangordening te maken van de invloed van verschillende voorspellers, voor zover die op dezelfde schaal gemeten zijn. De cijfers op natuurkunde en wiskunde B leveren dus een belangrijkere bijdrage aan de studievoortgang dan het cijfer op Nederlands, wat bij technische opleidingen natuurlijk niet verbazingwekkend is. De basisscore is in model twee aanzienlijk lager dan in de eerste kolom. Dit heeft te maken met de aanzienlijke positieve bijdragen van de vwo-cijfers en het feit dat de meeste studenten vrij hoge cijfers gehaald hebben (zie ook tabel 2). De afname in variantie op stu-

dentniveau door toevoeging van student-gerelateerde variabelen bedraagt 63,3 (173,4 - 108,1). Uitgedrukt als percentage van de variantie in model 1 betekent dit dat de student-gerelateerde variabelen dus 37,7% van de variantie op studentniveau kunnen verklaren.

Vervolgens is het effect van de verschillende curriculumkenmerken geschat. De modellen in kolom 3 tot en met 5 worden steeds door middel van Chi-kwadraat toetsen vergeleken met model 2. In kolom 3 van tabel 4 staat de invloed van de spreidingskenmerken. De toevoeging van spreidingskenmerken verklaart 77% van de variantie op curriculumniveau. Het aantal onderwijsweken per jaar verschilt behoorlijk tussen de drie opleidingen maar blijkt geen effect te hebben op het aantal behaalde studiepunten. Het aantal onderwijsvrije weken blijkt een positief effect te hebben op het aantal behaalde punten, het aantal vakken dat parallel geroosterd is heeft een negatief effect, dat wil zeggen, naarmate er meer vakken parallel geroosterd zijn halen studenten in dat curriculum minder studiepunten. De spreidingskenmerken variëren vooral tussen de opleidingen maar nauwelijks tussen de verschillende studiejaar. In kolom 4 is de invloed van de instructiekenmerken weergegeven. Als we model 4 met model 2 vergelijken blijkt dat de toevoeging van instructiekenmerken 78% van de variantie op curriculumniveau verklaart. Het aantal uren hoorcollege heeft geen effect op het aantal behaalde studiepunten. Het aantal uren werkcollege heeft geen effect op het aantal behaalde punten. Het aantal theoretische vakken per jaar heeft wel een effect, naarmate het aantal theoretische vakken toeneemt haalt men minder punten. Kolom 5 geeft het effect van de formele examenkenmerken. Uit de vergelijking met model 2 blijkt dat de examenkenmerken 67% van de variantie op curriculumniveau verklaren. Het aantal tentamens per jaar had geen effect op het aantal behaalde punten, het aantal eindcijfers wel. Bij een vak dat een aantal onderwijsperiodes doorloopt kunnen de deeltentamens aan het eind van elke periode apart gerekend worden of worden gemiddeld tot één eindcijfer. Bij de laatste mogelijkheid is er compensatie tussen deeltentamens. Naarmate het aantal eindcijfers toeneemt, neemt het aantal studiepunten af. Tenslotte dient te worden opgemerkt dat de verklaarde variantie in de modellen met verschillende sets van curriculum-gerelateerde predictoren steeds ongeveer even groot was en opgeteld boven de 1.0 zou uitkomen. Hoewel de verschillende variabelen binnen een set predictoren niet gecorreleerd waren, was er wel een samenhang tussen variabelen uit verschillende sets. Aangezien de verklarende variabelen in de verschillende modellen niet geheel onafhankelijk van elkaar variëren is er een overlap in de verklaarde variantie van de verschillende modellen.

De logistische meerniveau-analyse liet zien dat leeftijd en de cijfers op wiskunde en natuurkunde significante voorspelers waren van de kans dat een student binnen 1 jaar de studie staakt. De resultaten van deze analyse worden niet in een tabel gegeven omdat de regressiecoëfficiënten bij een logistische regressie niet eenvoudig te interpreteren zijn. We volstaan daarom met een korte weergave van de gevonden effecten. Jonge studenten en studenten met hoge cijfers op wiskunde en natuurkunde hebben minder kans om uit te vallen. Geslacht en het cijfer op Nederlands bleken geen effect te hebben op studiestaking. De organisatie van het curriculum heeft geen effect op de kans dat studenten na 1 jaar uitvallen. Spreidingskenmerken, instructiekenmerken en examenkenmerken hadden geen van allen een effect. Er is een extra analyse uitgevoerd om het verband tussen het aantal behaalde studiepunten en studiestaking te bepalen. Wanneer het aantal behaalde studiepunten wordt toegevoegd als voorspeller op studentniveau blijken de andere studentgebonden factoren (leeftijd, cijfers op wiskunde B en natuurkunde) geen additionele verklarende waarde te hebben. Een student met lage vwo-cijfers die evenveel punten heeft gehaald als een student met hoge vwo-cijfers heeft een even grote kans om uit te vallen. De beste voorspeller voor studiestaking na 1 jaar is dus het aantal behaalde punten in het eerste jaar en het aantal punten kan weer voorspeld worden op basis van de achtergrondgegevens over de student.

6. DISCUSSIE

Studievoortgang van eerstejaars studenten bij drie opleidingen aan de TU Delft hangt samen met leeftijd, geslacht en eindexamencijfers op natuurkunde, wiskunde B en Nederlands. Jonge studenten, vrouwen, en studenten met hoge VWO-cijfers halen meer studiepunten in hun eerste jaar. De studievoortgang van studenten techniek hangt sterk samen met de cijfers op natuurkunde en wiskunde B. Uit ons onderzoek blijkt dat het cijfer op Nederlands een additionele voorspellende waarde heeft. Het verband tussen leeftijd en studievoortgang wordt vaker gevonden (b.v. Jansen, 1996) en kan worden verklaard doordat oudere studenten al eerder in hun onderwijsloopbaan vertraging hebben opgelopen. Leeftijd bij aanvang van de studie zegt dus iets over algemene capaciteiten en inzet. De eindexamencijfers voor exacte vakken kunnen gezien worden als een indicator voor specifieke capaciteiten die nodig zijn voor het succesvol doorlopen van een technische opleiding. Leeftijd en vwo-cijfers op wiskunde B en natuurkunde blijken dan ook voorspellers te zijn van de kans dat een student in de loop van het eerste jaar uitvalt. Jongere studenten en studenten met hoge cijfers voor de exacte vakken hebben minder kans om uit te vallen. De effecten van studentgebonden factoren zijn dus in overeenstemming met de verwachtingen en met de resultaten van ander onderzoek.

Het belangrijkste doel van de analyses was om verbanden te zoeken tussen curriculumgebonden factoren en studievoortgang. Er is gekeken naar drie soorten curriculumkenmerken. De spreiding van onderwijsactiviteiten was zoals verwacht van belang voor de studievoortgang. Alle drie de opleidingen hadden een collegejaar van 42 weken verdeeld over vier onderwijsperiodes maar er waren verschillen tussen de opleidingen wat betreft de indeling van het jaar in onderwijsweken, tentamenweken en onderwijsvrije weken. Het aantal onderwijsweken had geen effect op het aantal behaalde studiepunten, maar het aantal onderwijsvrije weken wel. De aanwezigheid van onderwijsvrije weken had een positief effect op de studievoortgang. Dit is niet in overeenstemming met de veronderstelling dat het inroosteren van onderwijsvrije weken uitstelgedrag in de hand werkt en daardoor de studievoortgang vertraagt. Hierbij moet opgemerkt worden dat van de opleidingen die bij het onderzoek betrokken waren bij slechts één van de opleidingen onderwijsvrije weken ingeroosterd waren. We kunnen daarom niet uitsluiten dat er sprake is van een opleidingseffect. Het positieve effect van onderwijsvrije weken wil natuurlijk niet zegen dat de studievoortgang zal verbeteren naarmate men meer onderwijsvrije weken inroostert. Niettemin kunnen we zeggen dat één onderwijsvrije week voorafgaand aan elke tentamenweek geen aantoonbare negatieve effecten op de studievoortgang heeft. Het aantal parallel geroosterde vakken hangt ook samen met studievoortgang. Hoe meer vakken er parallel geroosterd zijn, hoe minder studiepunten de studenten halen. Dit is geheel in overeenstemming met de verwachtingen en met resultaten van eerder onderzoek. Wanneer er veel vakken parallel geroosterd zijn is de concurrentie tussen vakken groot en bestaat de kans dat studenten hun tijd niet efficiënt over de vakken verdelen.

Naast de spreidingskenmerken zijn er van elk curriculum drie instructiekenmerken in kaart gebracht. Het aantal uren hoorcollege had geen invloed op de studievoortgang van de studenten. Hoewel vaak gedacht wordt dat passieve contacturen zoals hoorcolleges een negatief effect op de studievoortgang kunnen hebben is dit ook in eerder onderzoek niet bevestigd (Jansen, 1996). We moeten daarbij wel bedenken dat het aantal ingeroosterde uren hoorcollege natuurlijk niet gelijk is aan het aantal gevolgde uren hoorcollege. Het is mogelijk dat het aantal colleges dat men overslaat toeneemt naarmate er meer uren zijn ingeroosterd, wat zou verklaren dat het aantal ingeroosterde uren hoorcollege geen verband heeft met studievoortgang. Het aantal uren werkcollege had ook geen effect op studievoortgang. Hoewel dit niet overeenkomt met de verwachting dat actieve contacturen zorgen voor gelegenheid tot oefenen en het krijgen van tussentijdse feedback en daardoor de studievoortgang bevordert is ook in eerder onderzoek (Jansen, 1996) geen positief verband gevonden tussen het aantal uren werkcollege en studievoortgang. Mogelijk maken de studenten weinig gebruik van de werkcolleges of gaat de tijd die aan het volgen van werkcolleges besteed wordt ten koste van zelf-

studie. Het aantal theoretische vakken had wel een consistent effect op de studievoortgang. Theoretische vakken die een aantal periodes doorlopen zijn in de analyse als één vak geteld. Hoe meer theoretische vakken er in het programma zijn opgenomen, hoe minder studiepunten de studenten halen. Dit was in overeenstemming met de verwachtingen. Een groot aantal theoretische vakken zorgt voor een versnippering van het programma. Het is dus beter om vakken te integreren in grotere gehelen met deelttentamens aan het eind van elke onderwijsperiode dan om een groot aantal losstaande kleine vakken te programmeren.

Tenslotte is nagegaan wat de invloed is van formele examenkenmerken op de studievoortgang. Daarbij is gekeken naar het aantal af te leggen (deel)tentamens het aantal eindcijfers. Het aantal tentamens had geen effect op de studievoortgang, het aantal eindcijfers wel. Hoe meer eindcijfers, hoe minder studiepunten men haalde. Het aantal eindcijfers hangt samen met het aantal tentamens en het aantal compensatiemogelijkheden door deelttentamens. Als er weinig of geen compensatiemogelijkheden zijn moeten studenten op (bijna) alle tentamens een voldoende halen, wat een negatief effect heeft op de studievoortgang. Mogelijk moet het effect van het aantal eindcijfers gezien worden in samenhang met het effect van het aantal theoretische vakken. Een groot aantal eindcijfers hangt mogelijk samen met een versnippering van het programma en leidt daardoor tot een tragere studievoortgang.

We kunnen concluderen dat spreidingskenmerken, instructiekenmerken en formele examenkenmerken alle drie van invloed zijn op de studievoortgang. Vanwege de ongunstige verhouding van het aantal predictoren tot de steekproefgrootte was het niet verantwoord meer dan drie curriculum-gerelateerde predictoren tegelijk in het model op te nemen. De spreidingskenmerken, instructiekenmerken en examenkenmerken variëren niet geheel onafhankelijk van elkaar en daardoor is het niet mogelijk om met zekerheid te zeggen welke curriculumkenmerken doorslaggevend zijn voor de studievoortgang. Niettemin bleek een groot deel van de variantie op het curriculumniveau verklaard te kunnen worden door de curriculum-gerelateerde factoren die voor dit onderzoek in kaart zijn gebracht. Hoewel een groot deel van de variantie in het aantal behaalde studiepunten verklaard wordt door individuele verschillen tussen studenten kunnen we concluderen dat de organisatie van het curriculum wel degelijk een bijdrage aan de studievoortgang levert.

Op basis van dit onderzoek kunnen drie aanbevelingen worden gedaan voor de organisatie van het curriculum, teneinde de studievoortgang te bevorderen. Een consistente bevinding in onderzoek naar curriculumkenmerken en studievoortgang die ook in dit onderzoek weer bevestigd werd is dat naarmate het aantal parallel geprogrammeerde vakken toeneemt, de studievoortgang afneemt. De eerste aanbeveling is daarom om zo weinig mogelijk vakken parallel te programmeren. Ten tweede bleek ook in dit onderzoek dat een toename van het aantal theoretische vakken samengaat met een afname van de studievoortgang. Het heeft dus de voorkeur om een beperkt aantal vakken met een grote omvang te programmeren in plaats van een groot aantal vakken met een kleine omvang. Daarnaast is gebleken dat compensatiemogelijkheden door deelttentamens de studievoortgang te bevorderen. De derde aanbeveling is daarom om grotere studieonderdelen met deelvakken te programmeren. In het algemeen kunnen opleidingen de studievoortgang van studenten bevorderen door concurrentie tussen vakken en versnippering van het programma tegen te gaan. De factoren die samenhangen met de studievoortgang hebben geen van allen een effect op de kans dat studenten binnen een jaar uitvallen. Een efficiënte organisatie van het curriculum kan dus de doorstroming verbeteren zonder dat de selectiviteit van de propaedeutische vermindert.

LITERATUUR

- Bax, A.E. (1992). Sneller herkansen leidt tot betere studievoortgang. *Onderzoek van Onderwijs*, februari, 14-15.
- Boom, J. de, Hofman, W.H.A., Van Wensveen, P. & Van den Berg, M.N. (1998). *Studeren en werken in het wetenschappelijk onderwijs. Een onderzoek naar de relatie tussen werken, lenen en studievoortgang*. Rotterdam: Rotterdams Instituut voor Sociaal-wetenschappelijk Beleidsonderzoek (RISBO).
- Bryk, A.S. & Raudenbush, S.W. (1992). *Hierarchical linear models. Application and data analysis methods*. Newbury Park, CA: Sage.
- Cohen-Schotanus, J. (1995). De praktijk van de compensatie. *Onderzoek van Onderwijs*, december, 60-62.
- Dijkman, M.W.J.W. (1991). Computertoetsen houden studenten bij de les. *Onderzoek van Onderwijs*, februari, 12-14.
- Drift, K.D.J.M. van der & Vos, P. (1987). *Anatomie van een leeromgeving. Een onderwijs-economische analyse van universitair onderwijs*. Lisse: Swets & Zeitlinger.
- Gijselaars, W.H. & Schmidt, H.G. (1993). Investeren in instructietijd: spaarzaamheid loont de moeite. *Onderzoek van Onderwijs*, november, 57-60.
- Jansen, E.P.W.A. (1996). *Curriculumorganisatie en studievoortgang*. Studies over Opvoeding en Onderwijs 1, Groningen: Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling (GION).
- Jansen, E.P.W.A. (1997). Curriculumorganisatie en studievoortgang. *Tijdschrift voor onderwijsresearch*, 2, 3-14.
- Johnson, J.L., Bloom, A.M. (1995). An analysis of the contribution of the five factors of personality to variance in academic procrastination. *Personality and Individual Differences*, 18, 127-133.
- KUO (1998). *Kengetallen Universitair Onderwijs*. Utrecht: VSNU.
- Macan, T., Shahani, C., Dipboye, R.L. & Phillips, A. (1990). College students' time management: correlations with academic performance and stress. *Journal of Educational Psychology*, 82, 760-768.
- Macan, T. (1994). Time management: test of a process model. *Journal of Applied Psychology*, 79, 391-391.
- Meerum Terwogt-Kouwenhoven, K. (1990). *Niet gewogen, toch te licht bevonden. Analyse van de rendementsproblematiek aan de universiteit*. Proefschrift, Universiteit van Amsterdam.
- Milgram, N., Marchevski, S. & Sadeh, C. (1995). Correlates of academic procrastination: discomfort, task aversiveness and task capability. *The Journal of Psychology*, 129, 145-155.
- Mislukking en vertraging van de studie* (1959). Verslag van een onderzoek verricht aan de Technische Hogeschool Delft door een werkgroep onder voorzitterschap van Prof. A.D. de Groot.
- Nonis, S.A., Hudson, G.I., Logan, L.B. & Ford, C.W. (1998). Influence of perceived control over time on college students' stress and stress-related outcomes. *Research in Higher Education*, 39, 587-605.
- Prins, J. (1997). *Studieuitval in het wetenschappelijk onderwijs. Studentkenmerken en opleidingskenmerken als verklaring voor studieuitval*. Proefschrift, Katholieke Universiteit Nijmegen. Nijmegen: Nijmegen University Press.
- Prins, J. (1998). Studieuitval in het wetenschappelijk onderwijs. *Onderzoek van Onderwijs*, maart, 8-10.
- Schouwenburg, H.C. (1992). Procrastinators and fear of failure: an exploration of reasons for procrastination. *European Journal of Personality*, 6, 225-236.
- Schouwenburg, H.C. (1994). *Uitstelgedrag bij studenten*. Proefschrift, Rijksuniversiteit Groningen.
- Senécal, C., Koestner, R. & Vallerand, R.J. (1995). Self-regulation and academic procrastination. *The Journal of Social Psychology*, 135, 607-619.
- Snijders, T.A.B. & Bosker, R.J. (1994). Explained variance in two-level models. *Sociological Methods and Research*, 22, 342-363.
- Vos, P. (1992). Het ritme van het rooster. *Onderzoek van Onderwijs*, oktober, 51-53.
- Wilders, P., Peereboom, M. & Staal, R.H.A. (1996). Regelmatige zelfstudie, direct tentamen doen en geen uitstelgedrag: kan dat? *Onderzoek van Onderwijs*, mei, 31-32.

Het ILS-instrument voor leerstijlen: een methodologische kanttekening*

P. Van Petegem en S. De Maeyer
Departement Didactiek en Kritiek, Universiteit Antwerpen

ABSTRACT

The ILS (Inventaris Leerstijlen), developed by Vermunt (1992), has been subject of many validation and replication studies in the Netherlands. These studies mainly focussed on the four typical learning styles. No attention was given to the subscales that are (by definition) the foundation of these learning styles. In this article we argue that the use of PCA for the construction of the subscales may not be taken for granted. Two alternative methods of analysis are suggested: PRINCALS and SEM. We describe the advantages and disadvantages of these methods within the framework of the ILS. Finally we analyse the subscales 'deep processing' and 'selfregulation' as examples, using the three methods of analysis (PCA, PRINCALS and SEM). These analyses are based on the data of students at the University of Antwerp. The three methods of analysis give broadly speaking analogous results. Nevertheless there are some differences between the results of these techniques.

1. INLEIDING

Het aantal generatiestudenten dat jaarlijks niet slaagt in het eerste jaar aan de universiteit is groot, in absolute zowel als in relatieve cijfers (Hendrickx, 1998): het gemiddeld slaagcijfer van de generatiestudenten ingeschreven in het eerste jaar bedroeg in het academiejaar '95-'96 44.01% (p. 17). Daarenboven heeft dit verschijnsel een permanent karakter en daarom spreekt men van een jaarlijks terugkerende 'hecatombe' (NFWO Contactgroep Academisch Onderwijs, 1993).

In het verleden zijn reeds tal van initiatieven ondernomen om het aantal mislukkingen terug te dringen; we verwijzen o.m. naar het tienpuntenprogramma uit het universiteitsdecreet van 1991¹. Verschillende universiteiten leveren dan ook grote inspanningen om studenten te begeleiden.

Op grond van (literatuur)onderzoek leren we dat een mix van oorzaken aan de basis ligt van deze lage slaagcijfers. Verwezen wordt o.m. naar de verkeerde oriëntering van studenten, de foute aanpak bij de studieplanning, motivationele problemen, enzovoort.

Tot nu toe is in Vlaanderen nagenoeg geen onderzoek verricht rond de wijze waarop universiteitsstudenten informatie verwerken bij het studeren als mogelijke verklaring voor het slagen en mislukken. Nochtans reveleert buitenlands onderzoek dat de *leerstijl* van studenten samenhangt met de studieresultaten (Ribich & Schmeck, 1979; Brown, 1987 in Vermunt, 1995; Vermunt, 1992; Santegoets & de Jong, 1993; Gadzella, Ginther & Bryant, 1997).

Om het verband tussen de leerstijl en de studieresultaten na te gaan is er behoefte aan een goede conceptualisatie en operationalisatie van het begrip leerstijl. Een leerstijl moet op één of andere manier meetbaar worden gemaakt aan de hand van een goed instrument. Een in de Nederlandstalige literatuur vaak gebruikt instrument is de ILS (Inventaris LeerStijlen) van Vermunt (1992). Via de ILS wordt een leerstijl gemeten als een combinatie van cognitieve verwerkingsactiviteiten, metacognitieve sturingsactiviteiten, studieopvattingen en studiemotieven.

Dit instrument is vaak voorwerp geweest van replicatiestudies en validiteitsstudies². De replicatiestudies zijn voornamelijk gericht op de vraag of men de vier leerstijlen die Vermunt vooropstelt ook terug vindt in andere data. Ze gaan hoofdzakelijk na in hoeverre een identieke factorstructuur als die van Vermunt (1992) wordt teruggevonden (zie Busato (1998); Busato, Prins, Hamaker & Visser (1995); Schouwenburg (1996); Weersink (1996)). De studies naar de validiteit van ILS richten zich vooral op de toetsing van de externe validiteit door de relaties met andere instrumenten en methoden na te gaan. We verwijzen hier voornamelijk naar Prins, Busato, Hamaker & Visser (1996); Prins, Busato, Elshout & Hamaker (1998); Schouwenburg (1996), Busato (1998).

Kenmerkend voor deze onderzoeken is de focus op de validiteit, bruikbaarheid en repliceerbaarheid van de vier leerstijlen. Weinig aandacht wordt geschonken aan de validiteit en repliceerbaarheid van de subschalen³ die per definitie de bouwstenen zijn van het begrip leerstijl in de ILS. In deze bijdrage zullen we beargumenteren dat onderzoekers de manier waarop deze bouwstenen geconstrueerd zijn niet als vanzelfsprekend mogen nemen. We zullen aantonen dat het maken van een kritische analyse van deze subschalen een onvermijdelijke stap is in het komen tot een valide instrument voor leerstijlen.

2. DE OPBOUW VAN HET ILS INSTRUMENT

Het ILS instrument bestaat uit 120 items (uitspraken) die moeten worden beoordeeld op een vijfpuntenschaal. Deze 120 items kunnen inhoudelijk onderverdeeld worden in vier categorieën: items m.b.t. verwerkingsstrategieën (cognitieve verwerking), items m.b.t. regulatiestrategieën (regulatie), items m.b.t. attitudes t.o.v. het onderwijs (affectieve processen) en items m.b.t. oriëntaties en motivaties voor het leren (mentale leermodellen en leeroriëntaties).

In het onderzoek van Vermunt (1992) worden vier principale componentenanalyses met Varimax rotaties uitgevoerd op de items van de vier onderdelen van ILS. Dit resulteert in 20 subschalen die getest worden op hun betrouwbaarheid aan de hand van Cronbach's alpha analyses.

Vervolgens voert Vermunt een nieuwe principale componentenanalyse met Varimax rotatie uit op de subschalen in de ILS wat resulteert in de vier typische leerstijlen⁴: de betekenisgerichte, de reproductiegerichte, de toepassingsgerichte en de ongerichte leerstijl.

Tabel 1 beschrijft de uitkomst van deze analyse.

3. METHODOLOGISCHE KANTTEKENINGEN BIJ ILS-ONDERZOEK

De meeste onderzoekers die de ILS hanteren, beperken zich tot het uitvoeren van een Cronbach's alpha analyse van de subschalen om vervolgens na te gaan in hoeverre de vier leerstijlen van Vermunt terug te vinden zijn in hun onderzoekspopulatie. Niettemin zijn er een aantal methodologische kanttekeningen te zetten bij deze procedure. In deze bijdrage richten we ons op het creëren van de subschalen. Meer bepaald op het gebruiken van een klassieke principale componentenanalyse (verder PCA genoemd) voor het extraheren van deze subschalen.

Tabel 1: De Varimax oplossing van de analyse van Vermunt (1992). De subschalen met een factorlading hoger dan 0.25 of lager dan -0.25 worden door Vermunt weergegeven.

De betekenisgerichte leerstijl:	De reproductiegerichte leerstijl:
- diepteverwerking	- stapsgewijze verwerking
- concrete verwerking	- externe sturing
- zelfsturing	- certificaat gericht
- opbouw van kennis	- testgericht
- persoonlijke interesse	- opname van kennis
De toepassingsgerichte leerstijl:	De ongerichte leerstijl:
- concrete verwerking	- de ongerichtheid
- beroepsgericht	- ambivalente oriëntatie
- certificaat gericht	- stimulerend onderwijs
- gebruik van kennis	- samen studeren

Een klassieke PCA vertrekt van een aantal veronderstellingen en vereisten voor de data.

Tacq (1997) beschrijft de PCA techniek (met Varimax-rotatie) als:

1. een techniek waarbij één of meerdere latente variabelen worden verondersteld; (zie o.m. Child, 1976; Kline, 1994; Krzanowski W.J. & Marriott, 1994);
2. een techniek waarbij de latente variabelen onafhankelijk zijn van elkaar; (zie o.m. Kline, 1994; Krzanowski W.J. & Marriott, 1994; Rencher, 1995);
3. een techniek waarbij alle variabelen minstens op intervalniveau gemeten zijn; (zie o.m. Kline, 1994; Krzanowski W.J. & Marriott, 1994);
4. een metrische techniek; (zie o.m. Rencher, 1995);
5. een techniek waarbij geen causaliteit wordt verondersteld; (zie o.m. Child, 1976; Rencher, 1995);
6. een exploratieve, eerder dan een confirmatieve techniek; (zie o.m. Kline, 1994; Krzanowski & Marriott, 1994).

Dit brengt de volgende eigenschap met zich mee:

7. een techniek waarbij alle variabelen *laden* op de resulterende componenten. (zie o.m. Kline, 1994).

Binnen de context van de Inventaris LeerStijlen geeft het hanteren van een PCA een aantal problemen.

Een eerste probleem is dat een PCA (met Varimax-rotatie) principale componenten (de subschalen in ILS) extraheert die orthogonaal zijn t.o.v. elkaar. Vertaald naar de situatie van de ILS houdt dit in dat de subschalen per definitie onderling niet correleren. In een volgende fase echter wordt een nieuwe PCA uitgevoerd op schaalniveau om op zoek te gaan naar de typische leerstijlen, wat contradictorisch is.

Het feit dat een PCA variabelen op minstens intervalniveau veronderstelt en dat het een metrische techniek is, vormt een tweede probleem. De items uit het ILS-instrument zijn van het ordinale meetniveau⁵. Bijgevolg is de toepassing van een niet-metrische eerder dan een metrische techniek eerder aangeraden.

Een derde probleem wordt veroorzaakt door de exploratieve aard van PCA⁶. De betekenis van de componenten wordt toegekend op basis van datagedreven resultaten (ladingen). Theoretisch gezien zou een omgekeerde werkwijze met als vertrekpunt de betekenis van de items interessanter zijn. De analyse zou moeten vertrekken van een aantal theoretische (latente) concepten waarop enkel bepaalde items laden. De analyse zou deze theoretische veronderstelling moeten nagaan (confirmatorische analyse).

Deze methodologische bezwaren vragen om een alternatief. In wat volgt worden twee alternatieve methoden naar voor gebracht: PRINCALS en Structural Equation Modelling (SEM).

4. PCA, PRINCALS EN SEM VERGELIJKENDERWIJS

PRINCALS

Als eerste alternatief wordt PRINCALS - de niet-metrische variant van klassieke PCA (zie o.m. Gifi, 1990; de Heus, van der Leeden & Gazendam, 1995; Tacq, 1997) - naar voor geschoven. PRINCALS staat voor principale componenten analyse aan de hand van het 'alternating least squares' algoritme⁷. Deze techniek laat het de onderzoeker toe om zelf het meetniveau van de geobserveerde variabelen te definiëren. Vervolgens wordt een homogeniteitsanalyse uitgevoerd die rekening houdt met het vooropgestelde meetniveau van de geobserveerde variabelen.

Niettemin is PRINCALS ook een exploratieve analysetechniek en resulteert deze techniek in principale componenten die orthogonaal zijn ten opzichte van elkaar (Gifi, 1990). Maar deze methode laat het toe na te gaan in hoeverre de resultaten van een metrische en een niet-metrische techniek van elkaar verschillen.

SEM

Eén van de belangrijkste toepassingen van SEM is confirmatorische factoranalyse (Bollen, 1989; Kelloway, 1998), meer bepaald confirmatorische factoranalyse om meeteigenschappen van bepaalde schalen na te gaan (Kelloway, 1998). Via SEM wordt een a priori gedefinieerd meetmodel getest waarbij zowel de significantie van het model als de significantie van de factorkladingen worden gemeten.

SEM stelt ons tevens in staat om te werken met ordinale variabelen. Hiervoor dient men de polychoric polyserial correlation matrix te analyseren aan de hand van de weighted least squares (WLS) methode, met de asymptotic covariance matrix als wegingsmatrix (Bollen, 1989; Jöreskog & Sörbom, 1996).

Een derde voordeel van SEM is dat men eveneens relaties tussen de gepostuleerde latente variabelen kan opnemen in het model. Deze relaties worden tenslotte ook op hun significantie getest. Binnen het kader van de constructie van de subschalen maakt dit dat het verband tussen de verschillende subschalen mee kan worden getest.

5. TOEPASSINGEN: 'DIEPTEVERWERKING' EN 'ZELFSTURING'

Ter illustratie worden hier zowel PCA, PRINCALS als SEM toegepast op de subschalen diepteverwerking en zelfsturing. In de bespreking van de resultaten van deze drie analyses zullen de verschillende items van beide schalen in korte notatie worden gepresenteerd (zie Tabel 2 voor de betekenis van deze korte notatie).

De geanalyseerde gegevens

De hier gepresenteerde analyses zijn gebaseerd op Vlaamse gegevens. Een vervlaamde⁸ versie van het ILS-instrument werd voorgelegd aan alle studenten uit de eerste kandidatuur van de Universiteit Antwerpen. In totaal hebben 1833 studenten de vragenlijst ingevuld, d.i. 44,81% van de beoogde populatie. Van deze 1833 studenten hebben er 1809 alle items rond diepteverwerking en 1825 alle items rond zelfsturing ingevuld.

Klassieke PCA

Diepteverwerking. Vermunt (1992) voerde een PCA uit op alle 27 items m.b.t. de verwerkingstrategieën die resulteerde in vijf principale componenten. Twee van deze principale componenten groepeerde hij later (op inhoudelijke basis) als de schaal diepteverwerking: de component 'relateren en structureren' en de component 'kritisch verwerken'.

We beperken ons voor deze bijdrage tot de 11 items van de schaal diepteverwerking. Hierbij

Table 2: De geanalyseerde items met hun korte notatie en de schaal waartoe ze behoren volgens de analyses van Vermunt.

Items :	Korte Notatie:	Schaal
Diepteverwerking		
Ik probeer onderwerpen die in een cursus afzonderlijk worden behandeld samen te brengen tot een geheel	Vraag 6	Relateren en structureren
Ik zoek uit wat de overeenkomsten en verschillen zijn tussen de theorieën die in een cursus worden behandeld	Vraag 10	Relateren en structureren
Ik breng specifieke feiten in verband met de grote lijn van een hoofdstuk of een artikel	Vraag 13	Relateren en structureren
Ik probeer studiestof die nieuw voor me is in verband te brengen met kennis die ik al heb over het behandelde onderwerp	Vraag 19	Relateren en structureren
Ik probeer te bedenken wat de onderwerpen uit verschillende hoofdstukken van een studieboek met elkaar te maken hebben	Vraag 25	Relateren en structureren
Ik vergelijk mijn visie op een onderwerp uit de cursus met de visie van de auteurs van het behandelde boek	Vraag 29	Kritisch verwerken
Ik probeer voor mezelf een totaalbeeld van een cursus op te bouwen	Vraag 34	Relateren en structureren
Ik vergelijk conclusies uit verschillende hoofdstukken met elkaar	Vraag 35	Relateren en structureren
Ik ga na of de conclusies van de auteurs van een studieboek logisch volgen uit de feiten waarop ze zijn gebaseerd	Vraag 39	Kritisch verwerken
Ik trek mijn eigen conclusies op basis van de gegevens die in een cursus worden gepresenteerd	Vraag 43	Kritisch verwerken
Ik probeer de interpretaties van deskundigen kritisch te benaderen	Vraag 49	Kritisch verwerken
Zelfsturing		
Ik bestudeer naast de examenstof ook andere literatuur die met de cursusinhoud te maken heeft	Vraag 16	Inhoud
Om mijn leervordering te toetsen probeer ik na het bestuderen van een studieboek de hoofdzaken ervan in eigen woorden te formuleren	Vraag 21	Proces
Als ik aan een nieuw hoofdstuk of artikel begin, denk ik eerst na over de manier waarop ik dat het best kan bestuderen	Vraag 24	Proces
Ik doe meer dan wat van me wordt gevraagd in een cursus	Vraag 28	Inhoud
Als ik moeite heb met een stuk studiestof, probeer ik te analyseren waarom dat moeilijk voor me is	Vraag 31	Proces
Om mijn leervordering te toetsen probeer ik een antwoord te formuleren op vragen over de studiestof die ik zelf bedenkt	Vraag 36	Proces
Ik voeg uit andere bronnen iets aan de studiestof toe	Vraag 42	Inhoud
Om te toetsen of ik de studiestof beheers, probeer ik andere voorbeelden en problemen te bedenken dan die in het studiemateriaal of door de docent worden gegeven	Vraag 46	Proces
Om mijn leervordering te toetsen probeer ik na een paragraaf de inhoud in mijn eigen woorden weer te geven	Vraag 50	Proces
Ik richt me bij het studeren ook op leerdoelen die niet door de docent worden gegeven, maar die ik mezelf stel	Vraag 51	Proces
Als ik een stuk studietekst niet goed begrijp, zoek ik er andere literatuur bij over dat onderwerp	Vraag 54	Inhoud

gaan we na in hoeverre een PCA op deze elf items resulteert in twee afzonderlijke principale componenten. De resultaten van deze analyse worden gepresenteerd in Tabel 3.

Deze analyse geeft bijna dezelfde resultaten als die van Vermunt⁹. Drie items vertonen echter een dubbele lading: vraag 13, vraag 19 en vraag 39. De dubbele ladingen veranderen weinig aan de inhoudelijke interpretatie van de twee componenten. De eerste component kunnen we nog steeds interpreteren als de component relateren en structureren. Vraag 39 verwijst naar het op zoek gaan van de logica van een studieboek en het daarna nagaan van de conclusies van de auteurs op basis van deze logica. Dit is ook een actie die we kunnen interpreteren als het relateren en structureren van de studiestof.

De tweede component duidt eveneens op het fenomeen 'kritisch verwerken'. Zowel vraag 13 als vraag 19 verwijzen naar het aanwenden van externe bronnen en het relateren ervan aan de nieuwe studiestof. Dit wijst op een kritische houding ten opzichte van de studiestof zoals gepresenteerd door de docent.

Tabel 3: Resultaten van de PCA op de 11 items i.v.m. diepteverwerking op basis van de gegevens van 1809 eerstejaars studenten van de Universiteit Antwerpen.

Diepteverwerking			
	Component 1	Component 2	
Vraag 6	.701		
Vraag 10	.681		
Vraag 13	.542	.289	
Vraag 19	.471	.308	
Vraag 25	.726		
Vraag 29		.738	
Vraag 34	.699		
Vraag 35	.700		
Vraag 39	.253	.559	
Vraag 43		.633	
Vraag 49		.767	
Eigenwaarde	4.046	1.222	
% Verklaarde variantie	36.785	11.113	
Totaal % verklaarde variantie:	47.898		
Zelfsturing			
	Component 1	Component 2	Component 3
Vraag 16	.786		
Vraag 21		.810	
Vraag 24			.681
Vraag 28	.627		
Vraag 31			.807
Vraag 36		.577	
Vraag 42	.792		
Vraag 46	.310	.541	
Vraag 50		.812	
Vraag 51	.393	.345	
Vraag 54	.665		
Eigenwaarde	2.397	2.141	1.290
% Verklaarde variantie	21.791	19.465	11.727
Totaal % verklaarde variantie:	52.982		

Zelfsturing. De analyses van Vermunt (1992) resulteerde m.b.t. de sturingsstrategieën in de schalen 'zelfsturing - inhoud' en 'zelfsturing - proces'. Hier voeren we een PCA uit op de 11 items m.b.t. zelfsturing. Tabel 3 geeft eveneens de resultaten weer van deze analyse.

In tegenstelling tot de analyse van Vermunt, resulteert deze analyse niet in twee, maar in drie principale componenten die samen bijna 53% van de variantie verklaren. Vragen 46 en 51 vertonen daarnaast een dubbele lading op de eerste twee componenten. Op de derde component laden slechts twee vragen (24 en 31).

De eerste component kan geïnterpreteerd worden als de schaal 'zelfsturing - inhoud'. Component twee lijkt de schaal 'zelfsturing - proces' te onderscheiden. De derde component zouden we 'zelfsturing (analyse van) - studiemethode' kunnen noemen.

PRINCALS

Diepteverwerking. Een eerste alternatieve analysetechniek is PRINCALS. We voeren in deze stap de elf items in die volgens Vermunt het concept 'diepteverwerking' meten. In PRINCALS moet de gebruiker zelf het aantal verwachte dimensies (latente componenten) opgeven. In dit geval wordt gevraagd naar een oplossing met twee principale componenten met ordinale restricties.

Uit de analyse blijkt vooreerst dat de vooropgestelde twee dimensies 'nodig zijn' aangezien de eigenwaarden van beide dimensies hoger zijn dan $0,09^{10}$ (0,377 voor dimensie één, 0,112 voor de tweede dimensie). Het is dus relevant om de items rond 'diepteverwerking' in te delen in twee schalen. De som van de eigenwaarden bedraagt 0,489. Er wordt m.a.w. 48,9% van de variantie in de getransformeerde variabelen verklaard (de Heus, van der Leeden & Gazendam, 1995).

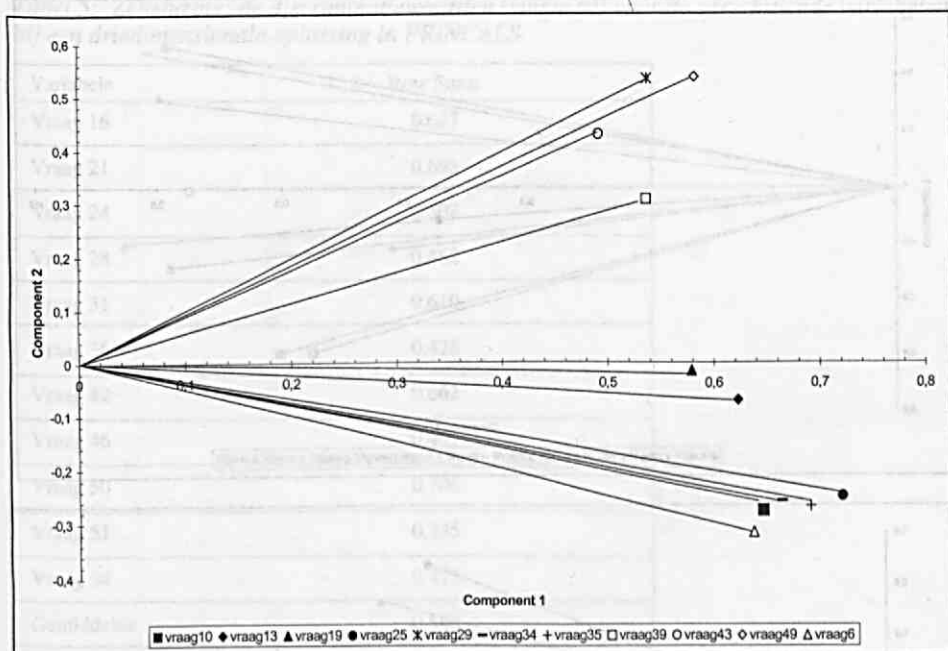
Om de resultaten van de analyse te interpreteren is Figuur 1 zeer nuttig. In deze figuur worden de componentladingen van de verschillende items in een tweedimensionale ruimte weergegeven en verbonden met de oorsprong. Hieruit blijken er twee groepen van lijnen te zijn die dicht tegen elkaar liggen: enerzijds de groep met vragen (6, 10, 25, 34 en 35) en anderzijds vragen (29, 43 en 49). Vragen 13 en 19 liggen tussen beide groepen in, ze passen minder in dit model. Vraag 39 is dan weer een twijfelgeval: ze leunt aan bij de groep van vragen (29, 43 en 49) maar hoort er niet echt bij.

De lengte van de lijn is recht evenredig met de hoeveelheid variantie die verklaard wordt door deze tweedimensionale oplossing (de Heus, van der Leeden & Gazendam, 1995). Hieruit leiden we af dat vragen 13, 19 en (in mindere mate) 39 minder goed passen bij deze oplossing. Dit wordt bevestigd door de Row Sums (zie Tabel 4): vragen (13, 19 en 39) scoren lager in vergelijking met de andere items.

Zelfsturing. Voor de schaal zelfsturing werd zowel een twee- als een driedimensionale analyse uitgevoerd. In de driedimensionale oplossing zijn de eigenwaarden van de drie dimensies hoger dan 0,09 (0,29 voor de eerste, 0,15 voor de tweede en 0,095 voor de derde dimensie). Dit wijst erop dat de drie dimensies nodig zijn om de data goed te vatten.

Figuur 2 geeft enerzijds de plot van de componentladingen van de items op de eerste en de tweede component en anderzijds de plot van de componentlading op de eerste en de derde component. De tweede component discrimineert duidelijk tussen de vragen 16, 28, 42, 54 (zelfsturing - inhoud) en de vragen 21, 24, 31, 36, 46, en 50 (zelfsturing - proces). Vraag 51 ligt tussen beide groepen in en vragen 24 en 31 passen slecht in deze plot (zie lengte van de lijn). De derde component discrimineert tussen vragen 24, 31 en de overige vragen.

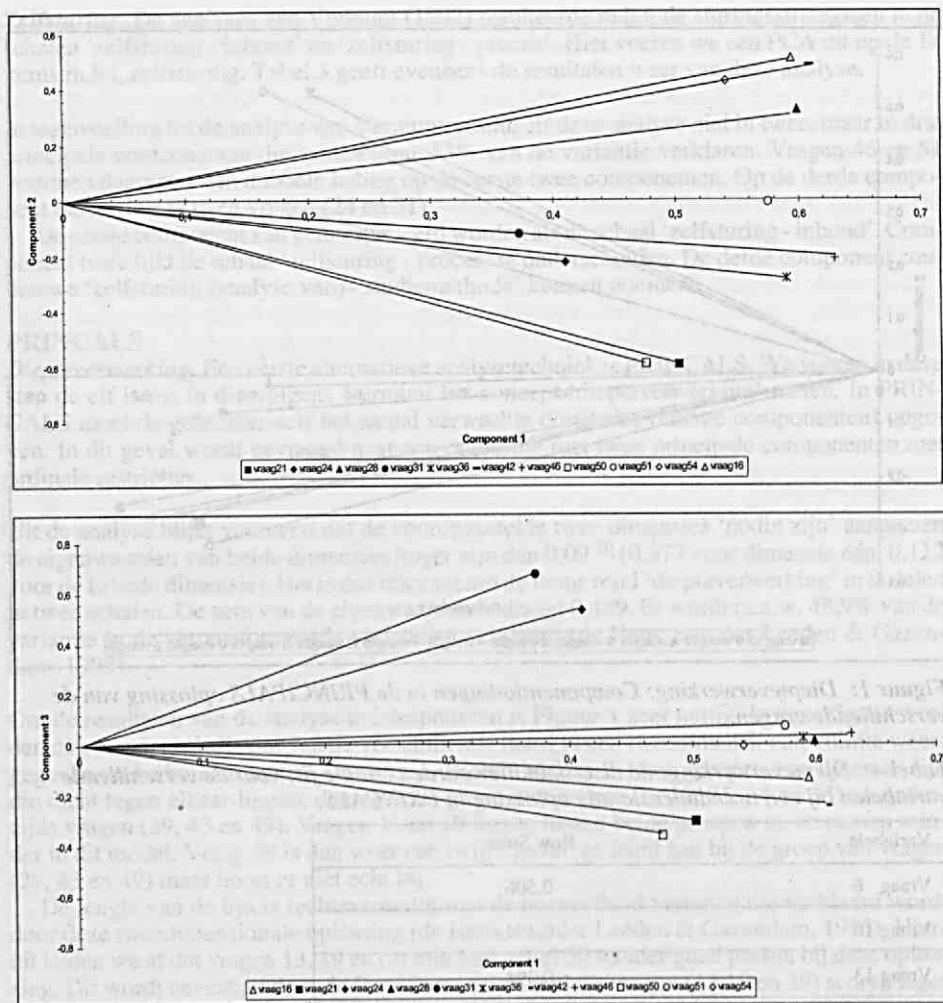
De Row Sums (single fit) (zie Tabel 5) duiden erop dat vraag 51 minder goed past binnen deze oplossing.



Figuur 1: Diepteverwerking: Componentladingen in de PRINCIPALS oplossing van de verschillende vragen.

Tabel 4: Diepteverwerking: de discriminatiewaarden (single fit) voor de verschillende variabelen bij een tweedimensionale oplossing in PRINCALS.

Variabele	Row Sums
Vraag 6	0.506
Vraag 10	0.494
Vraag 13	0.394
Vraag 19	0.335
Vraag 25	0.582
Vraag 29	0.573
Vraag 34	0.508
Vraag 35	0.550
Vraag 39	-0.382
Vraag 43	0.424
Vraag 49	0.625
Gemiddelde	0.489



Figuur 2: Zelfsturing: Componentladingen in de PRINCIPALS oplossing van de verschillende vragen (3 componenten).

SEM¹¹

Diepteverwerking. De tweede alternatieve techniek die werd voorgesteld was 'structural equation modeling'. Hiervoor maken we gebruik van het softwarepakket LISREL (Jöreskog & Sörbom, 1996). Figuur 3 geeft het a priori veronderstelde model (model D1) op basis van de analyses van Vermunt. Dit model analyseren we aan de hand van de polychoric polyserial correlation matrix van de verschillende items, met de weighted least squares (WLS) methode als algoritme en de asymptotic covariance matrix als wegingsmatrix (Bollen, 1989; Jöreskog & Sörbom, 1996).

In Tabel 6 worden de passingresultaten van model D1 gegeven. Volgens de Chi-kwadraat test blijkt dit geen goed model te zijn. De RMR en de AGFI wijzen echter wel op deugdelijkheid:

Tabel 5: Zelfsturing: de discriminatiewaarden (single fit) voor de verschillende variabelen bij een driedimensionale oplossing in PRINCALS.

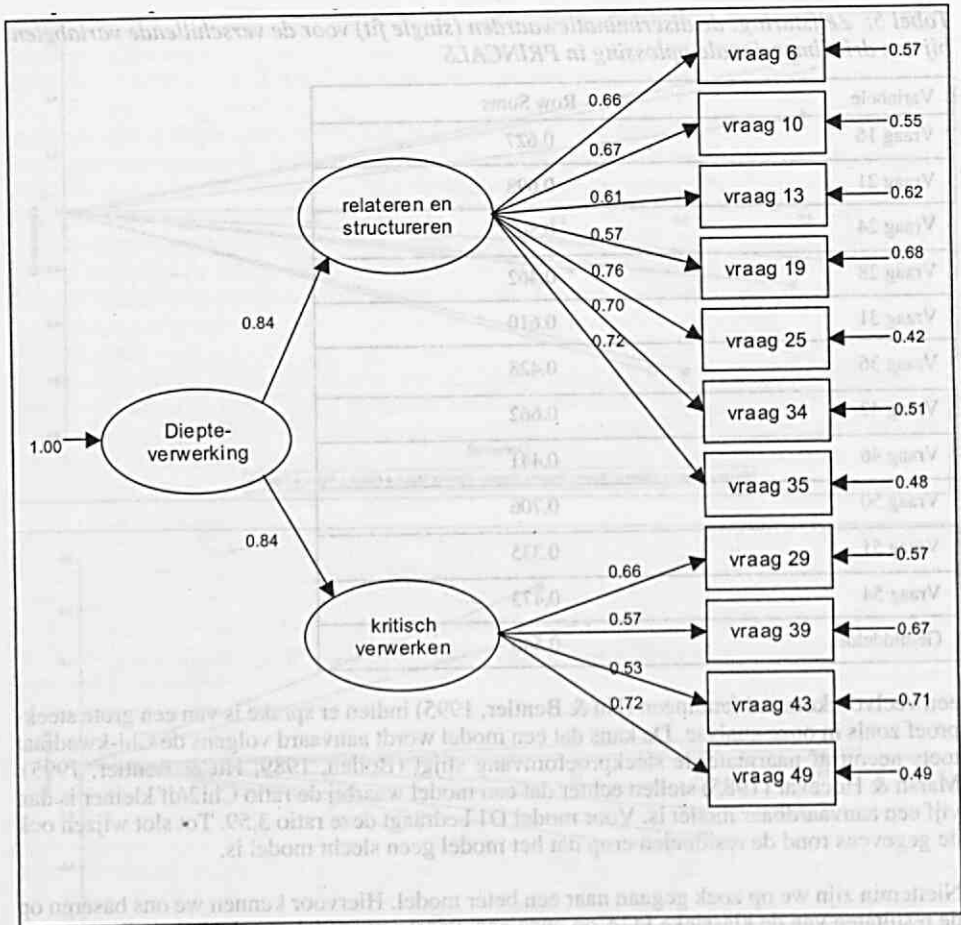
Variabele	Row Sums
Vraag 16	0.627
Vraag 21	0.698
Vraag 24	0.502
Vraag 28	0.462
Vraag 31	0.610
Vraag 36	0.428
Vraag 42	0.662
Vraag 46	0.441
Vraag 50	0.706
Vraag 51	0.335
Vraag 54	0.473
Gemiddelde	0.540

een veelvoorkomend fenomeen (Hu & Bentler, 1995) indien er sprake is van een grote steekproef zoals in onze analyse. De kans dat een model wordt aanvaard volgens de Chi-kwadraat toets neemt af naarmate de steekproefomvang stijgt (Bollen, 1989; Hu & Bentler, 1995). Marsh & Hocevar (1985) stellen echter dat een model waarbij de ratio Chi²/df kleiner is dan vijf een aanvaardbaar model is. Voor model D1 bedraagt deze ratio 3,59. Tot slot wijzen ook de gegevens rond de residuen erop dat het model geen slecht model is.

Niettemin zijn we op zoek gegaan naar een beter model. Hiervoor kunnen we ons baseren op de resultaten van de klassieke PCA op onze gegevens: een model met dubbele ladingen voor vraag 13, vraag 19 en vraag 39 (model D2, zie Figuur 4)¹². Tabel 6 maakt duidelijk dat dit model beter past bij de gegevens dan model D1: zowel de AIC als de CAIC waarden zijn duidelijk lager in het tweede model dan in het eerste. Niettemin zijn de resultaten van de modelpassing opnieuw tegenstrijdig: volgens de Chi-kwadraat toets gaat het om een slecht model, de andere toetsen spreken zulks tegen.

We hebben tenslotte een derde model getest: model D2 zonder vraag 35 (model D3). Dit aangezien de modification indices duidelijk maakten dat er een covariantie tussen de meetfouten van vraag 35 en twee andere vragen diende te worden opgenomen om tot een goed passend model te komen. De resultaten van de modelpassing (zie Tabel 6) voor dit model duiden op een zeer goed passend model. Tevens wijzen de AIC en de CAIC waarden erop dat dit model een beter model is dan de vorige twee modellen.

Zelfsturing. M.b.t. de schaal zelfsturing werd vooreerst het theoretisch veronderstelde model (zie Vermunt) getest (Figuur 5, model Z1). Tabel 7 geeft de passingresultaten voor dit model. Zowel de Chi-kwadraat toets als de ratio Chi²/df wijzen op een slecht model. Het theoretisch model van Vermunt wordt verworpen. Een eerste alternatief model is het model gebaseerd op de klassieke PCA (zie Figuur 6, model Z2). Echter, dit model blijkt ook niet goed te passen bij de gegevens.

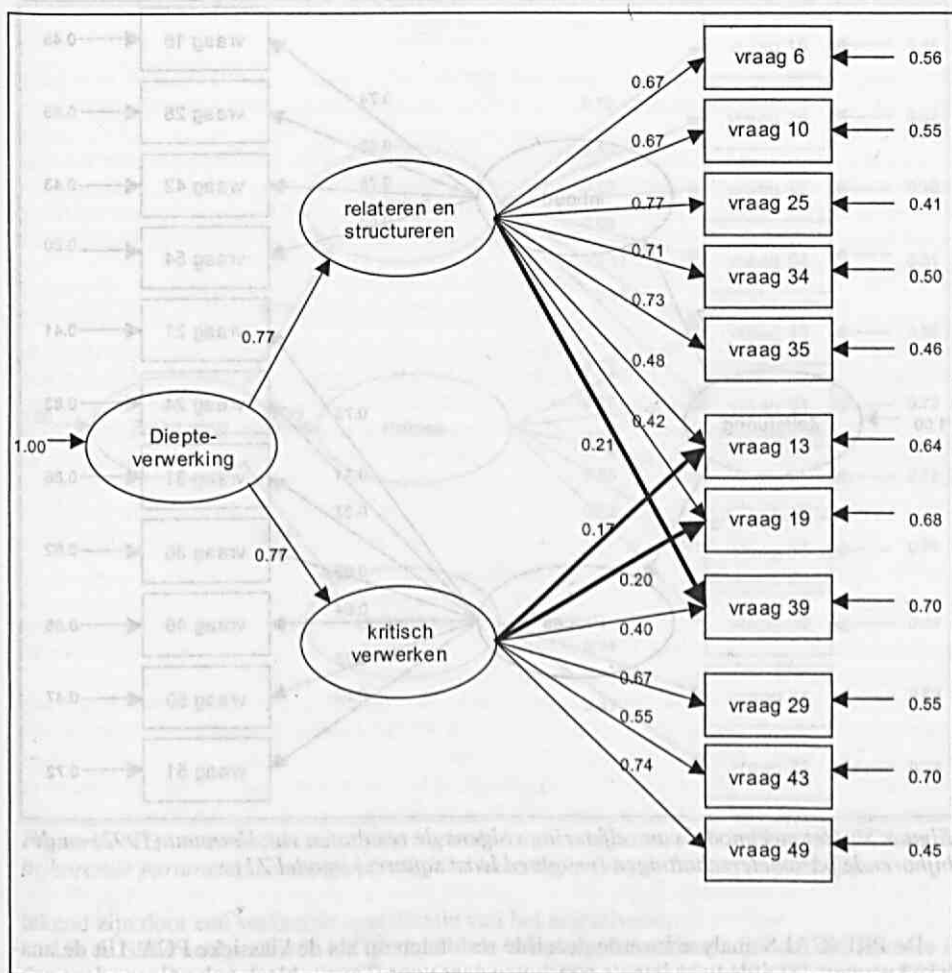


Figuur 3: Het meetmodel van diepte-verwerking volgens de resultaten van Vermunt (1992) en de bijbehorende parameterschattingen (weighted least squares) (model D1).

De Modification Indices bij de beide modellen suggereren een covariantie tussen de meetfouten van vragen 21 en 50. Inhoudelijk is deze covariantie ook zinvol; beide vragen verwijzen naar het persoonlijk evalueren van de leervordering. Modellen Z3 en Z4 zijn respectievelijk de modellen Z1 en Z2 met daarin deze covariantie opgenomen. Tabel 7 wijst erop dat

Tabel 6: Passingresultaten van de drie modellen m.b.t. Diepte-verwerking, getest in LISREL.

N=1809	Chi ²	df	Chi ² /df	p-waarde	AGFI	CFI	Hoogste Residu	AIC	CAIC
Model D1	157,79	44	3,59	0,000	0,99	0,99	0,09	201,79	344,97
Model D2	88,42	40	2,21	0,000	0,98	1,00	0,08	140,42	309,63
Model D3	40,48	31	1,3	0,119	0,99	1,00	0,06	110,00	244,68



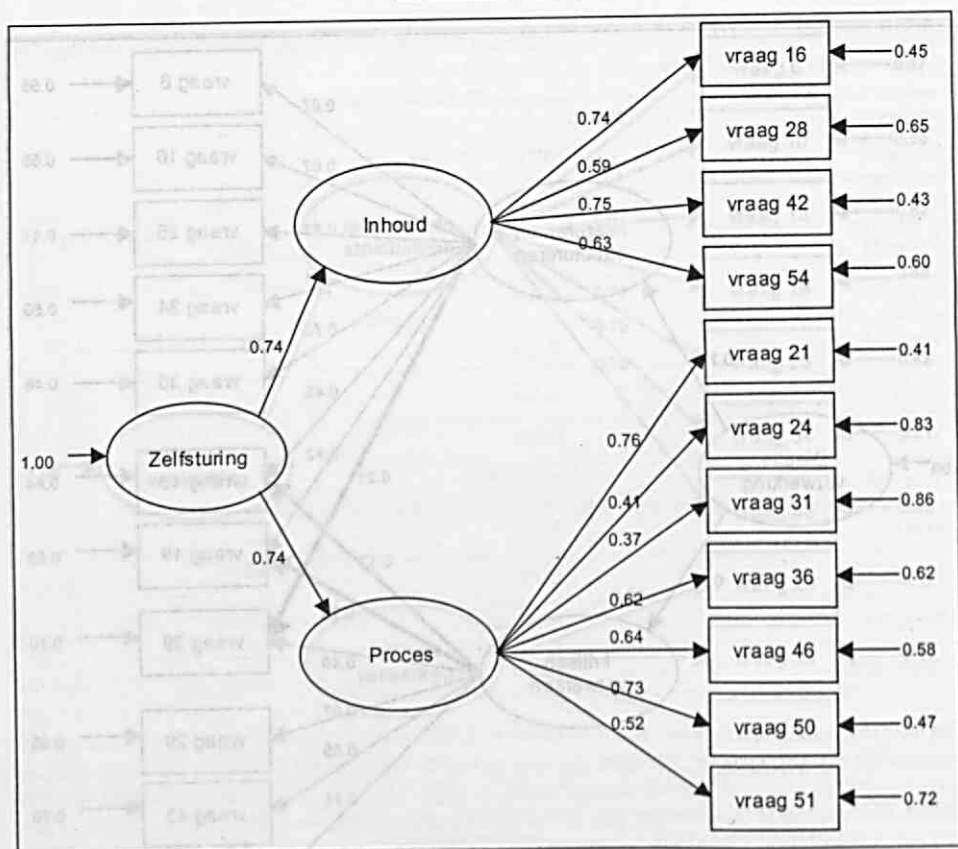
* De vette pijlen duiden de veranderingen aan t.o.v. model 1.

Figuur 4: Het meetmodel van diepteverwerking volgens de resultaten van de klassieke PCA en de bijhorende parameterschattingen (weighted least squares) (zie boven) (model D2); de vette pijlen duiden de verandering aan t.o.v. model 1.

model Z3 een redelijk goed model is, maar dat model Z4 het beste model is (zie AIC en CAIC-waarden en Chi²/df). Het model met drie componenten, een dubbele lading voor vragen 46 en 51 en een covariantie tussen de meetfouten van vragen 21 en 50 past het beste bij de data.

De resultaten naast elkaar gelegd

Diepteverwerking. Uit de klassieke PCA is gebleken dat de door Vermunt gevonden latente componenten ook terug te vinden zijn in de gegevens die we hier analyseerden. Wel bleken drie items (vraag 13, vraag 19 en vraag 39) te laden op beide componenten. Dit maakt dat deze drie items niet specifiek 'relateren en structureren' meten noch dat ze specifiek 'kritisch verwerken' meten, ze zijn een indicatie van beide.

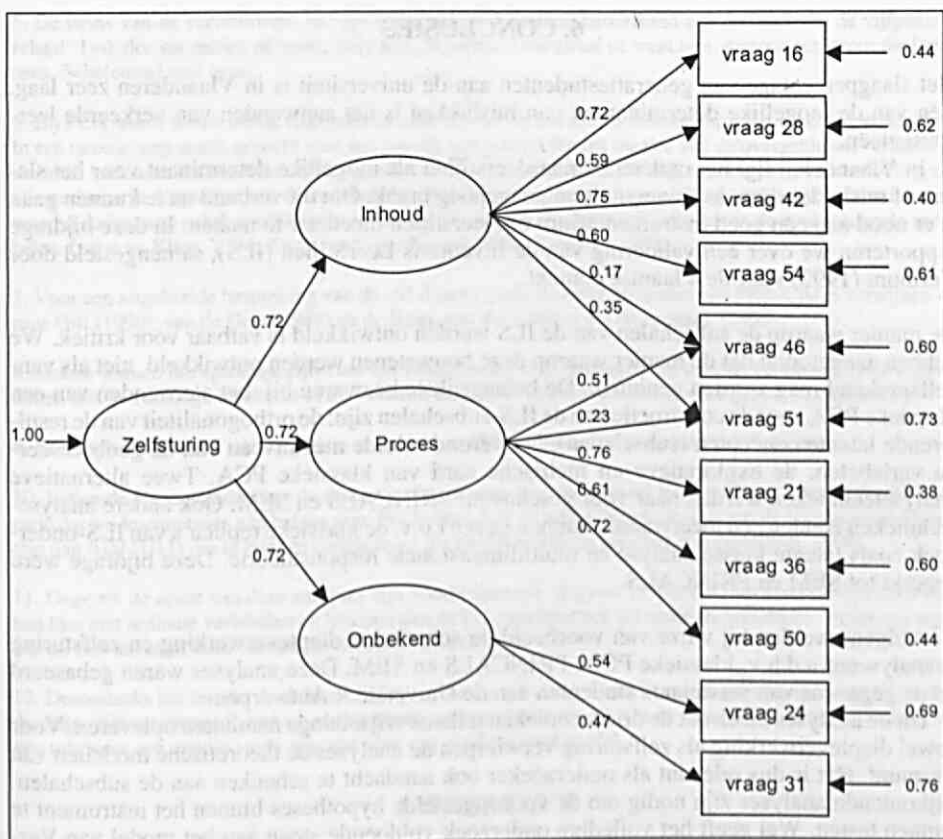


Figuur 5: Het meetmodel van zelfsturing volgens de resultaten van Vermunt (1992) en de bijhorende parameterschattingen (weighted least squares) (model Z1).

De PRINCALS analyse leverde dezelfde resultaten op als de klassieke PCA. Uit de analyse kwamen dezelfde twee latente concepten naar voor. Tevens bleek ook volgens deze analyse dat de drie items (vraag 13, vraag 19 en vraag 39) niet volledig pasten in de PRINCALS oplossing. Bijgevolg kunnen we concluderen dat de resultaten van de klassieke PCA niet ver-

Tabel 7: Passingresultaten van de vier modellen m.b.t. Zelfsturing, getest in LISREL.

N=1825	Chi ²	df.	Chi ² /df	p-waarde	AGFI	CFI	Hoogste Residu	AIC	CAIC
Model Z1	447,97	44	10,18	0,000	0,91	0,96	0,21	491,97	635,17
Model Z2	372,93	42	8,88	0,000	0,93	0,97	0,18	420,93	577,15
Model Z3	221,95	43	5,16	0,000	0,96	0,98	0,14	267,95	417,66
Model Z4	167,93	41	4,10	0,000	0,97	0,99	0,11	217,93	380,67



Figuur 6: Het meetmodel van zelfsturing volgens de resultaten van de klassieke PCA en de bijhorende parameterschattingen (weighted least squares) (zie boven) (model Z2).

tekend zijn door een verkeerde specificatie van het meetniveau.

Uit de SEM analyse bleek het model D3 de beste representatie te zijn van de structuur in de gegevens. Deze structuur is analoog aan de resultaten van de klassieke PCA en de PRINCALS analyse. Eén belangrijk verschil echter is dat in dit model de hypothese getest wordt dat beide latente concepten inderdaad een algemene tendens tot diepteverwerking meten (deze hypothese konden we overigens niet verwerpen). Tevens bleek uit de SEM analyse dat om tot een perfect passend model te komen de vraag 35 uit het model geweerd moest worden van- toe een algemeen karakter. De andere analysemethoden wezen niet op deze beperking.

Zelfsturing. Zowel de klassieke PCA, de PRINCALS als de SEM verwierpen het theoretisch model van Vermunt. De drie technieken resulteerden in een model met drie componenten, waarvan de derde component bestaat uit vragen 31 en 24 (Analyse van de studiemethode). Vragen 46 en 51 vertonen bij alle drie de analyses een dubbele lading op zowel 'zelfsturing - inhoud' als 'zelfsturing - proces'. Tot slot wees de SEM analyse op een covariantie tussen vragen 21 en 50 (toetsing van leervordering).

6. CONCLUSIES

Het slaagpercentage van generatiestudenten aan de universiteit is in Vlaanderen zeer laag. Eén van de mogelijke determinanten van mislukken is het aanwenden van verkeerde leerstrategieën.

In Vlaanderen ligt het onderzoek naar leerstijlen als mogelijke determinant voor het slaagen of mislukken aan de universiteit nog nagenoeg braak. Om dit verband na te kunnen gaan is er nood aan een goed instrumentarium om leerstijlen meetbaar te maken. In deze bijdrage rapporteren we over een validering van de Inventaris LeerStijlen (ILS), samengesteld door Vermunt (1992) voor de Vlaamse context.

De manier waarop de subschalen van de ILS werden ontwikkeld is vatbaar voor kritiek. We hebben aangetoond dat de manier waarop deze bouwstenen werden ontwikkeld niet als vanzelfsprekend mag worden genomen. De belangrijkste bezwaren bij het aanwenden van een klassieke PCA voor de constructie van de ILS-subschalen zijn: de orthogonaliteit van de resulterende latente concepten (subschalen); het veronderstelde meetniveau van de geobserveerde variabelen; de exploratieve en metrische aard van klassieke PCA. Twee alternatieve analysetechnieken werden naar voor geschoven: PRINCALS en SEM. Ook andere analysetechnieken zouden een meerwaarde kunnen geven t.o.v. de klassieke replica's van ILS-onderzoek zoals latente klasse analyse en multidimensionele responstheorie. Deze bijdrage werd beperkt tot SEM en PRINCALS.

Vervolgens werden bij wijze van voorbeeld de subschalen diepteverwerking en zelfsturing geanalyseerd a.d.h.v. klassieke PCA, PRINCALS en SEM. Deze analyses waren gebaseerd op de gegevens van eerstejaars studenten aan de Universiteit Antwerpen.

Uit de analyses bleek dat de drie technieken telkens vrij analoge resultaten opleverden. Voor zowel diepteverwerking als zelfsturing verwierpen de analyses de theoretische modellen van Vermunt. Het is dus relevant als onderzoeker ook aandacht te schenken aan de subschalen. Bijkomende analyses zijn nodig om de vooropgestelde hypothesen binnen het instrument te kunnen testen. Wel geeft het volledige onderzoek voldoende steun aan het model van Vermunt zodat toepassing van het instrument in de Vlaamse context gerechtvaardigd is.

Ten slotte kunnen we concluderen dat de drie gehanteerde analysetechnieken (PCA, PRINCALS en SEM) complementair zijn. Samen geven ze een duidelijker en betrouwbaarder beeld van de onderliggende structuur in de 120 uitspraken.

NOTEN

* (Deze bijdrage kwam tot stand binnen het onderzoeksproject 'Leerstijlgebonden determinanten van slagen en mislukken van eerstejaars generatiestudenten', gefinancierd vanuit de Kleine Projecten Onderzoeksraad 1998 van de UIA (Universiteit Antwerpen).

1. Decreet van 12 juni 1991 betreffende de universiteiten in de Vlaamse Gemeenschap *Belgisch Staatsblad*, 4 juli 1991, pp. 14907-14936.
2. Voor een overzicht zie Boekaerts, Otten & Simons (1997).
3. We verwijzen naar de summierende bespreking van het instrument in het deel twee van deze bijdrage.
4. Wierstra (1996) wees er echter op dat het onjuist is om de factoren te bestempelen als leerstijlen aangezien de factoranalyse doelt op de items en niet op de personen (studenten).

5. De items van de verwerkings- en regulatiestrategieën worden beoordeeld aan de hand van de vijfpunten-schaal: 1=ik doe dit zelden of nooit, 2=ik doe dit soms, 3=neutraal of weet niet, 4=voor een groot deel eens, 5=helemaal eens.
6. Bij PCA wordt in een eerste stap gezocht naar de component die de meeste variantie verklaart in de items. In een tweede stap wordt gezocht naar een tweede component die het meeste van de overgebleven variantie in de items verklaart; dit resulteert in niet-gelijkwaardige componenten. Daarbij is het bij PCA de bedoeling om alle variantie in de geobserveerde variabelen te verklaren. Dit staat in tegenstelling tot klassieke factoranalyse waarbij de factoren enkel moeten instaan voor de gemeenschappelijke variantie tussen de geobserveerde variabelen. (zie o.m. Kline, 1994; Krzanowski & Marriott, 1994; Rencher, 1995).
7. Voor een uitgebreide bespreking van de -ALS technieken en meer bijzonder van PRINCALS verwijzen we naar Gifi (1990), van de Geer (1988) en de Heus, van der Leeden & Ganzendam (1995).
8. Enkele termen die specifiek zijn voor het Nederlandse onderwijssysteem werden vervangen door equivalenten met een analoge betekenis in de Vlaamse context (bv. tentamen werd vervangen door examen).
9. Deze oplossing met Varimax rotatie staat in voor 47,9% van de totale variantie in de items.
10. Indien de eigenwaarden van de dimensies hoger zijn dan $1/p$, waarbij p =aantal items, dan is dit een indicatie dat beide dimensies van belang zijn. We hebben hier 11 items, bijgevolg moeten de eigenwaarden hoger zijn dan 0,09 (1/11) om te kunnen concluderen dat beide dimensies relevant zijn (zie SPSS, 1998).
11. Gegeven de opzet van deze analyses zijn voorbereidende analyses (PRELIS) niet noodzakelijk. We werken hier met ordinale variabelen en beschouwen ze in de analyse ook als ordinale variabelen. Bijgevolg moeten we niet testen of deze variabelen normaal verdeeld zijn.
12. Desondanks het feit dat de dubbele ladingen van zowel vraag 13 als vraag 19 inhoudelijk moeilijk te duiden zijn, zijn we datagedreven op zoek gegaan naar een beter model. Uit de data blijken deze dubbele ladingen trouwens een *conditio sine qua non* voor een goed passend model.

LITERATUUR

- Boekaerts M., Otten R. & Simons R.-J. (1997). Een onderzoek naar de bruikbaarheid van de ILS - Leerstijlen in de onderbouw van het voortgezet onderwijs. *Tijdschrift voor Onderwijsresearch*, 22(2), pp. 15-36.
- Bollen K.A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.
- Brown A.L. (1987). Learning strategies, student motivation patterns, and subjectively perceived success. In Kirby J.R. (Ed.), *Cognitive Strategies and Educational Performance* (pp. 111-134). New York: Academic Press.
- Busato V.V. (1998). *Leerstijlen nader geanalyseerd*. Academisch Proefschrift: Universiteit van Amsterdam.
- Busato V.V., Prins F.J., Hamaker C. & Visser K.H. (1995). Leerstijlonderzoek gerepliceerd; de samenhang tussen intelligentie en leerstijlen. *Tijdschrift voor Onderwijsresearch*, 20(4), pp. 332-340.
- Child D. (1976). *The essentials of Factor Analysis*. London: Holt, Rinehart and Winston.
- Decreet van 12 juni 1991 betreffende de universiteiten in de Vlaamse Gemeenschap, *Belgisch Staatsblad*, 4 juli 1991, pp. 14907-14936.
- Entwistle N. (1988). Motivational factors in students' approaches to learning. In Schmeck R.R. (Ed.), *Learning Strategies and Learning Styles* (pp. 21-51). New York: Plenum Press.
- Gadzella B.M., Ginther D.W. & Bryant W. (1997). Prediction of performance in an academic course by scores on measures of learning style and critical thinking. *Psychological Reports*, 81, pp. 595-602.
- Geer J.P. van de (1988). *De analyse van categorische gegevens*. Deventer: Van Loghum Slaterus.
- Gifi A. (1990). *Nonlinear multivariate analysis*. Chichester: Wiley.
- Hendrickx V. (1998). De eerstejaarsstudent aan de Vlaamse universiteiten in het academiejaar 1996/97, in: Ministerie van de Vlaamse Gemeenschap, Departement Onderwijs (red.). *Vlaamse onderwijsindicatoren in internationaal perspectief. Editie 1998* (pp. 213-223). Brussel: Departement Onderwijs.
- Heus P. de, Leeden R. van der & Ganzendam B. (1995). *Toegepaste data-analyse. Technieken voor niet-experimenteel onderzoek in de sociale wetenschappen*. Utrecht: Lemma.

- Hu Li-Tze & Bentler P.M. (1995). Evaluating Model Fit. In Hoyle K. (red.). *Structural Equation Modeling - Concepts, Issues, and applications* (pp. 76-99). Thousand Oakes: Sage.
- Jöreskog K. G. & Sörbom D. (1996). *LISREL 8 - User's Reference Guide*. United States of America: SSI Scientific Software International, Inc.
- Kelloway E.K. (1998). *Using LISREL for Structural Equation Modeling - A Researcher's Guide*. Thousand Oakes: Sage.
- Kline P. (1994). *An easy guide to factor analysis*. London & New York: Routledge.
- Krzanowski W.J. & Marriott F.H.C. (1994). *Multivariate Analysis - Part 2 Classification, covariance structures and repeated measurements*. Kendall's Library of Statistics 2. London: Arnold.
- Marsh H.W. & Hocevar D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, 97, pp. 562 - 582.
- Marton F. & Säljö R. (1976). On qualitative differences in learning -I: Outcomes and processes. *British Journal of Educational Psychology*, 46, pp. 4-11.
- NFWO Contactgroep Academisch Onderwijs (1993). *Van Secundair naar Hoger Onderwijs. Is er een einde aan die jaarlijkse hecatombe?*. Leuven/Apeldoorn: Garant.
- Pask G. (1976). Styles and strategies of learning. *British Journal of Educational Psychology*, 49, pp. 128-148.
- Prins F.J., Busato V.V., Hamaker C. & Visser K.H. (1996). Een bijdrage tot validatie van het (meta)cognitieve deel van de Inventaris Leerstijlen. *Pedagogische Studiën*, 73, pp. 108-122.
- Prins F.J., Busato V.V., Elshout J.J. & Hamaker C. (1998). Een nieuwe bijdrage tot de validatie van het (meta)cognitieve deel van de Inventaris Leerstijlen. *Pedagogische Studiën*, 75, pp. 73-93.
- Rencher A.C. (1995). *Methods of Multivariate analysis*. New York: John Wiley & Sons, inc.
- Ribich F. & Schmeck R.R. (1979). Multivariate Relationships between Measures of Learning Style and Memory. *Journal of Research in Personality*, 13(4), pp. 515-529.
- Santegoets H.M.C. & de Jong F. (1993). Leerstijlen en rendementsverbetering - De relatie tussen leerstijlen en leerprestatie bij eerstejaars MDGO-VP-studenten. *Onderwijs & Gezondheidszorg*, 17, 10, pp. 177-181.
- Schmeck R.R. (1983). Learning styles of college students. In R. Dillon & R.R. Schmeck (Eds.). *Individual differences in cognition*, 1 (pp. 233-279). New York: Academic Press.
- Schouwenburg H.C. (1996). Een onderzoek naar leerstijlen. *Tijdschrift voor Onderwijsresearch*, 21, 2, pp. 151-161.
- SPSS (1998). *SPSS Categories 8.0*. SPSS, inc.
- Taaj J. (1997). *Multivariate Analysis Techniques in Social Science Research - From Problem to Analysis*. Thousand Oakes: Sage Publications.
- 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 van recente onderzoeksgegevens. In H.C. Schouwenburg & J.T. Groenewoud (red.). *Studievaardigheid en leerstijlen* (pp. 51-72). Groningen: Wolters-Noordhoff.
- Wierstra R.F.A. & Beerends E.P.M. (1996). Leeromgevingspercepties en leerstrategieën van eerstejaars studenten sociale wetenschappen. *Tijdschrift voor Onderwijsresearch*, 21, 4, pp. 306 - 322.

De studentenrechtbank: een evaluatieonderzoek naar beoordelingslijsten voor presentatievaardigheden

F.M. Edens, F. Rink en M.J. Smilde

Correspondentie-adres: F.M. Edens, Heymans Instituut, Persoonlijkeids- en Onderwijs
psychologie, Grote Kruisstraat 2/1, 9712 TS Groningen, e-mail: f.m.edens@ppsw.rug.nl

ABSTRACT

This report concerns an investigation into the reliability, validity and efficiency of two rating lists which varied in length. Both lists were used to assess the presentation skills of law school students who had to plead a case or rule in a mock court. Results showed that both lists were internally consistent; inter rater agreement however varied between .17 and .71. Correlations between presentation grades and the big five personality dimensions were not significant, except for Extraversion. Raters found the short list more effective than the long list. For future use the short list was therefore recommended and suggestions for the improvement of inter rater reliability were made.

1. INLEIDING

Dit artikel betreft een onderzoek naar de kwaliteit van twee verschillende beoordelingslijsten voor presentatievaardigheden. Deze lijsten zijn ontwikkeld voor het onderdeel 'eindzitting' van het basisdoctoraalvak de 'Studentenrechtbank' van de faculteit der Rechtsgeleerdheid van de Rijksuniversiteit Groningen (Pater et al., 1998).

In de Studentenrechtbank moet de student in de rol van rechter, pleiter dan wel verdediger aan de hand van een casus een rechtsvraag oplossen in de vorm van een essay. Vervolgens neemt men deel aan een presentatietraining (Smilde, 1997) waarin studenten hun mogelijkheden op het gebied van presenteren kunnen onderzoeken en leren benutten. Tevens leren zij het eigen schriftelijke stuk mondeling overtuigend te presenteren. In een gesimuleerde rechtbankzitting houden de procespartijen hun betoog waarna de 'rechter' uitspraak doet. Alle deelnemers worden vervolgens op hun presentatie beoordeeld door een jury, bestaande uit een lid van de rechterlijke macht, een advocaat en de begeleider van de casus. Zo kent het project dus drie leerelementen: het oplossen van een rechtsvraag, het oefenen van schrijfvaardigheid en het verkrijgen van mondelinge uitdrukkingsvaardigheid.

De presentatietraining van de rechtenfaculteit is een onderdeel van een nieuw opgezet onderwijsprogramma. Het beoordelen van (mondelinge) communicatieve vaardigheden in het hoger onderwijs is een betrekkelijk nieuwe ontwikkeling. Vooral bij de medische wetenschappen (Bögels, 1994; Smit, 1995) en psychologische wetenschappen (Smit & van der Molen, 1996a en b) wordt al een aantal jaren gebruik gemaakt van gestructureerde simulaties om beroepsmatige communicatieve vaardigheden te beoordelen. In de sociale wetenschappen en de gezondheidszorg ligt het accent op één-op-één gesprekken, bijvoorbeeld een adviesgesprek of een anamnese-gesprek. In dergelijke simulaties maakt men over het algemeen gebruik van een casus, bestaande uit een instructie voor de student, een script voor de tegen-speler en een beoordelingslijst (Smit, 1995). Smit en Van der Molen (1996a) wijzen erop dat

een simulatie, in vergelijking met een papier-en-potloodtoets een valide, zij het tijdrovende vorm van vaardigheidstoetsing is. Bij het aanleren van presentatievaardigheden, dus het spreken voor publiek, is er sprake van een andere categorie van professionele communicatieve vaardigheden: het betreft hier niet de dyadische interactie die kenmerkend is voor interpersoonlijke vaardigheden, maar de 'unilaterale intergroepsvaardigheden' (Holsbrink-Engels, 1998).

Er is in de literatuur weinig bekend over presentatietrainingen als curriculumonderdeel in het hoger onderwijs in Nederland. Volgens Van der Meiden (1991) is het aandeel van presentatietrainingen in opleidingen niet erg groot, al wordt er nu meer aandacht aan besteed dan voorheen. Er wordt dan ook weinig gepubliceerd over het toetsen van deze vaardigheden, hoewel her en der ongetwijfeld presentatievaardigheden beoordeeld worden. Bij de Studentenrechtbank werd gekozen voor een aantal casussen, inhoudelijk aansluitend bij de studierichting van de student, op grond waarvan deze in de rol van rechter of één van de procespartijen een juridisch betoog leert schrijven en presenteren. Daarbij werd gekozen voor het ontwikkelen van twee beoordelingslijsten die in lengte en deels in inhoud van elkaar verschilden. Bij het construeren van deze beoordelingslijsten stonden drie onderzoeksvragen centraal: wat is de betrouwbaarheid, de validiteit en de praktische bruikbaarheid van deze lijsten (Smit, 1995)? Uiteindelijk doel van het onderzoek was de meest betrouwbare, valide en bruikbare lijst uit te kiezen. De betrouwbaarheid werd onderzocht door per lijst zowel de interne consistentie (coëfficiënt alpha) als de interbeoordelaarsbetrouwbaarheid te berekenen. De validiteit werd onderzocht door de relatie tussen presentatievaardigheden en persoonlijkheidsfactoren na te gaan.

Er is een groot aantal onderzoeken, met name in de selectiepraktijk, geweest waarin een relatie gelegd werd tussen communicatieve vaardigheden en persoonlijkheidsfactoren van deelnemers aan selectie-onderzoek (Barrick & Mount, 1991; Smit, 1995). De meeste zijn in de selectiepraktijk uitgevoerd. Hofstee (1992) constateert dat vriendelijkheid en ordelijkheid eigenschappen zijn die in een personeelselectie-context vaak als het belangrijkste genoemd worden. Dit geeft een indicatie dat persoonlijkheidseigenschappen een rol kunnen vervullen bij het voorspellen van werkprestaties (Smit, 1995).

Om een indicatie van de praktische bruikbaarheid van beide lijsten te krijgen werd de beoordelaars gevraagd een evaluatieformulier over de beoordelingslijsten in te vullen. Tenslotte is nagegaan in hoeverre de sekse en het beroep van beoordelaars een rol speelden bij de beoordeling, evenals de sekse en studierichting van de beoordeelde.

2. METHODE

2.1. Proefpersonen

De proefpersonen waren tweede- en derdejaars studenten Rechten. Men was afkomstig van de volgende studierichtingen: 41.6% privaatrecht, 22.3% strafrecht, 11.4% ondernemingsrecht, 13.1% publiekrecht, 4.8% fiscaalrecht en 6.6% Europees recht. De totale onderzoeksgroep bedroeg 166 personen. Er namen 62% vrouwelijke proefpersonen en 38% mannelijke proefpersonen deel aan het onderzoek. Deelname aan de eindsitting was verplicht. Het invullen van de onderzoeksinstrumenten de FFPI-vragenlijst en de evaluatie gebeurde wel op vrijwillige basis.

2.2 Beoordelaars

De groep van beoordelaars bestond uit 43 personen, waarvan 25 mannen en 8 vrouwen. De gemiddelde leeftijd bedroeg 44,7 jaar. De beoordelaars waren door de faculteit der Rechtsgeleerdheid van de rijksuniversiteit Groningen benaderd om de taak van jurylid op zich te nemen en namen hetzij op vrijwillige basis, hetzij uit hoofde van hun functie deel aan het onderzoek. Per zitting maakten een rechter, een advocaat en een vakdocent deel uit van de jury.

2.3 Instrumenten

Beoordelingslijsten

De korte lijst bestond uit 9 items en de lange lijst uit 15 items. In een aantal gevallen vonden de items samen een dimensie. De dimensies en items van de lijsten volgen hieronder (tabel 2.1 en 2.2).

Tabel 2.1.: Dimensies en items van de korte lijst

Dimensies	Items
Structuur betoog	- kop - romp - staart
Stemgebruik	- spreekt tempo - verstaanbaarheid - intonatie
Non-verbaal gedrag	- lichaamshouding - gebruik gebaren
Appellerende aspecten	- argumentatie - functionele contacten - gespannenheid

Tabel 2.2.: Dimensies en items van de lange lijst

Dimensies	Items
Non-verbaal	- houding en gebaren - enthousiasme - zelfvertrouwen
Spreekangst	- opvang onverwachte gebeurtenissen - omgaan met spanning
Inhoud	- herkenbaarheid structuur - overgangen - argumentatie - manier van formuleren - duidelijkheid - bondigheid
Stemgebruik	- spreekt tempo - gebruik pauzes - verstaanbaarheid - intonatie - articulatie - stemvolume

De dimensies en items van de lijsten zijn ontwikkeld op basis van de inhoud van de training enerzijds en literatuur (van Eemeren e.a., 1991; Bellack & Hersen, 1988) anderzijds. De items werden gescoord op een vijfpuntsschaal. Per item werd bij de hoogste en laagste score op de vijfpuntsschaal een trefwoord gegeven. Een voorbeeld uit de korte lijst is dat de structuur van het betoog vaag (score 1) of helder kan zijn (score 5). Alleen bij het item 'opvang van onverwachte gebeurtenissen' kon een 6, niet van toepassing, gescoord worden.

Toelichtingsformulieren

Op de toelichtingsformulieren van de lange, respectievelijk de korte lijst werden ten behoeve van de beoordelaars per item drie schaalpunten van de vijfpuntsschaal nader toegelicht. Een voorbeeld hiervan is dat de structuur bij de score één niet duidelijk was, er was geen sprake van een kop, romp, staartindeling, en de presentatie was moeilijk te volgen voor de toehoorders. Bij de score drie was de structuur niet onduidelijk maar ook niet duidelijk. Bij de score vijf was de structuur zeer duidelijk aanwezig in de presentatie, er was een duidelijke kop, romp, staartindeling en het betoog was zeer goed te volgen voor de toehoorders.

Ondersteuningsformulieren

Bij de verschillende beoordelingslijsten zijn twee ondersteuningsformulieren ontwikkeld. Hierop stonden de dimensies aangegeven waarop gescoord kon worden. Op de ondersteuningsformulieren konden de beoordelaars de plus- en minpunten van de presentatie van elke student opschrijven. De genoteerde opmerkingen konden gebruikt worden als geheugensteun bij het bepalen van het cijfer en het geven van feedback.

Evaluatieformulier beoordelaar

Op het evaluatieformulier werd naar de mening van de beoordelaar over beoordelingslijst, toelichtingsformulier en ondersteuningsformulier gevraagd. Het doel van het formulier was het onderzoeken van de praktische bruikbaarheid van de beoordelingslijsten.

FFPI-vragenlijst

De FFPI-vragenlijst (Hendriks, 1997) is gebaseerd op de vijf 'Big Five' persoonlijkheidskenmerken extraversie, mildheid, ordelijkheid, emotionele stabiliteit en autonomie. De lijst bestaat uit 100 korte, gedragsbeschrijvende items. Op een vijfpuntsschaal wordt vervolgens aangegeven in welke mate het betreffende item van toepassing is.

2.4 Procedure

Op het laatste dagdeel van de presentatietraining werd door de studenten de FFPI-vragenlijst ingevuld. Het invullen van deze lijst nam tien minuten in beslag. Voorafgaand aan de zittingen kregen de beoordelaars het toelichtingsformulier, de beoordelingslijst en het ondersteuningsformulier van de lange of de korte lijstversie ter bestudering toegestuurd.

Aan een eindzittingsdagdeel namen in principe vier trio's (rechter en twee procespartijen) deel. Tijdens de eindzitting werd per groep studenten de korte beoordelingslijst dan wel de lange beoordelingslijst gebruikt bij de beoordeling. De helft van de groepen werd beoordeeld met behulp van de lange beoordelingslijst en de andere helft van de groepen werd beoordeeld met de korte beoordelingslijst. De toewijzing van een lijstversie aan de beoordelaarsgroep vond at random plaats, zij het dat rekening gehouden werd met beoordelaars die meer dan één eindzitting beoordeelden: zij gebruikten steeds dezelfde lijstversie.

Bij de eindzittingen was een onderzoeker aanwezig die de beoordelingslijsten samen met de toelichtingsformulieren uitdeelde aan de beoordelaars. Elke student werd door iedere beoordelaar onafhankelijk beoordeeld: per student werden dus in totaal drie beoordelingslijsten ingevuld en drie afzonderlijke cijfers gegeven. De beoordelingslijsten werden ingevuld tijdens de presentaties die elk maximaal 10 minuten duurden.

Na één trio gezien te hebben, trokken de beoordelaars zich terug voor overleg en bepaal-

den gezamenlijk het uiteindelijke eindcijfer. Vervolgens ontving elke student persoonlijke feedback van de voorzitter van de jury. Vanwege een regeling van de rechtenfaculteit kon er op de presentatie geen onvoldoende gegeven worden; het cijfer liep van een zes tot een tien. Na afloop van het juryberaad middelde de onderzoeker de afzonderlijke schaalwaarden op de items per student en noteerde deze op een nieuwe beoordelingslijst. Ook de afzonderlijke cijfers van de beoordelaars werden gemiddeld. Deze 'gemiddelde' beoordeling gaf de onderzoeker aan de student ter informatie mee.

Na afloop van een eindsittingsdagdeel vulden de beoordelaars een evaluatieformulier in waarop naar hun mening gevraagd werd over de gevolgde beoordelingsprocedure.

3. RESULTATEN

3.1 Betrouwbaarheid.

Interne consistentie

De interne consistentie is berekend met coëfficiënt alfa. Hiertoe werd per lijstversie de totaalscore van de drie beoordelaars gemiddeld. De alfa van de lange lijst was .95 (N=78); en van de korte lijst .93 (N=81). Beide lijsten blijken dus een hoge interne consistentie te hebben.

Interbeoordelaarsbetrouwbaarheid

Tabel 3.1: *Correlaties tussen de drie beoordelaars op de lange lijst berekend met de Pearson correlatie coëfficiënt (N = 72).*

Hoofdcategorie	Beoordelaar 1-2	Beoordelaar 1-3	Beoordelaar 2-3
Non-verbaal	$r = .49^{**}$	$r = .66^{**}$	$r = .35^{**}$
Spreekangst	$r = .17$	$r = .48^{**}$	$r = .04$
Inhoud	$r = .29^*$	$r = .39^{**}$	$r = .13$
Stemgebruik	$r = .42^{**}$	$r = .66^{**}$	$r = .45^{**}$
Totaal	$r = .17$	$r = .51^*$	$r = .19$

** correlatie significant op niveau .01 (2-zijdig)

* correlatie significant op niveau .05 (2-zijdig)

Tabel 3.2: *Correlaties tussen de drie beoordelaars op de korte lijst berekend met de Pearson correlatie coëfficiënt (N = 51).*

Hoofdcategorie	Beoordelaar 1-2	Beoordelaar 1-3	Beoordelaar 2-3
Structuur betoog	$r = .39^{**}$	$r = .09$	$r = .34^{**}$
Stemgebruik	$r = .43^{**}$	$r = .34^{**}$	$r = .61^{**}$
Non-verbaal gedrag	$r = .31^*$	$r = .31^*$	$r = .44^{**}$
Appellerende aspecten	$r = .20^*$	$r = .32^*$	$r = .68^{**}$
Totaal	$r = .32^*$	$r = .31^*$	$r = .71^{**}$

** correlatie significant op niveau .01 (2-zijdig)

* correlatie significant op niveau van 0.05 (2-zijdig)

Uit tabel 3.1 blijkt dat de beoordelaars 1 (vakdocenten) en 3 (advocaten) de enigen zijn die significant met elkaar overeenstemmen op het totaal van de lange lijst. Met name 'spreekangst' en 'inhoud' lijken voor de lage correlaties verantwoordelijk te zijn.

Opvallend aan tabel 3.2 is dat de correlaties tussen beoordelaar twee en drie (rechter en advocaat) nu juist hoog zijn; dit in tegenstelling tot de lange lijst. De correlaties van de korte lijst zijn over het algemeen genomen ook iets hoger dan die van de lange lijst.

3.2 Validiteit

Five Factor Personality Inventory (F.F.P.I)

De gegevens van de F.F.P.I. zijn gecorreleerd met het gemiddelde eindcijfer dat de studenten hebben gekregen op hun presentatie. De resultaten staan in tabel 3.3. Alleen de correlatie tussen het cijfer en extraversie is significant.

Tabel 3.3: *Correlaties (Pearson pmc) tussen de vijf persoonlijkheidsfactoren en het cijfer behaald op de eindpresentatie (N = 166).*

	Extraversie	Mildheid	Ordelijkheid	Emotionele stabiliteit	Autonomie
Cijfer	$r = .27^*$	$r = .03$	$r = -.01$	$r = -.07$	$r = .12$

* De correlatie is significant op het niveau van 0.05 (2-zijdig)

Categorieën

Om de interne structuur van de lijsten te onderzoeken werden de items van de diverse categorieën bij elkaar opgeteld, waarna correlaties tussen de categorieën berekend werden (zie tabel 3.4). De lange lijst bevat vier categorieën. De correlaties variëren van $r = .70$ tot $r = .79$ en zijn significant op het niveau van 0.01. De korte lijst bevat ook vier categorieën. Deze correlaties variëren van $r = .67$ tot $r = .80$ en zijn ook significant op het niveau van 0.01. De categorieën hangen onderling dus redelijk tot sterk samen.

Ook het eindcijfer is gecorreleerd met de verschillende categorieën. Op de lange lijst variëren de correlaties van $r = .74$ tot $r = .82$. Op de korte lijst variëren de correlaties van $r = .68$ tot $r = .85$. Naarmate studenten hoger scoren op de categorieën van de lijst krijgen ze dus ook een hoger cijfer.

3.3 Derde variabelen

Om te beoordelen of de cijfers per lijstversie significant van elkaar verschillen, zijn de gemiddelden van beide lijsten getoetst. Het gemiddelde cijfer behaald op de korte lijst is 7.48 ($n=87$); op de lange lijst werd gemiddeld 7.40 ($n=71$) behaald. Bij een $\alpha = .05$ is het verschil niet significant.

Met behulp van een ANOVA in een 2 (rechter, pleiter) x 3 (docent, rechter, advocaat) proefopzet is de invloed van het beroep van de beoordelaar en de rol van de student op het cijfer onderzocht. Noch het beroep, noch de rol bleek van significante invloed op het cijfer. Ook de interactie tussen beide variabelen bleek niet significant.

Om vast te stellen of de studierichting van de student van invloed is op het cijfer zijn de gemiddelden van de studenten per studierichting getoetst. De gemiddelde cijfers staan in tabel 3.5.

Met behulp van de Bonferroni toets is getoetst of de gemiddelden significant van elkaar verschillen. Dat bleek niet het geval te zijn.

Ook werd onderzocht in hoeverre per trainer de gemiddelde cijfers van de groepen van

elkaar verschillen. Er is weer met behulp van Bonferroni getoetst of de verschillende gemiddelden significant van elkaar afwijken. Dat bleek wederom niet het geval te zijn.

Tabel 3.4: De correlaties tussen de hoofdcategorieën en het cijfer L1 = Algemeen Non-verbaal. L2 = Algemeen Spreekangst. L3 = Inhoud. L4 = stemgebruik. K1 = structuur betoog. K2 = stemgebruik. K3 = Non-verbaal gedrag. K4 = appellerende aspecten. L = lange lijst en K = korte lijst.

	Cijfer	L1	L2	L3	L4	K1	K2	K3	K4
Cijfer	1.00					.68	.85	.68	.83
L1	.80	1.00							
L2	.74	.73	1.00						
L3	.82	.73	.79	1.00					
L4	.82	.75	.70	.73	1.00				
K1						1.00			
K2						.67	1.00		
K3						.77	.67	1.00	
K4						.80	.79	.75	1.00

*Alle correlaties zijn significant op 0.01 niveau

Tabel 3.5: Het gemiddelde cijfer en standaarddeviatie per studierichting.

Studierichting	Europees recht	Privaatrecht	Publiekrecht	Strafrecht	Bedrijfsrecht	Totaal
M	7.17	7.37	7.59	7.62	7.39	7.45
Sd	.64	.74	.81	.66	1.41	.84

Verder is een variantie-analyse uitgevoerd op basis van twee factoren, geslacht student en geslacht beoordelaar. Het geslacht van de student bleek geen significant effect op het eindcijfer te hebben ($F .028$, $p .867$). Het geslacht van de beoordelaar had dat wel ($F 3.982$, $p .047$): vrouwelijke beoordelaars bleken een significant hoger cijfer te geven dan mannelijke beoordelaars. Ook het interactieeffect was significant ($F 5.695$, $p .017$). Mannelijke beoordelaars blijken hogere cijfers te geven aan vrouwelijke studenten dan aan mannelijke studenten, terwijl bij vrouwelijke beoordelaars het omgekeerde het geval is: zij geven aan mannen hogere cijfers dan aan vrouwen.

3.4 Praktische bruikbaarheid

Slechts 33 beoordelaars vulden de vragenlijst geheel of gedeeltelijk in. De helft hiervan gebruikte de lange lijst, de overigen gebruikten de korte lijstversie.

Op de vraag of ze genoeg tijd hadden om zich op de zitting voor te bereiden gaf 90% van de lange lijstgebruikers aan voldoende tijd te hebben; 10% had grotendeels voldoende tijd om zich voor te bereiden. Van de korte lijstgebruikers zei 75% voldoende tijd te hebben tegen 10% die zei grotendeels voldoende tijd te hebben om zich voor te bereiden en 15% vulde niets in.

Op de vraag of de toegestuurde instructie voldoende informerend was, antwoordde 85% dat de gestuurde instructie voldoende informerend was, 15% vond de instructie grotendeels voldoende informerend. Van de korte lijstgebruikers vond 75% de instructie voldoende informerend, 10% vond de instructie grotendeels informerend en 15% vulde niets in.

De vraag of de categorieën op de lijst begrijpelijk waren, beantwoordde 65% van de lange lijstgebruikers bevestigend. Terwijl 35% zei de categorieën grotendeels begrijpelijk te vin-

den. Van de korte lijst gebruikers gaf 75% aan de categorieën begrijpelijk te vinden, 25% vond ze grotendeels begrijpelijk.

Op de vraag of beoordelaars per student voldoende tijd hadden om de lijst in te vullen, antwoordde 50% van de lange lijst gebruikers bevestigend. 35% van de respondenten zei grotendeels voldoende tijd te hebben en 15% gaf aan per student onvoldoende tijd te hebben om de lijst in te vullen. Van de korte lijst gebruikers zei 65% voldoende tijd te hebben tegen 25% die grotendeels voldoende tijd had om de lijst in te vullen. Van de respondenten gaf 5% aan onvoldoende tijd te hebben gehad.

De laatste vraag betrof eventuele suggesties en opmerkingen met betrekking tot de gebruikte lijst. De meest voorkomende opmerkingen waren dat er te veel papierwerk was en te weinig tijd om alles in te vullen.

4. DISCUSSIE

De interne consistentie van beide lijsten blijkt hoog te zijn: per lijst vormen de items dus een consistent geheel. De interbeoordelaarsbetrouwbaarheid daarentegen is matig tot redelijk. Waarden boven .80 beschouwt men als voldoende (Van der Sande, 1986). Het probleem is gelegen in het feit dat er veel beoordelaars deelnamen, en dat deze van tevoren niet getraind werden. In het algemeen kan door getrainde observatoren een hoge interbeoordelaarsbetrouwbaarheid gehaald worden (Vu & Barrows, 1994). Daarnaast gaf de dimensie 'spreekangst' (tabel 3.1.) aanleiding tot verwarring; herformulering is waarschijnlijk op zijn plaats. Zo gaf de ene beoordelaar een studente die flauwviel een lage score op omgaan met spanning, terwijl de ander het opstaan en doorgaan na het flauwvallen waardeerde met een hoge score op hetzelfde item.

Met betrekking tot de interne structuur is het volgende op te merken: de vier hoofdcategorieën van beide beoordelingslijsten correleren redelijk sterk (.70-.80) met elkaar. De categorieën hangen dus, als operationalisaties van presentatievaardigheden, onderling samen zonder elkaar geheel te overlappen.

Bij vier van de vijf persoonlijkheidsfactoren is geen significante correlatie met de behaalde resultaten op de presentatie gevonden. De enige significante correlatie is de persoonlijkheidsfactor extraversie en deze correleert laag (.27) met het cijfer behaald op de eindpresentatie. Extraverte mensen hebben dus enig voordeel bij het houden van hun presentatie; extraversie werkt waarschijnlijk faciliterend. De hoogte van de correlatie geeft echter ook aan dat extraverte studenten niet als vanzelfsprekend hoger zullen scoren op de presentatie van hun betoog.

Over de praktische bruikbaarheid kunnen we opmerken dat de beoordelaars over het algemeen tevreden zijn. Bij de voorbereiding scoorde de lange lijst beter op aspecten als voldoende tijd hebben voor de voorbereiding en voldoende voorinformatie krijgen over de lijsten. Tijdens de zittingen blijkt de korte lijst het beste te voldoen met betrekking tot de begrijpelijkheid van categorieën en het hebben van voldoende tijd.

Geen van de derde variabelen bleek van invloed te zijn op de beoordeling, op het geslacht van beoordelaars en studenten na. Vrouwelijke beoordelaars blijken hogere cijfers te geven dan mannelijke beoordelaars. Mannelijke beoordelaars blijken vrouwen bovendien significant hoger te beoordelen dan mannen, terwijl bij de vrouwelijke beoordelaars het omgekeerde het geval is. In hoeverre is hier sprake van een systematische vorm van seksepartijdigheid bij beoordeling? Onderzoek naar sekseverschillen in de beoordeling van wetenschappelijke en artistieke producten wijst niet in deze richting. Zo vonden Swim e.a. (1989) in een meta-analyse dat de gemiddelde verschillen tussen de beoordelingen van mannelijke en vrouwelijke beoordelaars verwaarloosbaar waren. In een overzicht van onderzoeken naar seksepartijdigheid in de beoordeling van creatieve producten concludeert Top (1993) dat er geen algemene neiging is om het werk van vrouwen gunstiger of ongunstiger te beoordelen dan

dat van mannen. Het optreden van seksepartijdigheid blijkt afhankelijk van kenmerken van onder andere het beoordeelde product, de beoordeelde persoon, de beoordelaars en de wijze waarop oordelen worden gegeven. Hierbij blijkt de partijdigheid zowel ten nadele van vrouwen als van mannen op te treden.

Wat voor gevolgen heeft dit alles in onderwijskundig opzicht? Ten eerste verdient het gebruik van de korte lijst de voorkeur bij toekomstige beoordeling van studenten. Wel zou een beoordelaarstraining wenselijk zijn, al geeft een dergelijke beoordelaarstraining waarschijnlijk wel praktische problemen aangezien er veel mensen bijeengebracht moeten worden. Een oplossing hiervoor zou kunnen zijn om een cd-rom of een zelfinstructiepakket te maken waardoor de afzonderlijke beoordelaars hun eigen tijd kunnen indelen.

Als we naar de beoordelingsvorm in het algemeen kijken, dan betreft de simulatie een voor studenten en docenten nieuwe beoordelingsvorm. Diverse auteurs (Smit, 1995; Newble & Jaeger, 1983) maken melding van het feit dat simulaties als vorm van vaardigheidstoetsing over het algemeen positief beoordeeld worden. De studentenrechtbank vormt hierop geen uitzondering: het houden van een presentatie in een gesimuleerde rechtbankzitting voor de ogen en oren van de jury heeft 94% van de deelnemende studenten als leerzaam tot zeer leerzaam ervaren. Daarnaast wijst Smit op het effect dat studenten aangeven zich beter voor simulaties voor te bereiden dan voor een papier-en-potloodtest. Ook voor de studentenrechtbank geldt dat 89% van de studenten zich goed tot zeer goed had voorbereid op de eindzitting. Het lijkt daarbij in dit geval heel goed mogelijk dat een integratie plaatsvindt van kennis en inzichten die opgedaan zijn bij het schrijven van een juridisch betoog en de vaardigheden die nodig zijn om op basis van dit betoog een effectieve presentatie in een rechtszaal te houden.

Tenslotte lijkt het de overweging waard studenten tijdens hun optreden in toga gekleed te laten gaan. Ambtskleding benadrukt immers dat de drager in de eerste plaats als lid van een bepaalde professie gezien wil worden, zodat sekseverschillen mogelijk minder benadrukt worden.

Samenvattend kunnen we dan ook stellen dat we met de ontwikkelde beoordelingsprocedure op de goede weg zijn om recht te doen aan de doelstelling van het project de 'Studentenrechtbank': het verbeteren van de kwaliteit en studeerbaarheid van de studie.

LITERATUUR

- Barrick, M.R., & Mount, M.K. (1991). The Big Five Personality Dimensions and job performance: a meta-analysis, *Personnel Psychology*, 44, pp.1-23.
- Bellack, A.S., & Hersen, M. (1979). *Research and practice in social skills training*, New York: Plenum Press.
- Bögels, S.M. (1994). *Teaching and assessing diagnostic interviewing skills. An application to the mental health field*. Dissertatie, Rijksuniversiteit Limburg, Maastricht.
- Eemeren van F.H. e.a. (1991). *Argumenteren voor juristen*. Groningen: Wolters-Noordhoff
- Hendriks, A.A.J. (1997). *The construction of the Five-Factor Personality Inventory (FFPI)*. Dissertatie, s.n. s.l.
- Hofstee, W.K.B. (1992). *Persoonlijkheidseigenschappen: de caleidoscoop van individuele verschillen*. Rede. Vrije Universiteit, Brussel.
- Holsbrink-Engels, G.A. (1998). *Computer-based Role Playing for Interpersonal Skills Training*. Thesis Universiteit Twente, Enschede: Febodruk.
- Meiden, A. van der (1991). *Over spreken gesproken. Aspecten van moderne retorica*, Groningen: Wolters-Noordhoff.
- Newble, D.I. & Jaeger, K. (1983). The effect of assessments and examinations on the learning of medical students. *Medical Education*, 17, 165-171.
- Pater, J.; Straaten, E.E. van der; Bangma, I.; Wachters, B.E. & Rink, F.A. (1998). *De studentenrechtbank. Het ontwikkelen, geven en evalueren van een beoordelingsstelsel en een training*. Intern verslag bij de afstudeerrichting Persoonlijkheidspsychologie, Rijksuniversiteit Groningen

- Sande, J.P. van de (1986). *Gedragsobservatie. Een inleiding tot systematisch observeren*. Groningen: Wolters Noordhoff.
- Smilde, M. (1997). *Studentenrechtbank. Cursushandleiding*. Faculteit der Rechtsgeleerdheid, Rijksuniversiteit Groningen.
- Smit, G.N. (1995). *De beoordeling van professionele gespreksvaardigheden. Constructie en evaluatie van rollenspel-, video- en schriftelijke toetsen*. Baarn: Nelissen.
- Smit, G.N., & Molen, H.T. van der (1996a). Simulations for the Assessment of Counselling Skills. *Assessment & Evaluation in Higher Education*, (21) 4, pp.335-345.
- Smit, G.N. & Molen, H.T. van der (1996b). Three Methods for the Assessment of Communication Skills. *British Journal of Educational Psychology*, 66, 543-555.
- Swim, J., Borgida, E., Maruyama, G. & Myers, D.G. (1989). Joan McKay versus John McKay: Do gender stereotypes bias evaluations? *Psychological Bulletin*, 105, 409-429.
- Top, T.J. (1993). *Art and Gender. Creative achievement in the visual arts*. Dissertatie. s.l.,s.n.
- Vu, N.V. & Barrows, H.S. (1994). Use of standardized patients in clinical assessments. Recent developments and measurement findings, *Educational Researcher*, 23(3), pp23-30.

NOOT

1. Zo wordt bij de Rechtenfaculteit in Leiden al enige jaren met een beoordelingslijst bij een vergelijkbaar practicum gewerkt. In Nijmegen dienen de partijen bij de Rota Carolina, een verplicht vak in de doctoraalstudie bij privaatrecht, op basis van een casus schriftelijke stukken te produceren en vervolgens hun zaak te bepleiten. Hierbij wordt niet volgens een systematisch beoordelingssysteem gewerkt.

De relatie tussen leerstijl en persoonlijkheid

J. Adema, K.I. van der Zee, H.T. van der Molen

Rijksuniversiteit te Groningen

(Correspondentieadres: Johanna Adema, Sectie Persoonlijkeids- en Onderwijspsychologie, Vakgroep Psychologie, Grote Kruisstraat 2/1, 9712 TS Groningen, e-mail:

J.Adema@PPSW.RuG.nl)

ABSTRACT

The present research examined the relationship between learning style and personality. A sample of 193 second-year students of Psychology ($M = 21.4$ years) completed the Inventaris LeerStijlen (Vermunt, 1992) and the Five Factor Personality Inventory (Hendriks, 1997; Hendriks, Hofstee & De Raad, 1999a; 1999b). As expected, a meaning directed learning style was positively related to Autonomy. In line with the expectations, a reproduction directed learning style was positively correlated with Conscientiousness and negatively with Autonomy. Interestingly, an application directed orientation went along with high scores on Agreeableness. Finally, an undirected approach was negatively related to both Conscientiousness and Autonomy. To conclude, the present data suggest that, although learning styles seem to be modifiable, the possibilities for change are not without limits. Therefore, in developing education programs differences in learning style between students should be considered in order to make instruction suitable for all kinds of students. A possible approach is offering students instruction through several different methods to make sure that they are also instructed in a way that fits their own learning style.

1. DE RELATIE TUSSEN LEERSTIJL EN PERSOONLIJKHEID

De laatste tijd wordt steeds meer nadruk gelegd op de kwaliteit en studeerbaarheid van het onderwijs. In het voortgezet onderwijs resulteerde dat in de opkomst van het Studiehuis en op universiteiten en hogescholen is bijvoorbeeld een verschuiving te zien naar vormen van kleinschaliger onderwijs die een actievere inbreng van studenten vragen en afstandsonderwijs. Het effect van het onderwijs lijkt echter niet alleen van de kwaliteit van het onderwijs af te hangen, maar ook van de leerstijl van de studenten. Als een student een ineffektieve wijze van studeren hanteert, zullen diens studieresultaten minder goed zijn dan die van een student met een vergelijkbaar intelligentieniveau, die op een effectievere manier studeert. Daarnaast functioneren sommige studenten wellicht beter in de ene onderwijsvorm dan in de andere. Leerprocessen en leerstijlen krijgen dan ook steeds meer aandacht in de literatuur (Kolb, 1984; Vermunt, 1992). Het lijkt van belang te weten of leerstijlen gefundeerd zijn in de persoonlijkheid of worden beïnvloed door bijvoorbeeld motivatie of situatiekenmerken zoals onderwijstype. Onderzoeksgegevens hierover kunnen belangrijke implicaties hebben voor de manier waarop kwaliteit en studeerbaarheid van het onderwijs worden bereikt. Wanneer blijkt dat leerstijl sterk is gefundeerd in persoonlijkheid zal in het onderwijs rekening moeten worden gehouden met verschillen tussen lerenden. In dit artikel wordt de relatie tussen leerstijl en persoonlijkheid onderzocht. Voordat deze vraag wordt behandeld wordt eerst beschreven wat onder leerstijl en persoonlijkheid wordt verstaan.

De term 'leerstijl' lijkt in de 70-er jaren te zijn opgekomen als een meer alledaagse term die verwijst naar cognitieve stijl of als vervanging voor de term 'cognitieve stijl' (Riding & Cheema, 1991). Het begrip 'cognitieve stijl' werd door Allport (1937) omschreven als de voor een persoon karakteristieke of gebruikelijke wijze om problemen op te lossen, te denken, waar te nemen en te herinneren. Er zijn verschillende visies op het begrip leerstijl, wat onder andere in de definities ervan naar voren komt. Zo verschillen onderzoekers van mening over de mate waarin leerstijl wordt bepaald door aanleg- of omgevingsfactoren. Volgens Riding en Cheema (1991) wordt op drie manieren tegen (leer)stijlen aangekeken: als een structuur, als een proces en als beide. Wanneer stijl als een structuur wordt opgevat, betekent dit dat deze variabele wordt gezien als stabiel over tijd. Iemands stijl is dan een gegeven in een trainings- of onderwijssetting. Als de stijl van de lerende is vastgesteld kan het leermateriaal daaraan worden aangepast. Wanneer stijl wordt gezien als een proces, wordt deze als veranderbaar en dynamisch beschouwd. Een trainer kan daarbij proberen de stijl te beïnvloeden. Wanneer stijl wordt gezien als een combinatie van structuur en proces, betekent dat dat stijl wordt opgevat als redelijk stabiel, maar wel veranderbaar. Vermunt sluit bij deze laatste invalshoek aan en ziet leerstijl als het resultaat van de interactie tussen persoonsgebonden en omgevingsgebonden invloeden. Deze opvatting wordt gedeeld door de auteurs van dit onderzoek. In dit onderzoek wordt uitgegaan van de indeling in leerstijlen zoals Vermunt (1992) die formuleerde.

Leerstijldefinities verschillen voorts in de reikwijdte en de breedte van het begrip en in de te onderscheiden leerwijzen. Waar Kolb (1984) zich bij zijn metingen van leerstijl op leren in het algemeen richt, beperkt Vermunt zich heel nadrukkelijk tot onderwijssituaties. Vermunt onderzocht de leerstijlen van studenten uit het universitair en het hoger beroepsonderwijs. In tegenstelling tot andere auteurs, die leerstijl definiëren als de neiging om uit eigen beweging van dezelfde leerstrategieën gebruik te maken in verschillende (leer)situaties, vat Vermunt leerstijl in brede betekenis op (Moran, 1991; Schmeck & Geisler-Brenstein, 1989). In de smalle betekenis is leerstijl als het ware de concretisering van een voorkeur voor een bepaalde manier van denken en handelen in de leersituatie (zie Schouwenburg, 1996). Leerstijl kan zo worden opgevat als de expressie van een cognitieve stijl bij het leren of studeren. In de opvatting van Vermunt omvat leerstijl meer dan alleen leerstrategieën. Hij omschrijft leerstijl als een samenhangend geheel van de leer- en regulatie-activiteiten die studenten ontplooiën, hun studie-opvattingen en hun studiemotieven.

1.1 De indeling van Vermunt

Vermunt onderscheidt vier leerstijlen. Deze leerstijlen bestaan uit combinaties van *leerstrategieën*, *regulatiestrategieën*, *mentale leermodellen* oftewel opvattingen over het onderwijs en het leerproces en *leeroriëntaties* oftewel motieven om te studeren. Hieronder volgt een overzicht van de verschillende leerstijlcomponenten en van de vier leerstijlen.

Verwerkingsstrategieën. Uit onderzoek van Vermunt (1992, 1994) naar leerstijlen en -strategieën bleek dat over het algemeen vijf leeractiviteiten in zodanige combinaties worden toegepast, dat drie *leer- of verwerkingsstrategieën* ontstaan: diepteverwerking, stapsgewijze verwerking en concrete verwerking. Studenten die gebruik maken van een *diepteverwerkingsstrategie* proberen structuur aan te brengen in de studiestof, zoeken naar verbanden en hebben een kritische houding ten opzichte van de te bestuderen stof. Studenten die de leerstof *stapsgewijs* verwerken, herhalen de studiestof vaak, leren rijtjes uit hun hoofd en gaan gedetailleerd en stapsgewijs te werk. Een leerstrategie die door *concrete verwerking* wordt gekenmerkt richt zich op het toepassen en concreet maken van te bestuderen leerstof. Een vergelijkbaar onderscheid als dat tussen diepteverwerking en stapsgewijze verwerking is overigens ook bij andere auteurs te vinden: een holistische versus een serialistische (Pask, 1988), een holistische of non-analytische versus een analytische leerstrategie (Beyler & Schmeck, 1992) en een diepe versus een oppervlakkige leerstrategie (Marton & Säljö, 1984). Hierbij zijn de holistische en de non-analytische leerstrategieën te vergelijken met diepteverwerking en de serialistische, de analytische en de oppervlakkige leerstrategieën met stapsgewijze verwerking.

Regulatiestrategieën. Vermunt stelde tevens vast dat er drie *regulatiestrategieën* zijn, namelijk een intern gestuurde, een extern gestuurde en een stuurloze strategie. Bij de *intern gestuurde* strategie sturen studenten zelf hun leerprocessen, bewaken ze hun leerresultaten en bepalen ze vaak ook zelf wat ze leren. Zij raadplegen zelf bijvoorbeeld andere literatuur naast de verplichte literatuur. Studenten met een *extern gestuurde* strategie laten zich bij het sturen van leerprocessen en -inhoud en bij het bewaken van leerresultaten in hoge mate leiden door de sturing vanuit het onderwijs. Van een *stuurloze* strategie is sprake wanneer studenten aan de ene kant hun leerprocessen en -resultaten niet zelf weten te sturen, maar aan de andere kant ook te weinig houvast ondervinden aan de sturing vanuit de onderwijsomgeving.

Mentale leermodellen. Uit de analyses van Vermunt kwamen voorts vijf mentale *leermodellen* naar voren. Sommige studenten zien studeren als het *opnemen* van aangeboden kennis. Andere studenten vatten studeren op als het *opbouwen* van eigen kennis en inzichten. Weer andere studenten vinden dat het onderwijs *stimulerend* moet zijn of studeren en overleggen graag *samen met medestudenten*. Daarnaast zijn er studenten die vinden dat men de kennis die men verwerft moet *leren gebruiken*.

Leeroriëntaties. Tenslotte stelde Vermunt vast dat er vijf *leeroriëntaties* bestaan. Sommige studenten doen hun studie vanuit *persoonlijke interesse*, anderen zijn gericht op het *uittesten van eigen capaciteiten*. Verder zijn er studenten met een *certificaatgerichte* oriëntatie, met een *beroepsgerichte* oriëntatie en studenten die een *ambivalente*, onzekere oriëntatie tegenover hun studie hebben. De studenten uit de laatste categorie vragen zich af of ze de juiste studie hebben gekozen, of ze de studie wel aankunnen en twijfelen of ze wel verder zullen gaan.

Leerstijlen. Uit het onderzoek van Vermunt bleek dat de leer- en regulatie-activiteiten die studenten uitvoeren en de studie-opvattingen en -motieven die zij hebben zodanig met elkaar samenhangen, dat van vier *leerstijlen* kan worden gesproken (zie schema 1): een betekenisgerichte, een reproductiegerichte, een toepassingsgerichte en een ongerichte leerstijl.

Schema 1: Indeling leerstijlen

	<i>betekenisgericht</i>	<i>reproductief</i>	<i>toepassingsgericht</i>	<i>ongericht</i>
<i>leerstrategie</i>	diepteverwerking	stapsgewijze verwerking	concrete verwerking	
<i>regulatiestrategie</i>	zelfgestuurd	extern m.b.t. leerprocessen & -resultaten		stuurloos
<i>leermodel</i>	opbouw kennis & inzichten	opname van kennis	gebruik van kennis	stimulering samenwerking
<i>leeroriëntatie</i>	persoonlijke interesse	certificaatgericht uittesten eigen capaciteiten	beroepsgericht	ambivalent

Uit diverse onderzoeken is gebleken dat deze leerstijlen over het algemeen goed replicerbaar zijn (Busato, 1998; Busato, Prins, Hamaker & Visser, 1995; Schouwenburg, 1996). In replicatie-onderzoek van Roosendaal en Vermunt (1996) bleek de toepassingsgerichte leerstijl niet als afzonderlijke leerstijl naar voren te komen, maar eerder samen te gaan met de betekenisgerichte leerstijl.

1.2 Persoonlijkheid en leerstijl

Zoals gezegd is het huidige onderzoek gericht op de relatie tussen persoonlijkheid en leerstijl. Vermunt (1992) beschouwt leerstrategieën en -stijlen niet als eigenschappen van studenten,

maar als het gevolg van de wisselwerking tussen persoonsgebonden en omgevingsgebonden invloeden. Daarbij zouden persoonsgebonden invloeden als leeftijd, vooropleiding, geslacht, intelligentie en onderwijservaring voor consistentie zorgen in de wijze waarop mensen leren. Echter, buiten geslacht en in zekere mate intelligentie betreffen de persoonsgebonden variabelen die Vermunt noemt variabelen die naarmate de tijd verstrijkt niet zo zeer gelijk blijven: de leeftijd van mensen verandert voortdurend, de hoeveelheid onderwijservaring groeit vaak met de tijd en het opleidingsniveau stijgt eveneens vaak met de tijd. Op de relatie tussen meer stabiele persoonsgebonden invloeden, dat wil zeggen: persoonlijkheidseigenschappen, en leerstijlen van studenten gaat Vermunt niet in.

In het huidige onderzoek wordt bij het beschrijven van de persoonlijkheid uitgegaan van de 'Big Five'. In een groot aantal onderzoeken waarin persoonlijkheidsbeschrijvende termen werden gebruikt, zijn vijf factoren gevonden waarmee het domein van de persoonlijkheid kan worden weergegeven (Digman, 1990; Goldberg, 1993; McCrae & John, 1992). Deze 'Big Five' zijn in verschillende culturen gevonden, waardoor het vermoeden bestaat dat het hier universele eigenschappen betreft (Digman, 1990). Een Nederlands instrument voor het beschrijven van de persoonlijkheid in termen van vijf dimensies is de Five Factor Personality Inventory (FFPI). Deze vragenlijst meet de volgende vijf factoren: (I) Extraversie, (II) Mildheid, (III) Zorgvuldigheid, (IV) Emotionele Stabiliteit en (V) Autonomie. Elke dimensie heeft een positieve en een negatieve pool.

Personen die hoog scoren op Extraversie praten graag, terwijl laagscorers gezelschap vermijden en zich afzonderen. Milde personen houden rekening met andermans gevoelens en belangen, terwijl personen die laag op Mildheid scoren anderen hun wil opleggen. Zorgvuldigheid wordt aan de positieve pool gekenmerkt door eigenschappen als 'werkt graag volgens schema' en aan de negatieve pool door 'doet dingen zonder planning'. Emotioneel Stabiele personen kunnen problemen van zich afzetten, terwijl personen die minder stabiel zijn zichzelf juist problemen aanpraten. Autonomie tenslotte heeft betrekking op intellectuele zelfstandigheid: personen die hoog scoren op deze factor weten verbanden te leggen en willen zich een eigen mening vormen. Personen die laag scoren op Autonomie laten zich leiden door wat anderen doen. In eerder onderzoek is reeds een relatie tussen persoonlijkheid en leerstrategie of -stijl aangetoond. Zo vond Blickle (1996) bijvoorbeeld een sterke relatie tussen diepteverwerking en de persoonlijkheidsfactor 'Openness to Experience' ($r = .49$), vergelijkbaar met de factor Autonomie.

Persoonlijkheid en Leerstijl: hypothesen. De hypothesen in het onderhavige onderzoek zijn samengevat in schema 2. Te zien is dat er vier series verwachtingen zijn te onderscheiden met betrekking tot de relatie tussen leerstijl en persoonlijkheid. De eerste serie verwachtingen heeft betrekking op de *betekenisgerichte* leerstijl. Verwacht werd dat de *betekenisgerichte* leerstijl en alle componenten van deze leerstijl (*diepteverwerking*, *zelfsturing*, een mentaal leermodel waarin leren wordt opgevat als het *opbouwen van kennis* en een leeroriëntatie gebaseerd op *persoonlijke interesse*) positief samenhangen met *Autonomie*. In verschillende onderzoeken is hiervoor empirische ondersteuning gevonden. Zo vonden Slaats, Van der Sanden & Lodewijks (1996) dat leerlingen uit het middelbaar beroepsonderwijs met een *constructieve leerstijl* significant hoger scoorden op *Openheid*, vergelijkbaar met *Autonomie* (V), dan leerlingen met een *reproductieve* of een *passieve* leerstijl. Zoals gezegd vond Blickle (1996) een samenhang tussen de leerstrategie *diepteverwerking* en de Big-Five factor '*Openness to Experience*' (.49). Gadzella en Penland (1995) vonden een relatie tussen enerzijds vaardigheden die samenhangen met *diepteverwerking*, zoals gevolgtrekkingen maken, bewijsmateriaal wegen en argumenten evalueren en anderzijds *creativiteit*, dikwijls geassocieerd met factor V. Wat betreft *zelfsturing* ligt het voor de hand dat gedragingen als het toetsen van leervorderingen door zelf bedachte vragen, het zelf stellen van leerdoelen en zelf toevoegen van extra literatuur aan de opgegeven stof positief samenhangen met zich een eigen mening willen vormen (V^+) en negatief met niets in zijn/haar eentje ondernemen, kenmerkend voor mensen die laag scoren op factor V. Een gerichtheid op het *opbouwen van eigen kennis en inzichten* duidt

op zelfstandigheid en onafhankelijkheid bij het leren, hetgeen een positieve relatie met *Autonomie* aannemelijk maakt. Met betrekking tot *leeroriëntatie* bleek uit onderzoekresultaten van Eison (1982) dat studenten die op leren waren gericht (vergelijk Vermunts schaal persoonlijke interesse) meer fantasierijk en onafhankelijk (V^+) waren dan studenten die meer op het behalen van cijfers waren gericht.

Schema 2: Verwachtingen m.b.t. de relatie tussen leerstijl en persoonlijkheid

	Karakterisering ¹		Karakterisering
<i>betekenisgericht</i>	V^+	<i>reproductief</i>	III^+ & V^-
diepteverwerking	V^+	stapsgewijze verwerking	III^+ & V^-
zelfsturing	V^+	extern gestuurd	III^+ & V^-
opbouw kennis & inzichten	V^+	opname van kennis	III^+ & V^-
persoonlijke interesse	V^+	certificaatgericht uittesten eigen capaciteiten	
<i>toepassingsgericht</i>		<i>ongericht</i>	III^- & V^-
concrete verwerking	V^+	stuurloos	III^- & V^-
gebruik van kennis		stimulerend onderwijs	V^-
beroepsgericht	III^+ & V^+	samenwerking	V^-
		ambivalent	IV^-

1. I = Extraversie, II = Mildheid, III = Zorgvuldigheid, IV = Emotionele Stabiliteit, V = Autonomie, een ⁺ duidt op een positieve relatie, een ⁻ op een negatieve relatie

De tweede serie hypothesen heeft betrekking op de *reproductiegerichte* leerstijl. Op inhoudelijke gronden werd verwacht dat de *reproductiegerichte* leerstijl en de componenten *stapsgewijze verwerking*, *externe sturing* en *leren als het opnemen van kennis* een positief verband vertonen met *Zorgvuldigheid* en een negatief verband met *Autonomie*. Een relatie met *Zorgvuldigheid* is aannemelijk gezien de systematische en grondige werkwijze die karakteristiek is voor de reproductiegerichte leerstijl en die tevens naar voren komt bij het grondig stap voor stap doorwerken van de studiestof, kenmerkend voor *stapsgewijze verwerking*, bij het nauwgezet volgen van aanwijzingen en bij het werken volgens een vast patroon kenmerkend voor *externe sturing*. Een neiging tot *opname van kennis* wordt gekenmerkt door veel inzet, discipline en zorgvuldigheid in werkwijze, hetgeen eveneens duidt op een positieve relatie met *Zorgvuldigheid*. Daarnaast komt uit de schaal *opname van kennis* met betrekking tot het volgen van studieaanwijzingen een afhankelijkheid van docent en onderwijs naar voren die wijst op volgzzaamheid en gebrek aan eigen mening en daarmee op een negatieve relatie met *Autonomie*.

De derde serie hypothesen betref de relatie van de toepassingsgerichte leerstijlelementen met persoonlijkheid. De relatie tussen de toepassingsgerichte hoofdschaal en persoonlijkheid werd exploratief onderzocht. Hoewel uit recent onderzoek van Busato (1998; Busato, Prins, Elshout & Hamaker, 1999) aanwijzingen zijn gevonden voor een relatie met Ontwikkeling (vergelijkbaar met Autonomie) en Mildheid, werden geen hypothesen betreffende deze relaties opgesteld, omdat er inhoudelijk gezien niet een logische verklaring voor genoemde relaties lijkt te zijn. Wel werd op inhoudelijke gronden een positieve correlatie verwacht tussen de *beroepsgerichte leeroriëntatie* en *Zorgvuldigheid* en tussen enerzijds de componenten *concrete verwerking* en *beroepsgerichte oriëntatie* uit de toepassingsgerichte leerstijl en anderzijds *Autonomie*. Het ligt namelijk voor de hand dat iemand met een *beroepsgerichte leeroriëntatie*, gekenmerkt door handelen met het oog op de toekomst, systematisch en volgens plan handelt, typisch voor de Big Five-schaal *Zorgvuldigheid*. Aangezien studenten met deze oriëntatie waarschijnlijk ook meer op de voor hen relevante en nuttige leerstof zijn gericht en

daarom wellicht minder zorgvuldig omgaan met leerstof die hun minder relevant lijkt, wordt tevens verwacht dat de relatie met Zorgvuldigheid slechts matig is. De schalen *concrete verwerking* en *beroepsgerichte leeroriëntatie* vertegenwoordigen daarnaast een zelfstandigheid die samen zou kunnen hangen met *Autonomie*.

De vierde serie hypothesen betrof de *ongerichte* leerstijl. Op inhoudelijke gronden werd verwacht dat de *ongerichte* leerstijl en de component *stuurloze leerstrategie* negatief aan *Zorgvuldigheid* zijn gerelateerd, dat de *ongerichte* leerstijl en de componenten *stuurloze leerstrategie*, *stimulerend onderwijs* en *samen studeren* negatief aan *Autonomie* zijn gerelateerd en dat de component *ambivalente leeroriëntatie* negatief samenhangt met *Emotionele Stabiliteit*. Het gebrek aan systematiek in de werkwijze van studenten met de ongerichte leerstijl maakt een negatief verband tussen deze leerstijl en Zorgvuldigheid aannemelijk. Stuurloos leergedrag wordt gekarakteriseerd door weinig zelfdiscipline, wat derhalve duidt op een negatief verband met Zorgvuldigheid, en door onzelfstandigheid, wat wijst op een negatief verband met Autonomie. De schalen *stimulerend onderwijs* en *samen studeren* vertonen een afhankelijkheid van docent en onderwijs respectievelijk medestudenten, die een negatieve relatie met Autonomie logisch maakt. Uit de ambivalente leeroriëntatie spreekt twijfel en gebrek aan zelfvertrouwen, kenmerkend voor hoog neurotische personen (IV).

3. METHODE

3.1 Proefpersonen en procedure

De proefpersoongroep bestond uit 193 tweedejaarsstudenten psychologie (53 mannen en 140 vrouwen), die in het jaar 1997/1998 deelnamen aan een verplichte tweedejaarscursus gespreksvoering. De leeftijd varieerde van 19 tot 45 jaar ($M = 21.4$, $SD = 3.35$). Het onderzoek werd gepresenteerd als onderdeel van de cursus. De gegevensverwerking was anoniem. Alle cursisten vulden de vragenlijsten in op het eerste dagdeel van de cursus. De proefpersonen werden enkele maanden na deelname geïnformeerd over hun scores op beide vragenlijsten.

3.2 Instrumenten

De Inventaris Leerstijlen (ILS). De ILS (Vermunt & Van Rijswijk, 1990) is een vragenlijst voor het in kaart brengen van de leerstrategieën, regulatiestrategieën, mentale leermodellen en leeroriëntaties van studenten. De ILS bevat 16 schalen, waarvan vier elk in twee subschalen zijn onderverdeeld. In totaal bevat de ILS 120 items: 27 uitspraken over leerstrategieën, 28 over regulatiestrategieën, 40 over mentale leermodellen en 25 over leeroriëntaties. Het aantal items per schaal varieert van vijf tot elf over de schalen en over de subschalen van vier tot negen. Voorbeelden van items zijn 'ik vergelijk mijn visie op een onderwerp uit de cursus met de visie van de auteurs', 'ik leer alles precies zoals het in de studieboeken staat' en 'mijn belangrijkste doel bij deze studie is om mezelf voor te bereiden op de uitoefening van een beroep'. Elk item dient op een 5-puntsschaal te worden beantwoord, waarbij (1) 'helemaal niet van toepassing' betekent, (3) 'neutraal' of 'weet niet' en (5) 'helemaal van toepassing'. Per schaal wordt het totaal van de zelfoordelen berekend door de scores op te tellen. De leerstijlscores kwamen tot stand door de totalen van de bijbehorende subschalen te sommeren. De betrouwbaarheid van de hoofdschalen was hoog met $\alpha = .82$ voor de toepassingsgerichte leerstijl tot $\alpha = .91$ voor de betekenisgerichte leerstijl. De betrouwbaarheden van de subschalen varieerden van .72 (beroepsgerichte oriëntatie) tot .90 (samen studeren) met drie uitschieters naar beneden, namelijk persoonlijke interesse (.52), externe sturing (.65) en certificaatgerichtheid (.64).

De Five Factor Personality Inventory. De FFPI (Hendriks, 1997; Hendriks, Hofstee & De Raad, 1999a; 1999b) meet vijf persoonlijkheidsfactoren, namelijk Extraversie, Mildheid, Zorgvuldigheid, Emotionele Stabiliteit en Autonomie. Elke factor van de FFPI wordt gerepresenteerd door 20 items. Per schaal zijn er tien items die betrekking hebben op de

positieve pool en tien items die betrekking hebben op de negatieve pool van de factor (zie Schema 3). Daarnaast hebben de items dikwijls een secundaire lading op een andere factor.

Schema 3: Voorbeelditems uit de FFPI

Factoren	Voorbeelditems
Extraversie	Babbelt graag (+) Gaat contacten uit de weg (-)
Mildheid	Houdt rekening met andermans gevoelens (+) Legt anderen zijn/haar wil op (-)
Zorgvuldigheid	Werkt graag volgens schema (+) Doet dingen zonder planning (-)
Emotionele Stabiliteit	Kan problemen van zich afzetten (+) Praat zichzelf problemen aan (-)
Autonomie	Weet verbanden te leggen (+) Sluit zich bij de meerderheid aan (-)

Voor het verkrijgen van onafhankelijke factorscores is een orthogonaliseringsprocedure nodig (Hendriks, 1997). De 100 items moeten op een vijfpuntsschaal worden beantwoord, waarbij (1) 'helemaal **niet** van toepassing' betekent en (5) 'helemaal van toepassing'. De tijd die nodig is om de test in te vullen bedraagt 10 tot 15 minuten. De testscores worden omgezet in factorscores, waarbij een score nul op bijvoorbeeld Zorgvuldigheid inhoudt dat de betreffende persoon noch zorgvuldig noch onzorgvuldig is. De betrouwbaarheden van de factorscores, berekend met behulp van gestratificeerde alpha's, zijn .83 voor de Autonomie-schaal en .89 voor de overige schalen (Hendriks, 1997).

4. RESULTATEN

De relaties tussen leerstijl en persoonlijkheid zijn samengevat in *schema 4*. Eerst zijn de ruwe correlaties berekend tussen leerstijl en persoonlijkheid (zie *Tabel 1*, p. 283). Vervolgens werden stapsgewijze regressie-analyses uitgevoerd om te bepalen in hoeverre de leerstijlen terug zijn te voeren op de vijf fundamentele persoonlijkheidsdimensies (zie *Tabel 2*, p. 283). Bij deze analyses werden telkens alle vijf persoonlijkheidsfactoren in de vergelijking opgenomen: Extraversie, Mildheid, Zorgvuldigheid, Emotionele Stabiliteit en Autonomie.

Uit de resultaten blijkt dat vooral Autonomie en Zorgvuldigheid gerelateerd zijn aan leerstijlscores. Zoals verwacht ging een hoge score op de *betekenisgerichte* leerstijl samen met een hoge score op *Autonomie* (zie *Tabel 1*). Dit verband werd volgens verwachting weerspiegeld in een significante samenhang met de componenten diepteverwerking, zelfsturing en persoonlijke interesse. Uit *Tabel 2* blijkt dat de *betekenisgerichte* leerstijl het best werd voorspeld door *Autonomie* en *Extraversie*, waarbij *Autonomie* de meeste variantie verklaarde ($\beta = .37$). Voor *Extraversie* werd een zwak negatief verband gevonden. Een vergelijkbare relatie met *Extraversie* werd gevonden voor drie componenten van de *betekenisgerichte* leerstijl.

Studenten die *reproductiegericht* te werk gingen behaalden hogere scores op *Zorgvuldigheid* en lagere scores op *Autonomie*, hetgeen overeenkwam met de verwachtingen. Uit *Tabel 2* blijkt dat *Zorgvuldigheid* met $\beta = .30$ de meeste variantie verklaarde, gevolgd door *Autonomie* met $\beta = -.28$. De relatie met deze twee persoonlijkheidsfactoren werd bij een aantal componenten van de *reproductieve leerstijl* teruggevonden. In tegenstelling tot de verwachtingen bleek stapsgewijze verwerking niet met lagere scores op *Autonomie* samen te gaan. Onverwacht bleek ook *certificaatgerichtheid* positief met *Zorgvuldigheid* en negatief met *Autonomie* samen te hangen.

De relatie tussen de hoofdschaal *toepassingsgerichtheid* en persoonlijkheid werd explora-

Schema 4: Karakterisering van leerstijlen in termen van persoonlijkheidsfactoren, gevolgd door de resultaten van het onderhavige onderzoek

	Karakterisering ¹	Resultaten
<i>betekenisgericht</i>	V ⁺	I ⁻ & V ⁺
diepteverwerking	V ⁺	I ⁻ & V ⁺
zelfsturing	V ⁺	I ⁻ & V ⁺
opbouw kennis & inzichten	V ⁺	I ⁻ & II ⁺
persoonlijke interesse	V ⁺	IV ⁺ & V ⁺
<i>reproductief</i>	III ⁺ & V ⁻	III ⁺ & V ⁻
stapsgewijze verwerking	III ⁺ & V ⁻	III ⁺
extern gestuurd	III ⁺ & V ⁻	I ⁺ , III ⁺ & V ⁻
opname van kennis	III ⁺ & V ⁻	II ⁺ , III ⁺ & V ⁻
certificaatgericht		III ⁺ & V ⁻
uittesten eigen capaciteiten		
<i>toepassingsgericht</i>		II ⁻
concrete verwerking	V ⁺	V ⁺
gebruik van kennis		II ⁺ & IV ⁺
beroepsgericht	III ⁺ & V ⁺	II ⁺ & III ⁺
<i>ongericht</i>	III ⁻ & V ⁻	III ⁻ & V ⁻
stuurloos	III ⁻ & V ⁻	III ⁻ & IV ⁻
samenwerking	V ⁻	I ⁺ III ⁻ & V ⁻
stimulerend onderwijs	V ⁻	III ⁻
ambivalent	IV ⁻	I ⁻ III ⁻ & IV ⁻

I. I = Extraversie, II = Mildheid, III = Zorgvuldigheid, IV = Emotionele Stabiliteit, V = Autonomie, een + duidt op een positieve relatie, een - op een negatieve relatie

tief onderzocht. De scores op de *toepassingsgerichte* leerstijl hingen samen met drie persoonlijkheidsdimensies, namelijk *Mildheid*, *Autonomie* en *Zorgvuldigheid* (zie tabel 2). Opvallend was dat voor *Autonomie* en *Zorgvuldigheid* geen significante ruwe correlaties met de hoofdschaal werden gevonden (zie tabel 1). Twee van de drie componenten van de toepassingsgerichte leerstijl bleken verantwoordelijk voor de relatie van de hoofdschaal met *Mildheid*, namelijk gebruik van kennis en beroepsgerichtheid. Wat betreft de verwachting dat zowel concrete verwerking als een beroepsgerichte leeroriëntatie met *Autonomie* correleren, werd alleen voor de relatie met concrete verwerking ondersteuning gevonden. Voorts bleek beroepsgerichtheid positief te correleren met *Zorgvuldigheid*.

Tenslotte werd wat betreft de *ongerichte* leerstijl ondersteuning gevonden voor de geformuleerde hypothese: zowel met *Zorgvuldigheid* als met *Autonomie* werd een negatief verband gevonden. Vergeleken met de overige leerstijlen werd van deze leerstijl de meeste variantie verklaard door de persoonlijkheidsfactoren: $R^2 = .22$. Drie persoonlijkheidsfactoren, namelijk *Zorgvuldigheid*, *Autonomie* en *Emotionele Stabiliteit*, bleken de score op de *ongerichte* leerstijl goed te kunnen voorspellen (zie tabel 2). De negatieve samenhang met *Zorgvuldigheid* kwam op het niveau van de componenten niet alleen bij stuurloos leergedrag terug, maar bij alle componenten. Uit de resultaten bleek verder dat samen studeren met een lagere score op *Autonomie* samenging. Wat betreft stuurloos leergedrag en stimulerend onderwijs werd tegen de verwachtingen in geen relatie met *Autonomie* gevonden. Verder bleek niet alleen een ambivalente leeroriëntatie, maar ook een stuurloze leerstrategie negatief gerelateerd aan *Emotionele Stabiliteit*.

Tabel 1: Correlaties tussen persoonlijkheid en de leerstijlen volgens Vermunt.

Schalen (N)	Extraversie	Mildheid	Zorgvuldigheid	Emotionele Stabiliteit	Autonomie
<u>Betekenisgerichte leerstijl</u> (185)	-.19**	.01	.05	.03	.39***
Diepteverwerking (191)	-.18*	-.08	-.01	.01	.44***
Zelfsturing (192)	-.15*	-.02	.11	-.01	.32***
Opbouw van kennis (189)	-.16*	.15*	-.01	.03	.13
Persoonlijke interesse (192)	.00	.04	.10	.15*	.23**
<u>Reproductiegerichte leerstijl</u> (189)	.11	.12	.34***	.05	-.32***
Stapsgewijze verwerking (193)	.07	.04	.31***	.04	-.06
Externe sturing (193)	.16*	.10	.31***	.19**	-.25***
Opname van kennis (189)	.09	.15*	.20**	.03	-.39***
Certificaatgericht (193)	.04	.05	.21**	-.01	-.29***
Testgericht (193)	.06	-.07	.04	-.05	-.12
<u>Toepassingsgerichte leerstijl</u> (188)	-.01	.28***	.12	.08	.12
Concrete verwerking (193)	-.07	.10	-.05	-.07	.30***
Gebruik van kennis (190)	-.04	.32***	.01	.15*	-.01
Beroepsgericht (191)	.04	.22**	.28***	.07	-.02
<u>Ongerichte leerstijl</u> (187)	-.04	.08	-.34***	-.09	-.24**
Stuurloos (193)	-.10	-.07	-.30***	-.22**	-.11
Stimulerend onderwijs (190)	-.10	.06	-.23**	-.08	-.07
Samen studeren (188)	.15*	.14	-.19*	.12	-.33***
Ambivalent (192)	-.20**	.08	-.24**	-.28**	-.13

* p < .05 ** p < .01 *** p < .001

Tabel 2: Resultaten uit stapsgewijze regressie-analyses van de leerstijlen volgens Vermunt op persoonlijkheid.

Leerstijl	Stap	Variabele	R	R ²	R ² change	F	p <	β ₁	β ₂	β ₃
Betekenisgericht	1	V (Auton.)	.39	.15	.15	F(1,183) = 32.37	.0001	.39***	.37***	
	2	I (Extrav.)	.42	.18	.03	F(2,182) = 19.35	.0001		-.16*	
Reproductiegericht	1	III (Zorgv.)	.34	.11	.11	F(1,187) = 24.18	.0001	.34***	.30***	
	2	V (Auton.)	.44	.19	.08	F(2,186) = 22.41	.0001		-.28***	
Toepassingsgericht	1	II (Mildh.)	.28	.08	.08	F(1,186) = 15.47	.0005	.28***	.31***	.31***
	2	V (Auton.)	.33	.11	.03	F(2,185) = 10.99	.0001		.17*	.19**
	3	III (Zorgv.)	.36	.13	.02	F(3,184) = 8.87	.0001			.14*
Ongericht	1	III (Zorgv.)	.34	.11	.11	F(1,185) = 23.73	.0001	-.34***	-.37***	-.38***
	2	V (Auton.)	.44	.19	.08	F(2,184) = 21.73	.0001		-.28***	-.33***
	3	IV (Em.St.)	.47	.22	.03	F(3,183) = 17.58	.0001			-.19**

* p < .05 ** p < .01 *** p < .001

5. DISCUSSIE

De hoofdvraag in het onderhavige onderzoek betrof de relatie tussen leerstijl en persoonlijkheid. De eerste hypothese was dat de *betekenisgerichte* leerstijl en alle componenten van deze leerstijl (diepteverwerking, zelfsturing, een mentaal leermodel waarin leren wordt opgevat als het opbouwen van kennis en een leeroriëntatie gebaseerd op persoonlijke interesse) positief samenhangen met *Autonomie*. Deze hypothese vond een tamelijk sterke ondersteuning. Ook de resultaten uit het onderzoek van Busato (1998; Busato et al, 1999) wijzen op een samenhang tussen de betekenisgerichte leerstijl en de vijfde factor. Daarnaast bleken het grondig en kritisch doornemen van de leerstof, zelf bepalen wat belangrijk is en daartoe soms nog extra literatuur opzoeken ter ondersteuning van het leerproces en een gerichtheid op het opbouwen van kennis samen te gaan met *Introversie*, wat niet onaannemelijk is. Introverte mensen zoeken minder contacten met andere mensen dan extraverte personen. Het kan zijn dat hun 'naar-binnengekeerdheid' maakt dat ze actiever en geconcentreerder met hun leerstof omgaan dan mensen die juist actiever zijn in contacten met andere mensen.

Ten tweede werd verwacht dat studenten met een *reproductiegerichte* leerstijl en studenten die hun stof stapsgewijs verwerken, die zich extern laten sturen en die leren zien als het opnemen van kennis hoger scoren op *Zorgvuldigheid* en lager op *Autonomie*. Ook deze hypothese werd grotendeels ondersteund. Hoewel het aannemelijk lijkt dat de oppervlakkige leerwijze die kenmerkend is voor de stapsgewijze werkwijze negatief samenhangt met de intellectuele zelfstandigheid die karakteristiek is voor *Autonomie*, werd geen relatie tussen deze variabelen gevonden. Dit is wellicht te wijten aan het feit dat stapsgewijze verwerking twee aspecten bevat, namelijk *memoriseren & herhalen* en *analyseren*. Alleen het eerste aspect, dat betrekking heeft op het uit het hoofd leren van feiten, rijtjes en definities, is verwant aan het klakkeloos overnemen van dingen, kenmerkend voor de negatieve pool van *Autonomie*.

De relatie van persoonlijkheid met de *toepassingsgerichte* leerstijl werd exploratief onderzocht. Deze leerstijl bleek net als bij Busato (1998; Busato et al, 1999) positief gerelateerd aan *Mildheid*. Een verklaring voor deze opmerkelijke relatie ligt misschien in het feit dat het onderhavige onderzoek, evenals bij Busato, bij een selecte steekproef is uitgevoerd, namelijk bij psychologiestudenten in de eerste helft van hun studie. Een deel van deze studenten volgt deze studie met als doel mensen te helpen, een nobel streven dat naar alle waarschijnlijkheid samengaat met *Mildheid*. Het is dan ook verklaarbaar dat in deze selecte steekproef *Mildheid* en *toepassingsgerichtheid* met elkaar samenhangen. Uit nader onderzoek zou moeten blijken of dit in een meer diverse steekproef eveneens het geval is. Het zou kunnen zijn dat toepassingsgerichtheid binnen studierichtingen verband houdt met de voor functies in de praktijk wenselijke eigenschappen. Dat zou betekenen dat wanneer hetzelfde onderzoek wordt uitgevoerd bij een groep studenten die voor een functie in de praktijk erg zorgvuldig moeten zijn, zoals studenten Informatica, *toepassingsgerichtheid* samengaat met hogere scores op *Zorgvuldigheid*. Wat betreft de subschalen werd verwacht dat studenten die datgene wat ze leren proberen te toetsen aan de alledaagse werkelijkheid evenals studenten met beroepsgerichte studiemotieven een hogere score behalen op *Autonomie* en dat het laatste type student bovendien ietwat hoger op *Zorgvuldigheid* scoort. Concrete verwerking bleek inderdaad met *Autonomie* en een beroepsgerichte leeroriëntatie inderdaad met *Zorgvuldigheid* samen te hangen, maar de relatie tussen beroepsgerichte motieven en *Autonomie* werd niet gevonden. Wel bleken gebruik van kennis en beroepsgerichtheid met *Mildheid* te correleren.

Wat betreft de vierde serie hypothesen stemden de resultaten niet geheel overeen met de verwachtingen. Zoals verwacht werd een negatief verband gevonden tussen de *ongerichte* leerstijl als geheel en *Zorgvuldigheid* en *Autonomie*. De relatie met *Zorgvuldigheid* werd zelfs bij alle componenten van de ongerichte leerstijl teruggevonden, terwijl de relatie met *Autonomie* op het niveau van de componenten alleen werd bevestigd voor de schaal samen studeren. Tenslotte bleek, wat betreft de hypothesen, niet alleen een ambivalente leeroriëntatie, maar ook een stuurloze leerstrategie samen te gaan met minder *Emotionele Stabiliteit*. Dat

ook stuurloos leergedrag negatief met Emotionele Stabiliteit samenhangt is niet vreemd: het feit dat mensen te weinig houvast aan sturing vanuit het onderwijs ondervinden, maar ook niet in staat zijn zichzelf te sturen impliceert een bepaalde mate van onzekerheid. Verder werd een zwak verband gevonden tussen samen studeren en *Extraversie*, hetgeen niet onlogisch is gezien het feit dat men contacten moet leggen, alvorens samen met anderen te kunnen studeren. Daarnaast bleek een ambivalente leeroriëntatie licht met *Introversie* samen te hangen. Dit stemt overeen met bevindingen van Schouwenburg en Stevens (1996), die in hun Checklist voor Studiemotivatie het begrip sociale teruggetrokkenheid, dat duidelijk betrekking heeft op het studentenleven, definiëren als een factor die remmend werkt op studiemotivatie.

Zoals uit de gegevens van de regressie-analyse bleek, verklaart persoonlijkheid een deel van de variantie in leerstijl van personen. Gezien de resultaten lijkt het verdedigbaar te stellen dat leerstijl inderdaad in de persoonlijkheid is verankerd; toch kunnen persoonlijkheid en leerstijl niet als verschillende termen voor hetzelfde begrip worden opgevat. Leerstijl houdt meer in dan persoonlijkheid alleen. Illustratief hiervoor zijn de percentages onverklaarde variantie van de leerstijlen, variërend van 78% tot 87%. De vraag is nu wat naast persoonlijkheid bepalend is voor de leerstijl die een persoon heeft. Waarschijnlijk speelt de motivatie van studenten hierbij een rol. Zo lijkt een beroepsgerichte leeroriëntatie bijvoorbeeld samen te hangen met extrinsieke motivatie, net zoals persoonlijke interesse duidelijk in verband lijkt te staan met intrinsieke motivatie (vergelijk Schouwenburg & Stevens, 1996). Toekomstig onderzoek zou zich op de relatie tussen motivatie en leerstijl kunnen richten. Uit diverse onderzoeken is reeds gebleken dat aspecten van motivatie samenhangen met het gebruik van leer- en regulatiestrategieën (Pintrich & De Groot, 1990; Pintrich, Roeser & De Groot, 1994; Pintrich & Schrauben, 1992; Wolters & Pintrich, 1998). Een eerste aanwijzing uit Nederlands onderzoek dat leerstijl met motivatie samenhangt is afkomstig van Busato (1998; Busato et al., 1999): hij vond een positief verband van prestatiemotivatie met de betekenis-, reproductie- en toepassingsgerichte leerstijl en een negatieve relatie met de ongerichte leerstijl.

Het kan ook zijn dat situatietekenen zoals bijvoorbeeld onderwijsmethode of toetsingswijze van invloed zijn op leerstijl. In vervolgonderzoek zou aandacht kunnen worden besteed aan de invloed van het type onderwijservaring, bijvoorbeeld traditioneel onderwijs versus zelfinstructieonderwijs, op leerstijl. Vermunt (1992; 1994) geeft aan dat er een samenwerking tussen leer- en docerstrategieën is en dat docenten studenten kunnen uitdagen die leer- en regulatiestrategieën toe te passen die ze nog niet goed beheersen of nog niet uit eigen beweging gebruiken. De huidige invoering van het studiehuis sluit hierbij aan. De kerngedachte achter het studiehuis is dat leerlingen uit het voortgezet onderwijs naast kennis en inzicht een aantal vaardigheden verwerven die voor zelfstandig leren van groot belang zijn². Daarbij fungeert de docent meer als coach voor de leerling dan als een docent die frontaal lesgeeft. Volgens Starren (1996) kan ook de toetsvorm invloed uitoefenen op de manier waarop mensen stof bestuderen. Zo zou bijvoorbeeld een mondeling tentamen stimuleren tot diep-terverwerking. Ook dit is een interessant onderwerp voor vervolgonderzoek.

Zoals aan het begin van dit artikel werd opgemerkt, lijkt het effect van onderwijs niet alleen afhankelijk van de onderwijskwaliteit, maar tevens van de wijze waarop studenten met leren en met de leerstof omgaan. Toekomstig onderzoek zou zich dan ook kunnen richten op de invloed van persoonlijkheid en leerstijl op studieprestaties. Daarnaast is het interessant na te gaan of persoonlijkheid en leerstijl van invloed zijn op het effect van verschillende onderwijsvormen. In de '50-er en '60-er jaren werd een onderwijsmethode geïntroduceerd die scherp contrasteerde met de gebruikelijke frontale manier van lesgeven, namelijk 'zelfontdekkend leren' (Eysenck, 1996). Hoewel uit empirische studies geen verschil in onderwijseffect bleek, bleek later uit onderzoek van Leith (1974) dat sommige typen studenten beter functioneerden in de zelfontdekkende onderwijsvorm en dat andere typen studenten juist meer profijt hadden van frontaal onderwijs. Bij deze twee onderzoeksvragen rijst met het oog op spaarzaamheid de interessante vraag of leerstijl en persoonlijkheid samen studieprestaties of effect van onderwijsvormen beter voorspellen dan persoonlijkheid of leerstijl alleen. De toegevoegde

waarde van het leerstijleninstrument blijkt pas wanneer behalve persoonlijkheid ook leerstijl een significant deel van de variantie in studieprestaties verklaart of wanneer de leerstijl van studenten een mediërende invloed blijkt te hebben op het effect van persoonlijkheid. Daarbij is het tevens interessant het effect van persoonlijkheid en leerstijl op studieprestaties en op effect van type onderwijs, bijvoorbeeld zelfinstructie versus begeleid onderwijs, te koppelen aan het type onderwijservaring dat studenten bezitten, bijvoorbeeld zelfontdekkend leren versus traditioneel onderwijs of ervaring in het studiehuis versus frontaal onderwijs.

Wanneer blijkt dat iemands leerstijl, persoonlijkheid of beide goed kunnen bijdragen aan de voorspelling van prestaties in studie of onderwijsvorm, kan dit voor de psychologische praktijk en voor de inrichting van het onderwijs gevolgen hebben bij de behandeling van geschiktheidsvragen voor studies en voor bepaalde typen onderwijs. Zo kan voor een persoon met een toepassingsgerichte leerstijl een meer theoretische studie, zoals wijsbegeerte, minder geschikt zijn dan een meer praktische opleiding in psychologie of pedagogiek. Wat betreft de inrichting van onderwijs kan de docent kiezen voor verschillende onderwijsvormen, zodat verschillende typen studenten de leerstof op meerdere manieren, waaronder de voor hen meest geschikte wijze, kunnen verwerken. Een ander optie is dat de docent studenten die onderwijsvorm laat kiezen die bij hen past, zodat elke student zijn eigen weg door het onderwijs volgt.

Samengevat lijkt het erop dat, hoewel veranderingen in type onderwijs zowel in het voortgezet onderwijs als in universitaire studies studenten wellicht in de richting van een andere leerstijl sturen, deze invloed niet mag worden overschat. Het onderwijs zou dan ook rekening kunnen houden met verschillen tussen studenten in hun neiging om op een bepaalde manier met leerstof om te gaan, bijvoorbeeld door gebruik te maken van verschillende instructiemethoden.

NOTEN

1. De auteurs willen Prof. Dr. Gerrit Lang, Prof. Dr. Willem K.B. Hofstee en Mario Schönrock bedanken voor hun kritische kanttekeningen en hun suggesties ter verbetering
2. Zie website <http://www.aps.nl/vo/studhuis.htm>

LITERATUUR

- Allport, G.W. (1937). *Personality: A psychological interpretation*. New York: Holt & Co.
- Berge, J.M.F. ten & Siero, F.W. (1994). Factoranalyse. In A. van Knippenberg & F.W. Siero (Eds.), *Multivariate analyse; beknopte inleiding en toepassingen*. Houten: Bohn Stafleu Van Loghum.
- Beyler, J. & Schmeck, R.R. (1992). Assessment of individual differences in preferences for holistic-analytic strategies: Evaluation of some commonly available instruments. *Educational and psychological measurement*, 52, 709-719.
- Blickle, G. (1996). Personality traits, learning strategies, and performance. *European Journal of Personality*, 10, 337-352.
- Busato, V.V. (1998). *Leerstijlen nader geanalyseerd*. Dissertatie. Amsterdam: Universiteit van Amsterdam.
- Busato, V.V., Prins, F.J., Elshout, J.J., & Hamaker, C. (1999). The relation between learning styles, the Big Five personality traits and achievement motivation in higher education. *Personality and Individual Differences*, 26, 129-140.
- Busato, V.V., Prins, F.J., Hamaker, C. & Visser, K.H. (1995). Leerstijlonderzoek gerepliceerd; de samenhang tussen intelligentie en leerstijlen. *Tijdschrift voor Onderwijsresearch*, 20, 332-340.
- Digman, J.M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, 41, 417-440.
- Eison, J.A. (1982). Educational and Personal dimensions of learning- and grade-oriented students. *Psychological Reports*, 51, 867-870.

- Eysenck, H.J. (1996). Personality and the experimental study of education. *European Journal of Personality*, 10, 427-439.
- Gadzella, B.M., & Penland, E. (1995). Is creativity related to scores on critical thinking? *Psychological Reports*, 77, 817-818.
- Goldberg, L.R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48, 26-34.
- Hendriks, A.A.J. (1997). *The construction of the Five-Factor Personality Inventory (FFPI)*. Dissertatie, Rijksuniversiteit Groningen.
- Hendriks, A.A.J., Hofstee, W.K.B. & De Raad, B., (1999a). *Handleiding bij de Five-Factor Personality Inventory (FFPI)*. Lisse: Swets & Zeitlinger.
- Hendriks, A.A.J., Hofstee, W.K.B. & De Raad, B., (1999b). The Five-Factor Personality Inventory (FFPI). *Personality and Individual Differences*, 27, 307-325.
- Kolb, D.A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, N.J.: Prentice Hall.
- Leith, G.O.M. (1974). Individual differences in learning: Interactions of personality and teaching methods. In *Personality and Academic Progress*. London: Association of Educational Psychologists.
- Marton, F., & Säljö, R. (1984). Approaches to learning. In F. Marton, D. Hounsell & N. Entwistle (Eds.). *The experience of learning: Implications for teaching and studying in higher education*. Edinburgh: Scottish Academic Press.
- McCrae, R.R., & John, O.P. (1992). An introduction to the Five-Factor Model and its applications. *Journal of Personality*, 60, 175-215.
- Moran, A. (1991). What can learning styles research learn from cognitive psychology? *Educational Psychology*, 11, 239 - 245.
- Nunnally, J.C. (1978). *Psychometric theory*. New York: Mc Graw Hill.
- Pask, G. (1988). Learning strategies, teaching strategies, and conceptual or learning styles. In R.R. Schmeck (Ed.), *Learning strategies and learning styles*. New York: Plenum Press.
- Pintrich, P.R. & Groot, E.V. de (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.
- Pintrich, P.R., Roesser, R.W. & Groot, E.A.M. de (1994). Classroom and individual differences in early adolescents' motivation and self-regulated learning. *Journal of Early Adolescence*, 14, 139-161.
- Pintrich, P.R. & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom tasks. In D.H. Schunk & J. Meece (Eds.). *Student Perceptions in the Classroom: Causes and Consequences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Riding, R. & Cheema, I. (1991). Cognitive Styles - an overview and integration. *Educational Psychology*, 11, 193 - 215.
- Roosendaal, A. & Vermunt, J.D.H.M. (1996). Leerstijlen en zelfstandig leren in het voorportaal van het studiehuis. *Tijdschrift voor Onderwijsresearch*, 21, 336-347.
- Schmeck, R.R. & Geisler-Brenstein, E., (1989). Individual differences that affect the way students approach learning. *Learning and individual differences*, 1, 85-124.
- Schouwenburg, H.C. (1996). Een onderzoek naar leerstijlen. *Tijdschrift voor Onderwijsresearch*, 21, 151-161.
- Schouwenburg, H.C., & Stevens, I.L. (1996). *De Checklist Studiemotivatie CSM'96*. Groningen: Rijksuniversiteit Groningen.
- Slaats, A. Van der Sanden, J.M.M., & Lodewijks, J.G.L.C. (1996). Leerstijlen en persoonlijkheidskenmerken in het Middelbaar Beroepsonderwijs. In J.G.L.C. Lodewijks (Ed.). *Onderwijsonderzoek in Nederland en Vlaanderen 1996*. Tilburg: Sectie Onderwijs- en Opleidingspsychologie.
- Starren, J.C.M.G.M. (1996). De toets als hefboom voor gewenst leergedrag. *De Psycholoog*, 31, 294-295.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs. Naar procesgerichte instructie*. Dissertatie. Amsterdam/Lisse: Swets & Zeitlinger.
- Vermunt, J. (1994). Leerstijlen en leerstrategieën van studenten: Recente onderzoeksgegevens. *VELON tijdschrift*, 15, 8-15.
- Vermunt, J.D.H.M. & Van Rijswijk, F.A.W.M. (1990). *Inventaris Leerstijlen*. Heerlen/Tilburg: Open Universiteit / Katholieke Universiteit Brabant.
- Wolters, C.A. & Pintrich, P.R. (1998). Contextual differences in student motivation and self-regulated learning in mathematics, English, and social studies classrooms. *Instructional Science*, 26, 27-47.

Studievoortgang in het wetenschappelijk onderwijs

M. van den Berg en A. Hofman

M.N. van den Berg, Rotterdams Instituut voor Sociaal-Wetenschappelijk Beleidsonderzoek (RISBO), Erasmus Universiteit Rotterdam.

W.H.A. Hofman, RISBO, Erasmus Universiteit Rotterdam, e-mail: hofman@risbo.eur.nl

ABSTRACT

The impact of student characteristics, curriculum and context factors on careers in higher education is subject of this study. A theoretical model including diverse operationalisations of these factors was evaluated by means of linear structural modelling.

Students under the "prestatiebeurs"-regime achieve better results compared to "tempobeurs"-students. Women gain more study points than men. Students in technical education gain less study points than others, although they spend more time and have gained higher exam scores in high school. Labour activities have negative effects on study progress in general, while time spent on study exerts a stronger impact for men in comparison to women.

1. INLEIDING

Sinds de invoering van de Wet Tweefasenstructuur (1981) heeft het Nederlandse Hoger Onderwijsstelsel een aantal ingrijpende wijzigingen ondergaan. Voorbeelden hiervan zijn de Wet Heroriëntering Studiefinanciering (1991), waardoor de studiefinancieringsduur van de meeste studies in het Hoger Onderwijs bekort werd en de Wet op de Studievoortgangscontrole (1993), die beter bekend staat onder de naam 'tempobeursnorm'. In 1996 is deze wet nog verder aangescherpt door het invoeren van de 'prestatiebeursnorm'.

De genoemde wijzigingen hadden alle de doelstelling om de gemiddelde studieduur van studenten in het Wetenschappelijk Onderwijs te verlagen. Onderzoek toont aan, dat de gemiddelde studieduur sinds de invoering van de Wet Tweefasenstructuur is gedaald van ruim zeven jaar naar ruim vijf en half jaar (Prins, 1997). Waarschijnlijk spelen de wijzigingen in het Hoger Onderwijsstelsel een rol bij het verklaren van de studievoortgang van studenten. Deze wijzigingen, maatregelen van buitenaf, kunnen worden aangeduid met de term 'omgevingsfactoren'.

Daarnaast is het waarschijnlijk dat tal van andere factoren hierop van invloed zijn, zoals de motivatie en leeftijd van de student en zijn geschiktheid voor de opleiding; deze eigenschappen worden samengevat met de term 'studentfactoren'. Een derde groep factoren betreft zaken zoals de curriculumorganisatie van een opleiding, de moeilijkheid van de opleiding en het aantal tentamenmogelijkheden dat jaarlijks aangeboden wordt; deze vormen gezamenlijk de 'opleidingsfactoren'.

Verwacht wordt dat deze factoren, ieder afzonderlijk en in interactie met elkaar, invloed uitoefenen op de studievoortgang van studenten in het Wetenschappelijk Onderwijs.

De probleemstelling van dit onderzoek luidt dan ook: 'In welke mate kan de studievoortgang van studenten in het Wetenschappelijk Onderwijs verklaard worden uit studentfactoren, opleidingsfactoren, omgevingsfactoren en hun onderlinge interacties?'

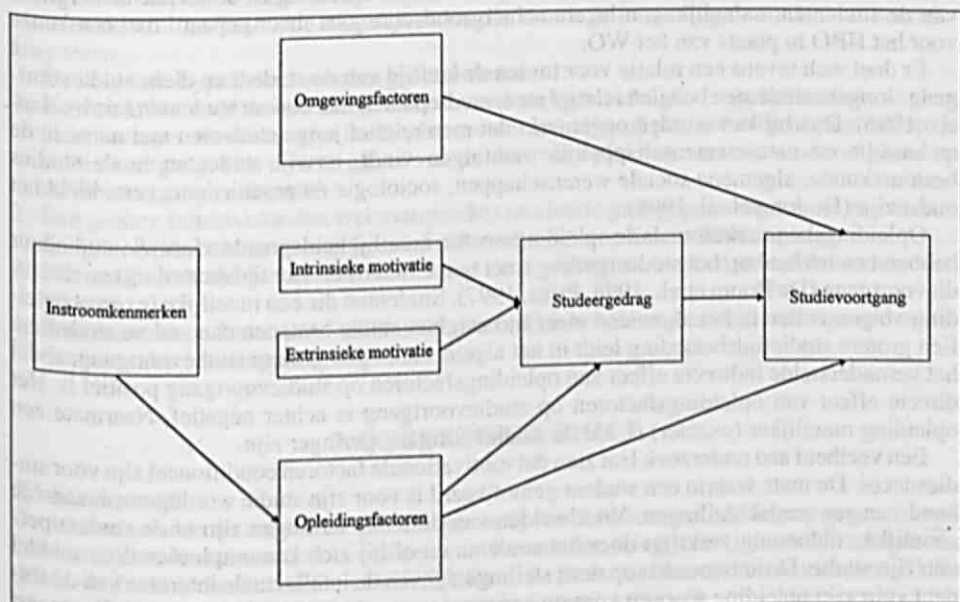
2. THEORETISCH KADER EN HYPOTHESEN

In deze studie zal gebruik gemaakt worden van de interactionele benadering van studievoortgang. De theoretische basis voor interactionele modellen, dus modellen waarbij het verschijnsel studievoortgang (of studiestaking) verklaard wordt vanuit studentfactoren, opleidingsfactoren en omgevingsfactoren en hun interacties, is gelegd door Spady (1970). Modellen van voor zijn tijd legden veelal eenzijdig de nadruk op studentfactoren. Enkele navolgers van Spady zijn Tinto (1987), Bean en Metzner (1985), Pascarella en Terenzini (1983) en Webb (1990).

In Nederland heeft Bijleveld (1993) een interactioneel model ter verklaring van studiestaking ontworpen. Na uitgebreide literatuurstudie komt hij tot de conclusie dat onderzoek naar studiestaking en studievoortgang meestal beperkt wordt tot één opleiding. Bijleveld richt zich evenwel op het verklaren van verschillen in studiestaking tussen opleidingen, waarbij vooral het effect van opleidingsfactoren belicht wordt. Jansen (1996) heeft in haar dissertatie een model ter verklaring van studievoortgang ontworpen en getoetst, waarbij de invloed van de organisatie van het curriculum het uitgangspunt vormt. Daarnaast speelt een aantal studentkenmerken een rol.

Prins (1997) stelt vast dat het model van Bijleveld zeer bruikbaar is voor de verklaring van studiestaking. Hij is echter van mening dat Bijleveld een aantal cruciale studentkenmerken buiten beschouwing heeft gelaten. Prins maakt onderscheid in drie groepen factoren die samenhangen met studieuitval: stelsel factoren, studentfactoren en opleidingsfactoren. De stelsel factoren worden niet in het theoretisch model opgenomen omdat er binnen zijn onderzoeksgroep geen verschillende stelselomstandigheden (studiefinancieringsregimes) optreden. Prins heeft uiteindelijk een empirisch verklaringsmodel opgesteld, waarin alleen nog de studentkenmerken opgenomen zijn. Toetsing met behulp van LISREL (Jöreskog en Sörbom, 1993) leverde een verklaarde variantie van studieuitval van ruim 26%.

In deze studie wordt op basis van interactionele theorieën een model opgesteld ter verklaring van studievoortgang (figuur 1), waarin zowel opleidingsfactoren als omgevingsfactoren



Figuur 1: Theoretisch model ter verklaring van studievoortgang.

en studentfactoren zijn opgenomen. Deze laatste groep factoren wordt onderverdeeld in de instroomkenmerken van de student, het studeergedrag van de student en de mate waarin de student intrinsiek en extrinsiek gemotiveerd is voor zijn opleiding.

Studeergedrag, omgevingsfactoren en opleidingsfactoren worden geacht directe effecten uit te oefenen op de studievoortgang in het Wetenschappelijk Onderwijs. De invloed van de opleidingen op de studievoortgang verloopt naar verwachting deels via het studeergedrag. Ook de motivationele factoren oefenen hun invloed naar verwachting indirect uit via het studeergedrag. De instroomkenmerken van de studenten, tenslotte, zullen invloed kunnen uitoefenen op de studievoortgang via meervoudige indirecte relaties. Er worden hier slechts indirecte relaties verondersteld omdat we vooral geïnteresseerd zijn in de mechanismen die een verband tussen instroom en studievoortgang kunnen verklaren.

In het conceptueel model zijn die relaties weergegeven waarvan op basis van theoretische overwegingen en empirische evidentie significante verbanden mogen worden verondersteld. Daarbij hebben we ons met betrekking tot de theorie internationaal georiënteerd (vgl. de eerder genoemde literatuur) terwijl als het gaat om de empirische verbanden de Nederlandse situatie als uitgangspunt is genomen. Dit laatste is van belang omdat met name de Amerikaanse setting, en daarmee naar verwachting de daar aangetroffen modelverbanden, afwijkt van de Nederlandse situatie. We gaan nu nader in op de gepostuleerde verbanden.

Het veronderstelde rechtstreekse verband tussen de instroomkenmerken en opleidingsfactoren is gebaseerd op het uitgangspunt dat studenten in het algemeen voor een opleiding in het Wetenschappelijk Onderwijs zullen kiezen die aansluit bij de samenstelling van hun eind-examenpakket en behaalde eindexamenresultaten bij de vooropleiding. Zo zullen studenten met hogere cijfers voor de exacte vakken naar verwachting vaker kiezen voor een exacte opleiding dan studenten met lagere cijfers. De Jong et. al. (1998) concluderen: "*studenten in technische richtingen hebben in het algemeen de hoogste schooladviezen, de hoogste gemiddelde eindexamencijfers, meer bètavakken. Bovendien zijn ze minder vaak gedoubleerd....*".

De verschillen in sociale achtergrond van de studenten tussen uiteenlopende studierichtingen in het Wetenschappelijk Onderwijs zijn veelal klein. Wel wordt een relatie geconstateerd in de onderzoeksliteratuur tussen het niveau van de opleiding en de sociale achtergrond van de studenten, namelijk: een lagere achtergrondscore gaat meer gepaard met een keuze voor het HBO in plaats van het WO.

Er doet zich tevens een relatie voor tussen de leeftijd van de student en diens studievoortgang. Jongere studenten behalen relatief meer studiepunten dan oudere studenten (zie bv. Jansen 1996). Daarbij kan worden opgemerkt dat men relatief jonge studenten met name in de technische en natuurwetenschappelijke richtingen vindt, terwijl studenten in de studies bestuurskunde, algemene sociale wetenschappen, sociologie en geschiedenis gemiddeld het oudst zijn (De Jong et. al, 1998).

Opleidingskenmerken zoals de opleidingssector, moeilijkheidsgraad en formele studieduur hebben een invloed op het studeergedrag (met name in termen van tijdsbesteding) en de studievoortgang (De Boom et.al., 1998; Prins, 1997). Studenten die een moeilijke (exacte) opleiding volgen zullen in het algemeen meer tijd aan hun studie besteden dan andere studenten. Een grotere studietijdsbesteding leidt in het algemeen tot een grotere studievoortgang, zodat het veronderstelde indirecte effect van opleidingsfactoren op studievoortgang positief is. Het directe effect van opleidingsfactoren op studievoortgang is echter negatief. Naarmate een opleiding moeilijker (exacter) is, zal de studievoortgang geringer zijn.

Een veelheid aan onderzoek laat zien dat motivationele factoren conditioneel zijn voor studiesucces. De mate waarin een student gemotiveerd is voor zijn studie wordt gemeten aan de hand van een aantal stellingen. Voorbeelden van dit soort stellingen zijn of de student persoonlijke voldoening verkrijgt door het studeren en of hij zich kan ontplooiën door middel van zijn studie. De antwoorden op deze stellingen geven de intellectuele interesse van de student voor zijn opleiding weer en vormen een maat voor zijn intrinsieke motivatie. Daarnaast worden vragen gesteld over de materiële interesse van de student voor zijn opleiding. Deze

vragen vormen een maat voor de extrinsieke motivatie. Voorbeelden van dit soort vragen zijn of de student een hogere positie denkt te bereiken door zijn opleiding of dat hij hierdoor een hoog inkomen denkt te verwerven (Verbeek en De Jong, 1995; Verrijt, 1996). Studenten die sterk (intrinsiek) gemotiveerd zijn besteden in het algemeen meer tijd aan hun studie en behalen een grotere studievoortgang dan minder sterk gemotiveerde studenten. Het belang van studiemotivatie inzake studievoortgang kan worden geïllustreerd aan de hand van het onderzoek van Van Dyck et.al. (1985). Dit toont aan dat 33% van een groep HBO-studenten die hun propeuse niet behaald hebben, dit wijt aan motivationele factoren.

De omgeving waarin de student zijn studie tracht te voltooien oefent invloed uit op de studievoortgang. Uitgangspunt is hier de toetsing van invloeden op de studievoortgang van een tweetal factoren die in de omgeving van de student worden gesitueerd, namelijk: het studiefinancieringsregime waaraan hij of zij onderhevig is en het werkgedrag en de werkhoudingen van de student. Het studiefinancieringsregime valt uiteen in tempo- en prestatiebeursregime die met name van elkaar verschillen in de duur van de beurs en de te behalen studievoortgang.¹

Een belangrijke component van de omgevingsfactoren wordt gevormd door het werkgedrag van studenten. Het verrichten van (betaald) werk naast de studie heeft na wijzigingen in de studiefinanciering in 1992 een drastische groei doorgemaakt.

Tevens kan worden gesteld dat de huidige krappe en steeds flexibeler wordende arbeidsmarkt situatie uitgebreide mogelijkheden biedt tot werken naast de studie (De Boom, et.al., 1998). Het aantal gewerkte uren per week oefent naar verwachting een negatieve invloed uit op de studievoortgang. Zo blijkt uit onderzoek van Verbeek en De Jong (1995) dat werkende studenten die meer dan twaalf uur per week bijverdienen een lager studietempo hebben dan studenten die minder werken.

Door Bean en Metzner (1985) wordt het aantal uren dat door de student wordt besteed aan betaald werk als omgevingsvariabele in het verklaringsmodel geplaatst. Wij sluiten daarbij aan gezien het gegeven dat het al dan niet werken in belangrijke mate omgevingsbepaald is, namelijk afhankelijk van de mogelijkheden die de (regionale en lokale) arbeidsmarkt biedt, met name in bepaalde sectoren, de vrijheid in tijdsbesteding die de opleidingen in deze toestaan en de overige financiële mogelijkheden die vanuit de omgeving de student ter beschikking staan.

De volgende hypothesen worden geformuleerd:

1. Naarmate het voor de opleiding vereiste wiskundig inzicht groter is, besteedt de student meer uren per week aan zijn studie, hetgeen een (indirect) positief effect uitoefent op de studievoortgang. Het directe effect van de door de opleiding vereiste wiskundigheid op studievoortgang is negatief.
2. Een grotere (intrinsieke dan wel extrinsieke) motivatie gaat gepaard met meer aan de studie bestede uren, hetgeen een positief effect uitoefent op de studievoortgang.
3. Naarmate meer tijd aan betaald werk naast de studie wordt besteed neemt de studievoortgang af.
4. Studenten die onder het prestatiebeursregime vallen behalen meer studievoortgang dan studenten onder het tempobeursregime.

3. ONDERZOEKSOPZET

3.1 Steekproef

De data waarvan in dit onderzoek gebruik gemaakt wordt hebben betrekking op gegevens van 2401 voltijdstudenten van de Universiteit van Amsterdam (UvA), de Universiteit Maastricht (UM), de Technische Universiteit Delft (TUD) en de Erasmus Universiteit Rotterdam (EUR) over het collegejaar 1996-1997. De gegevens zijn in een telefonisch survey verzameld in het kader van het onderzoek "Studeren en werken in het Wetenschappelijk Onderwijs", dat door

het Rotterdams Instituut voor Sociaal-Wetenschappelijk Beleidsonderzoek (RISBO) wordt uitgevoerd in opdracht van de directie Wetenschappelijk Onderwijs van het Ministerie van Onderwijs, Cultuur en Wetenschappen (De Boom et al., 1998).

De analyses zijn gebaseerd op een gestratificeerde, niet-proportionele steekproef van studenten in het Wetenschappelijk Onderwijs. De onderscheiden strata zijn de universitaire instelling, het cohort en de sector waarin de student studeert en het geslacht van de student.² De keuze voor een niet-proportionele aselechte steekproef heeft tot gevolg dat er geen sprake is van een evenredige vertegenwoordiging van populatiekenmerken in de steekproef. Om uitspraken te kunnen doen over de populatieparameters van universitaire instelling, cohort, sector en geslacht is een weging toegepast op de steekproef, zodanig dat de steekproefgrootte hetzelfde is gebleven.³ (De Boom et al., 1998).

Het responspercentage voor de steekproef als geheel bedraagt 93%. Per cel (universitaire instelling * cohort * sector) variëren deze percentages van 80% tot 100%. De representativiteit van de steekproef is voor een aantal kenmerken nagegaan door vergelijking met de populatie voltijdstudenten van de vier universiteiten en de voltijdstudenten in Nederland (tabel 1). Hieruit blijkt dat het percentage mannen in de steekproef iets lager is dan op grond van de populatiegegevens en landelijke gegevens werd verwacht. Het percentage studenten met de Nederlandse nationaliteit wijkt niet af van de populatiegegevens en de landelijke gegevens. Het percentage leners bij de Informatie Beheer Groep (IBG) komt overeen met de landelijke gegevens (De Vos en Fontein, 1997).

Tabel 1: Algemene kenmerken

	ongewogen steekproef	populatie	Nederland
% mannen	51	58	54
% nl. nationaliteit	96	96	95
% leners IBG	18	niet bekend	17

3.2 Variabelen

In de volgende subparagrafen worden de variabelen besproken en zullen enkele beschrijvende statistieken gepresenteerd worden voor de totale steekproef, voor mannen en vrouwen, voor de beide studiefinancieringsregimes en voor de zes onderscheiden sectoren.

Studentfactoren

De studentfactoren worden opgesplitst in vier 'blokken', te weten de instroomkenmerken, het studeergedrag en de intrinsieke en extrinsieke studiemotivatie van de student.

De instroomkenmerken die in het theoretisch model opgenomen worden zijn het geslacht en de leeftijd van de student. Hierbij wordt de leeftijd bij instroom in de opleiding gehanteerd, welke kan worden opgevat als uitdrukking van een al dan niet vertraagde (of onderbroken) onderwijs carrière. Voor de studenten met een VWO-vooropleiding (82,6%, N = 1983) wordt het gemiddelde eindexamencijfer voor de exacte vakken (natuurkunde, scheikunde, biologie, wiskunde a en b) en het aantal exacte vakken in het eindexamenpakket opgenomen (vergeleijk Bijleveld, 1993).

De studiemotivatie wordt gemeten aan de hand van acht stellingen (zie bijlage I). Vier stellingen hebben betrekking op de intrinsieke motivatie van de student, dit is de intellectuele interesse van de student voor zijn opleiding. De overige stellingen vormen een maat voor de extrinsieke motivatie, hiermee wordt de materiële interesse van de student voor zijn opleiding weergegeven (Verbeek en De Jong, 1995; Verrijt, 1996). Per stelling hadden de studenten de keuze uit de volgende antwoordmogelijkheden: 1. helemaal mee oneens, 2. mee oneens, 3. mee eens en 4. helemaal mee eens. Alle vragen zijn positief ge(her)formuleerd, zo dat een hogere score een sterkere motivatie impliceert.

Het studeergedrag wordt geoperationaliseerd door middel van het gemiddeld aantal uren dat de student in het collegejaar 1996-1997 per week besteedde aan passieve onderwijsvormen zoals hoorcolleges, actieve onderwijsvormen zoals practica en werkcolleges, en zelfstudie.

Tabel 2: Studentfactoren

	N	N exact	gem. exact	N	leef-tijd begin studie	intrinsic	extrinsic	passief	actief	zelfstudie
totaal	1983	2.5	6.6	2401	19.1	12.6	12.0	8.4	7.0	14.2
man	1192	2.7	6.8	1400	19.0	12.5	12.2	9.0	7.0	13.6
vrouw	791	2.1	6.4	1001	19.2	12.7	11.7	7.6	6.9	15.0
tempo-beurs	1471	2.5	6.6	1854	19.2	12.6	12.0	7.7	6.8	14.4
prestatie-beurs	512	2.4	6.7	547	18.8	12.5	12.0	10.6	7.6	13.5
n & t *	556	3.5	7.2	606	18.7	12.5	12.2	11.9	8.9	12.6
gezondh.	268	3.7	6.6	318	19.1	12.9	12.1	9.5	9.0	13.7
economie	464	2.1	6.7	552	18.9	12.5	12.4	7.1	6.7	15.4
recht	268	1.3	6.1	327	19.3	12.4	12.0	6.2	6.1	15.8
g & m	239	1.7	6.0	329	19.6	12.6	11.4	6.4	5.1	12.9
t & c	188	1.3	6.1	269	19.8	12.7	11.0	6.7	4.5	15.6

* afkortingen zijn verklaard in noot 2.

Uit het gegeven dat in tabel 2 onderscheid gemaakt wordt tussen mannen en vrouwen, tempobeurs- en prestatiebeursstudenten en de zes onderscheiden sectoren, valt op te maken dat op voorhand verschillen verwacht worden tussen deze groepen. Uit vergelijking van de gemiddelde waarden voor mannen en vrouwen (t-toets) blijkt dat de eerste groep significant hogere cijfers behaald heeft voor de exacte vakken. Mannen kiezen ook significant meer exacte vakken in het eindexamenpakket dan vrouwen. Vrouwen hebben een significant grotere intrinsieke motivatie en een significant kleinere extrinsieke motivatie dan mannen. Gemiddeld wordt door studenten 29.6 uur per week aan de studie besteed. Het volgen van passieve onderwijsvormen maakt daarvan 28% uit. Actieve onderwijsvormen en zelfstudie nemen 24% respectievelijk 48% van de studietijd voor hun rekening. Mannen besteden significant meer tijd aan passief onderwijs en vrouwen significant meer tijd aan zelfstudie.

Prestatiebeursstudenten zijn op het moment van instroom jonger dan tempobeursstudenten. Zij besteden significant meer tijd aan passieve en actieve onderwijsvormen, maar minder tijd aan zelfstudie.

Een variantie-analyse toont significante verschillen tussen de onderscheiden sectoren. Opvallend zijn de uiteenlopende scores tussen sectoren met betrekking tot het gemiddelde eindexamencijfer voor de exacte vakken. De studenten uit de sector natuur & techniek behalen voor deze vakken hogere cijfers dan de studenten uit de andere sectoren. Het gemiddelde cijfer verschilt niet tussen de sectoren gezondheid en economie. Deze studenten hebben een hoger gemiddelde dan studenten uit de sectoren recht, gedrag & maatschappij en taal & cultuur, die onderling niet verschillen.

Studenten uit de sector natuur & techniek besteden meer tijd aan passief onderwijs dan studenten uit de sector gezondheid. Natuur & techniek studenten en gezondheid-studenten besteden meer tijd aan passief en actief onderwijs dan studenten uit de overige sectoren. Studenten uit de sector taal & cultuur besteden minder tijd aan actieve onderwijsvormen dan de andere studenten.

Opleidingsfactoren

De opleidingsfactoren betreffen de sector, de opleidingsduur en het beroep dat de opleiding doet op het wiskundig inzicht van de student. De in het onderzoek onderscheiden sectoren zijn natuur & techniek, gezondheid, economie, recht, gedrag & maatschappij en taal & cultuur. De formele opleidingsduur is uitgedrukt in het aantal cursusjaren. De meeste studies in het Wetenschappelijk Onderwijs hebben een officiële cursusduur van vier jaren. Een uitzondering hierop vormt de sector natuur & techniek. Het merendeel van deze studies kent een vijfjarige studieduur.

Het beroep dat de studie doet op het wiskundig inzicht van de student wordt uitgedrukt in de mate van wiskundigheid van de vooropleidingsvereisten. Deze kunnen als volgt gerangschikt worden: 1. geen vereisten, 2. wiskunde a of b, 3. natuurkunde en scheikunde, 4. wiskunde b en natuurkunde.

Omgevingsfactoren

De omgevingsfactoren die in het model opgenomen worden zijn het werkgedrag van de student en het gegeven of hij onder het tempobeursregime of prestatiebeursregime valt. In de steekproef behoren 547 studenten tot het prestatiebeursregime. Dit zijn de eerstejaars studenten in het collegejaar 1996-1997. De overige studenten (N=1854) vallen onder het tempobeursregime.

Het werkgedrag van de student valt uiteen in twee delen. Enerzijds wordt het aantal gewerkte uren per week in het collegejaar 1996-1997 als variabele opgenomen (Bean en Metzner, 1985). Anderzijds wordt aan de werkende studenten een drietal stellingen over de inhoud van het werk voorgelegd (zie bijlage I). De antwoordmogelijkheden zijn wederom gerangschikt van 1. helemaal mee oneens tot en met 4. helemaal mee eens. Het opnemen van deze werkstellingen geschiedt om werk dat gerelateerd is aan de studie te kunnen onderscheiden van overige werkzaamheden. Het gemiddeld aantal gewerkte uren per week en de gemiddelde waarden voor de werkhoudingen worden weergegeven in tabel 3.

Studievoortgang

Het te verklaren concept, studievoortgang, wordt gemeten aan de hand van twee indicatoren. De eerste indicator is het aantal studiepunten dat de studenten in het collegejaar 1996-1997 hebben behaald.⁴ De tweede indicator wordt gevormd door het percentage gerealiseerde studielast van het officiële curriculum. Deze indicator biedt zicht op de voorsprong dan wel achterstand in studieprogressie van de studenten, gecorrigeerd voor eventuele jaareffecten. Een negatieve waarde voor deze variabele duidt op een studie-achterstand.

De studenten behaalden in studiejaar 1996-1997 gemiddeld 29.1 punten. De gemiddelde achterstand op het officiële curriculum bedroeg 11%. Vrouwen behaalden in het collegejaar 1996-1997 significant meer studiepunten dan mannen, de achterstand die vrouwen gedurende hun studie oplopen is kleiner. Prestatiebeursstudenten hebben een substantieel kleinere achterstand opgelopen dan tempobeursstudenten, waarbij het gegeven dat de eerstgenoemde groep uit eerstejaarsstudenten bestaat en de laatstgenoemde groep uit ouderejaars studenten, is verdisconteerd. Opvallend is wel, dat prestatiebeurs-studenten in het collegejaar 1996-1997 significant meer studievoortgang boekten dan tempobeursstudenten. Tempobeursstudenten scoeren hoger op stellingen waarmee de werkhoudingen gemeten worden.

Tabel 3: Omgevingsfactoren en studievoortgang.

	N	werkattitudes	N	uren werk per week	% voorsprong op curriculum	studiepunten
totaal	1676	7.2	2401	8.0	-11	29.1
man	964	7.2	1400	8.0	-13	28.0
vrouw	712	7.2	1001	7.9	-8	30.7
tempobeurs	1340	7.4	1854	8.8	-12	28.5
prestatiebeurs	336	6.6	547	5.0	-7	31.1
n & t	380	7.1	606	5.1	-15	28.1
gezondheid	202	7.2	318	6.5	-5	32.0
economie	402	7.6	552	9.0	-11	30.3
recht	248	7.1	327	10.0	-14	25.9
g & m	262	7.1	329	10.7	-10	29.1
t & c	182	7.0	269	8.1	-8	29.5

Studenten uit de sectoren natuur & techniek en gezondheid besteden beduidend minder tijd aan betaald werk dan studenten uit de sectoren economie, recht en gedrag & maatschappij. Natuur & techniek-studenten besteden ook minder tijd aan deze activiteit dan taal & cultuur-studenten. Gedrag & maatschappij-studenten besteden substantieel meer tijd aan betaald werk dan taal & cultuur-studenten.

De achterstand van natuur- en techniek en rechtenstudenten op de volgens het studierooster te behalen studiepunten is, in vergelijking met de andere sectoren, groot. De achterstand van studenten in de sectoren gezondheid en taal & cultuur is relatief gering.

De studievoortgang van rechtenstudenten en natuur- & techniekstudenten verschilt in het collegejaar 1996-1997 niet significant. Beide groepen behaalden beduidend minder studiepunten dan studenten uit de sector gezondheid. Daarnaast behaalden rechtenstudenten minder studievoortgang dan studenten uit de sectoren economie, gedrag & maatschappij en taal & cultuur.

4. RESULTATEN

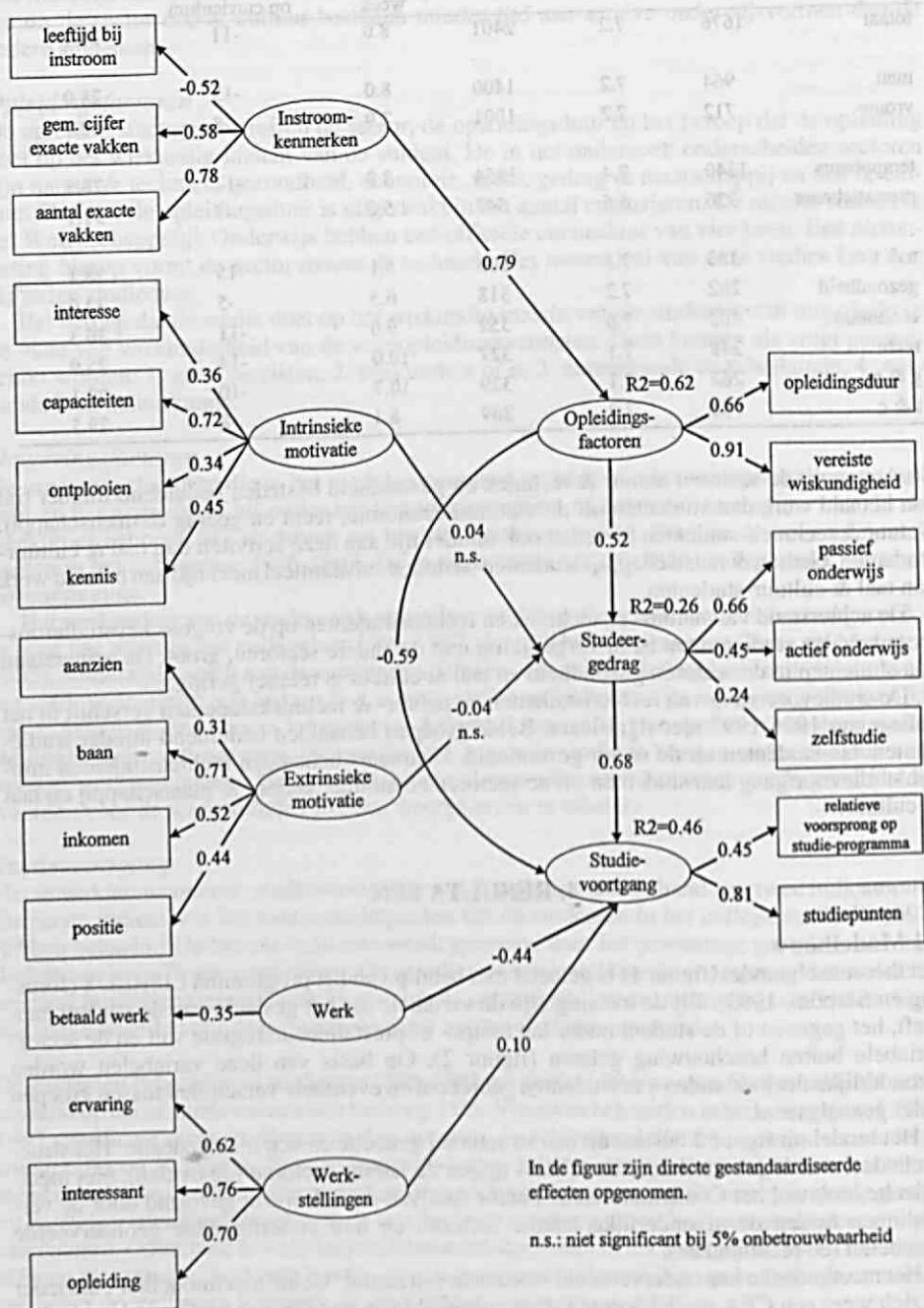
4.1 Modelbouw

Het theoretisch model (figuur 1) is getoetst met behulp van het programma LISREL 8 (Jöreskog en Sörbom, 1993). Bij de toetsing zijn de variabele die het geslacht van de student aangeeft, het gegeven of de student onder het tempo- of prestatiebeursregime valt en de sectorvariabele buiten beschouwing gelaten (figuur 2). Op basis van deze variabelen worden afzonderlijke deelbestanden van studenten gecreëerd en eventuele verschillen tussen groepen nader geanalyseerd.

Het model uit figuur 2 bestaat uit een structureel gedeelte en een meetgedeelte. Het structurele deel wordt gevormd door de relaties tussen de latente factoren (de ovals). Het meetgedeelte, ook wel het Confirmatorische Factor Analyse-model, wordt gevormd door de verbandingen tussen de afzonderlijke latente factoren en hun onderliggende geobserveerde variabelen (de rechthoeken).

Het meetgedeelte kan onderverdeeld worden in een aantal 'kleine meetmodellen', die ieder op zich weer een CFA-model vormen. Een voorbeeld van een CFA-model is de latente factor studeergedrag, die wordt gerepresenteerd door de variabelen 'passief onderwijs', 'actief

onderwijs' en 'zelfstudie'. Figuur 2 laat zien dat deze drie variabelen significante indicatoren vormen voor het concept studeergedrag.



Figuur 2: Model ter verklaring van studievoortgang voor de totaalgroep.

Het opbouwen van het model uit figuur 2 vindt plaats in drie stappen. De eerste stap betreft het construeren van 'kleine meetmodellen' met een acceptabele fit. Een acceptabele fit houdt in, dat het model goed gespecificeerd is, statistisch overgeïdentificeerd is en bij toetsing niet verworpen wordt. Indien de fit van de kleine meetmodellen acceptabel is, worden zij in stap 2 samengevoegd tot het 'grote meetmodel'. Nadat van het grote model een acceptabele fit verkregen is, worden de relaties tussen de latente factoren gespecificeerd (stap drie). Het uiteindelijk verkregen model heeft een χ^2 -waarde van 126.10 bij 103 vrijheidsgraden ($p = 0.061$). De waarde voor de Comparatieve Fit Index is gelijk aan 1.00 en de standardized root mean square residual bedraagt 0.016. Dit houdt in dat de fit van het model acceptabel is⁵.

4.2 Toetsing van het generieke model

Instroomkenmerken (leeftijd bij instroom, aantal exacte vakken, gemiddeld eindexamencijfer voor exacte vakken) hangen samen met opleidingsfactoren (duur en wiskundigheid van de opleiding). Dit verband kan als volgt worden geïnterpreteerd: naarmate de student met minder voorafgaande vertragsjaren aan de studie is begonnen, jonger is en meer exacte vakken in zijn eindexamenpakket had en hogere cijfers voor de exacte vakken behaalde, kiest hij vaker voor een opleiding die een groter beroep doet op het wiskundig inzicht en voor een opleiding met een langere cursusduur. De variantie in het construct opleidingsfactoren kan met behulp van de instroomkenmerken voor 62% worden gebonden.

Het studeergedrag van de student, ofwel het aantal uren per week dat hij aan de diverse studieactiviteiten (passief onderwijs, actief onderwijs, zelfstudie) besteedt hangt positief samen met de opleidingsfactoren. Studenten die een meer wiskundige opleiding volgen en/of een opleiding met een langere cursusduur, besteden derhalve meer tijd aan hun studie. Zesentwintig procent van de variantie in het studeergedrag wordt door het model verklaard.

Intrinsieke alsook extrinsieke gemotiveerdheid voor de studie blijken geen wezenlijke invloeden op het studeergedrag van de studenten te bewerkstelligen.

De studievoortgang van studenten in het Wetenschappelijk Onderwijs wordt positief beïnvloed door het aantal uren per week dat de student aan de diverse onderwijsactiviteiten besteedt (studeergedrag). Opleidingsfactoren (vereiste wiskundigheid en opleidingsduur) vertonen een negatieve samenhang met de studievoortgang. De invloed van het aantal door studenten gewerkte uren is negatief, terwijl de werkatitudes een positieve invloed uitoefenen op de studievoortgang. De verklaarde variantie van de uiteindelijk te verklaren factor studievoortgang bedraagt 46%.

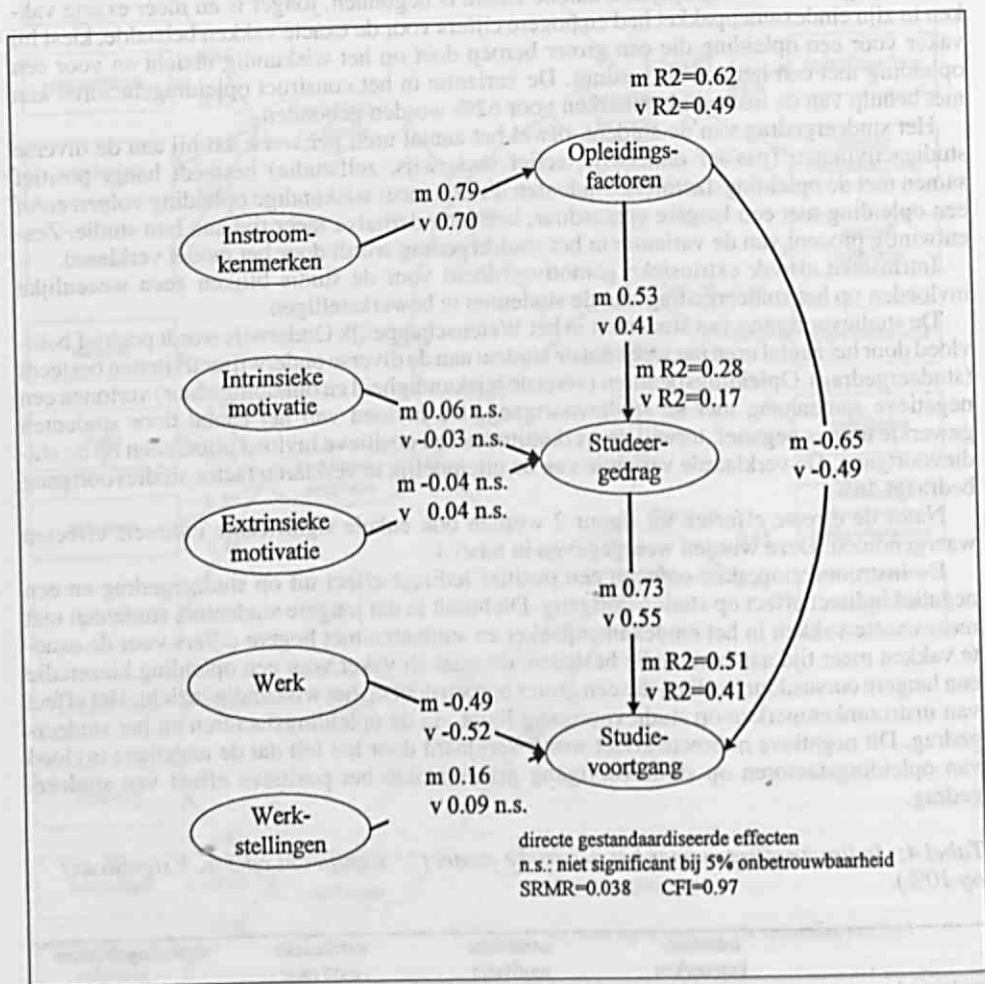
Naast de directe effecten uit figuur 2 worden ook enkele significante indirecte effecten waargenomen. Deze worden weergegeven in tabel 4.

De instroomkenmerken oefenen een positief indirect effect uit op studeergedrag en een negatief indirect effect op studievoortgang. Dit houdt in dat jongere studenten, studenten met meer exacte vakken in het eindexamenpakket en studenten met hogere cijfers voor de exacte vakken meer tijd aan hun studie besteden, doordat zij vaker voor een opleiding kiezen die een langere cursusduur heeft en die een groter beroep doet op het wiskundig inzicht. Het effect van instroomkenmerken op studievoortgang loopt via de opleidingsfactoren en het studeergedrag. Dit negatieve indirecte effect wordt verklaard door het feit dat de negatieve invloed van opleidingsfactoren op studievoortgang groter is dan het positieve effect van studeergedrag.

Tabel 4: Indirecte effecten voor het generieke model (** significant op 5%, * significant op 10%).

	instroom- kenmerken	intrinsieke motivatie	extrinsieke motivatie	opleidingsfactoren
studeergedrag	0.41**			
studievoortgang	-0.19**	0.03	-0.02	0.35**

Uit tabel 4 kan worden afgeleid dat het indirecte effect van opleidingsfactoren op studievoortgang positief is (gestandaardiseerde bèta van 0.35). Het totale effect van de opleidingsfactoren op studievoortgang blijft echter negatief: het directe effect van -0.59 plus het indirecte effect van 0.35 levert een totaaleffect ter grootte van -0.24 (gestandaardiseerde bèta). Dit betekent dat studenten die een studie volgen waarvoor meer wiskundigheid is vereist en die langer duurt, significant minder studievoortgang behalen dan andere studenten. Zij besteden echter meer tijd aan hun studie, hetgeen een positief effect heeft op studievoortgang. Het negatieve effect is echter sterker dan het positieve effect. En daarmee is de eerdere constatering ongewijzigd: naarmate de officiële cursusduur langer is en het voor de opleiding vereiste wiskundig inzicht groter is, behaalt de student een geringere studievoortgang. De gerealiseerde studielast uitgedrukt in een percentage van het officiële curriculum is voor deze studenten geringer en zij behaalden ook in het collegejaar 1996-1997 een kleiner percentage van de formele jaarlijkse studielast, die voor alle studies gelijk is aan 42 studiepunten.



Figuur 3: Modellen ter verklaring van studievoortgang voor mannen en vrouwen.

4.3 Sekse-specifieke verklaringen voor studievoortgang

Een nadere analyse van sekse-specifieke determinanten van studievoortgang (1400 mannen en 1001 vrouwen) door middel van een meergroepen analyse levert de volgende resultaten:

De eerste gevolgtrekking uit figuur 3 is dat met behulp van de in het analysemodel opgenomen factoren de studievoortgang van mannelijke studenten beter kan worden voorspeld dan die van hun vrouwelijke collega's (R^2 studievoortgang mannen 51%; vrouwen 41%).

Bezien we de directe effecten die de afhankelijke latente variabelen op elkaar uitoefenen dan vallen drie zaken op. De positieve effecten van opleidingsfactoren op studeergedrag en studeergedrag op studievoortgang lijken voor de mannen groter dan voor de vrouwen. Hetzelfde geldt voor het negatieve directe effect van opleidingsfactoren op studievoortgang.

De onafhankelijke constructen instroomkenmerken en werk oefenen in beide modellen significante invloeden uit op respectievelijk de opleidings- en studievoortgangsfactor. De relatie tussen werkstellingen en studievoortgang is voor de vrouwen niet significant. Voor mannen lijkt het instroomeffect op de opleidingsfactoren en de invloed van de werkhoudingen op studievoortgang sterker dan voor vrouwen.

Om vast te stellen hoe substantieel deze verschillen zijn wordt voor de man/vrouw modellen per structurele vergelijking nagegaan of er daadwerkelijk significante verschillen tussen de seksen optreden. Daartoe zijn steeds twee modellen met elkaar vergeleken, waarbij in model a de desbetreffende parameter is gelijkgesteld voor beide groepen terwijl deze in model b van elkaar verschillen. Dit laatste model schat dus 1 parameter meer. Indien het schatten van deze extra parameter een verbetering in chi-kwadraat waarde van 3.84 of meer teweeg brengt (kritieke waarde 1 vrijheidsgraad chi-kwadraat op 5%) dan concluderen we dat de parameters van beide groepen significant van elkaar verschillen op het 5%-niveau. Een verbetering van chi-kwadraat waarde tussen de 2.7 en 3.84 duidt op een verschil op het 10%-niveau.

Uit de toetsing blijkt dat het effect van studeergedrag op studievoortgang tussen mannen en vrouwen inderdaad significant van elkaar afwijkt op het 5%-niveau (effect mannen: 0.73; vrouwen: 0.55). Geconcludeerd kan worden dat extra studie-uren voor mannen meer studievoortgang opleveren dan voor vrouwen. Ten aanzien van de andere relaties wordt geen significant verschil geconstateerd.

Tabel 5: Indirecte effecten voor mannen en vrouwen (** significant op 5%, * significant op 10%).

		instroom- kenmerken	intrinsicke motivatie	extrinsicke motivatie	opleidings- factoren
studeergedrag	mannen	0.42**			
	vrouwen	0.29**			
studievoortgang	mannen	-0.21**	0.04	-0.03	0.39**
	vrouwen	-0.18**	-0.02	0.02	0.23**

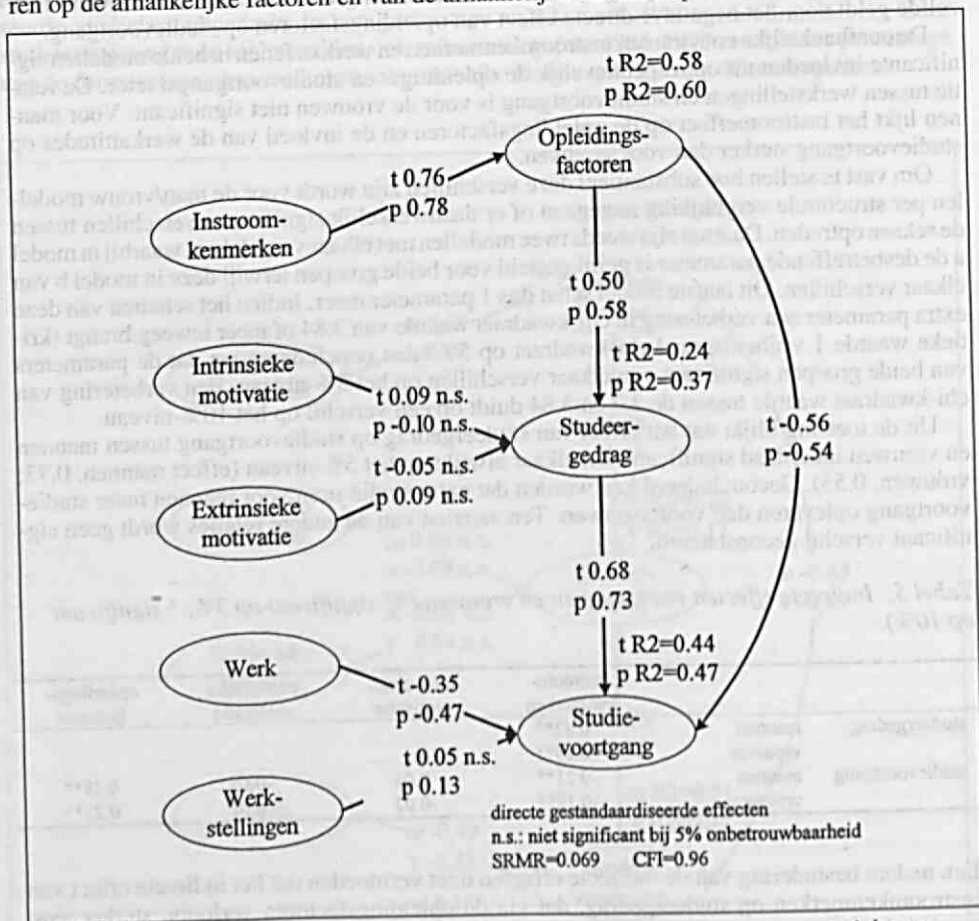
Een nadere bestudering van de indirecte effecten doet vermoeden dat het indirecte effect van instroomkenmerken op studeergedrag, dat via de opleidingsfactoren verloopt, sterker aanwezig is in het mannenmodel (tabel 5). Dezelfde constatering wordt gedaan voor het indirecte effect van opleidingsfactoren op studievoortgang. Deze indirecte relatie verloopt via het studeergedrag.

De constatering dat jongere studenten, studenten met meer exacte vakken en hogere cijfers voor die vakken meer tijd aan hun studie besteden, en dat die studie vaker een relatief lange cursusduur en een hogere moeilijkheidsgraad kent, geldt dus meer voor mannen dan voor vrouwen.

4.4 Effecten van onderwijsbeleid

Een relevante vraag betreft die naar de mate waarin effecten van het Nederlandse onderwijsbeleid op de studievoortgang in het Wetenschappelijk Onderwijs zijn te traceren. Daartoe is een meergroepen analyse uitgevoerd met als uitgangspunt de studiefinancieringsregimes waaronder de studenten functioneren. We constateren allereerst dat de modellen voor temporebeursstudenten ($N=1854$) en prestatiebeursstudenten ($N=547$) geen sterk uiteenlopende verklaringskracht blijken te hebben ten aanzien van de studievoortgang (R^2 prestatiebeurs: 0.47; temporebeurs: 0.44).

Figuur 4 bevat de gestandaardiseerde directe effecten van de onafhankelijke latente factoren op de afhankelijke factoren en van de afhankelijke factoren op elkaar.



Figuur 4: Modellen ter verklaring van studievoortgang voor tempo- en prestatiebeurs.

De invloed van instroomkenmerken op opleidingsfactoren is voor tempo- en prestatiebeursstudenten gelijk. Blijkbaar is er in de loop van de tijd weinig veranderd in de relatie tussen keuze en aanleg voor exacte vakken zoals gemeten in het Voortgezet Onderwijs en de keuze voor een exacte studie. Het directe effect van intrinsieke motivatie, hoewel zij in geen van beide modellen significant van nul afwijkt, verschilt tussen de prestatiebeursgroep en de temporebeursgroep op het 10%-niveau. De intrinsieke motivatie werkt naar het zich laat aanzien positiever uit onder het temporebeursregime.

Zowel in het tempobeurs- als prestatiebeursmodel is de invloed van het aantal gewerkte uren op de studievoortgang significant negatief. Dit betekent dat een groter aantal gewerkte uren voor de studenten leidt tot een significant kleinere studievoortgang. Voor prestatiebeursstudenten is het negatieve effect van werken op studievoortgang evenwel nog sterker (het effect verschilt tussen beide groepen op het 10% onbetrouwbaarheidsniveau). Een mogelijke verklaring voor dit verschil is dat de prestatiebeurs nog maar recent in werking is getreden en de prestatiebeursstudenten dus meer jongerejaars zijn dan tempobeursstudenten. Ouderejaars studenten zijn over het algemeen beter geïntegreerd in de universitaire samenleving, zowel op sociaal als academisch gebied. Werken naast de studie levert voor deze groep daardoor minder studievertraging op dan voor jongerejaars studenten, die hun draai nog moeten vinden.

Het is ook mogelijk dat het verschil in beursregime ertoe bijdraagt dat de studievoortgang van prestatiebeursstudenten anders beïnvloed wordt dan de studievoortgang van tempobeursstudenten. Op dit moment is de impact van de invoering van de prestatiebeurs echter nog niet eenduidig vast te stellen. Vervolgmetingen, die in de komende jaren worden verricht, kunnen antwoord geven op de vraag of de invloed van het werkgedrag op de studievoortgang van prestatiebeursstudenten anders verloopt dan die van tempobeursstudenten, of dat het werkgedrag van jongerejaars studenten een andere invloed uitoefent op de studievoortgang dan het werkgedrag van ouderejaarsstudenten, of dat het een combinatie is van beide factoren.

Uit figuur 4 kan eveneens worden opgemaakt dat alle onderlinge invloeden van de afhankelijke factoren significant zijn. Er wordt vastgesteld dat het effect van opleidingsfactoren op studeergedrag voor de prestatiebeursgroep (eerstejaars studenten) iets sterker is dan voor de tempobeursgroep (ouderejaars studenten). Prestatiebeursstudenten die een wiskundige (technische) opleiding volgen met een langere cursusduur besteden meer tijd aan hun studie dan tempobeursstudenten die een vergelijkbare opleiding volgen.

Tabel 6: Indirecte effecten voor tempobeurs en prestatiebeurs (** significant op 5%, * significant op 10%).

		instroom- kenmerken	intrinsieke motivatie	extrinsieke motivatie	opleidings- factoren
studeergedrag	tempobeurs	0.38**			
	prestatiebeurs	0.45**			
studievoortgang	tempobeurs	-0.17**	0.06	-0.03	0.34**
	prestatiebeurs	-0.10*	-0.07	0.06	0.42**

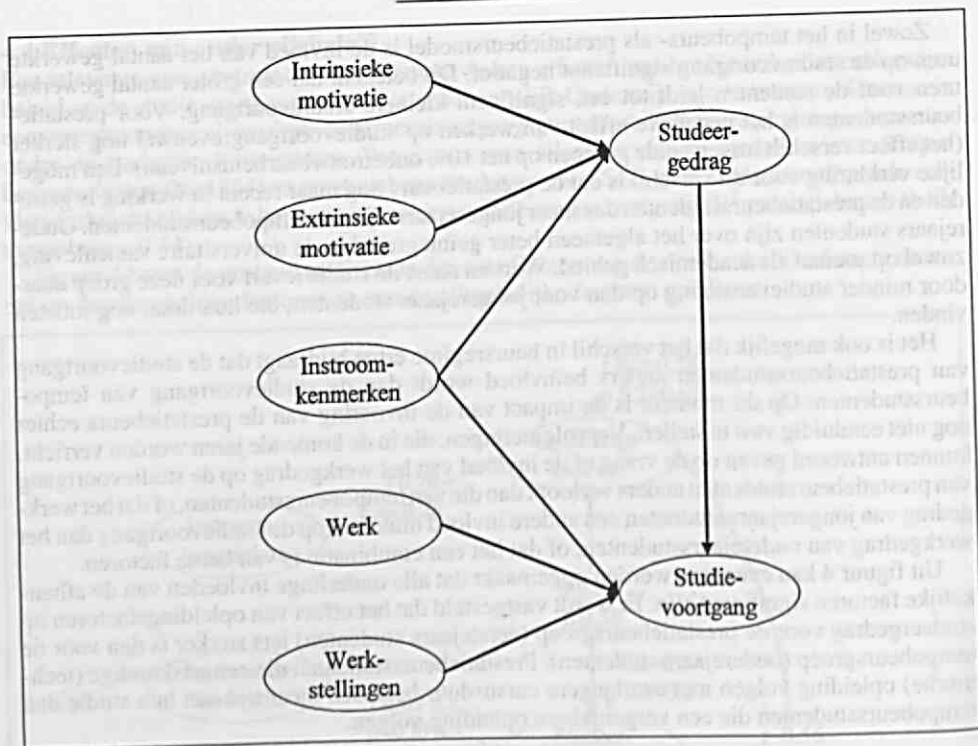
Het indirecte positieve effect van opleidingsfactoren op studievoortgang, dat via het studeergedrag verloopt, bevestigt dit. De prestatiebeursgroep lijkt zich, meer dan de tempobeursgroep, te realiseren dat een technische (meer exacte) studie een grote tijdsinvestering vergt.

4.5 Sector-specifieke verklaringen voor studievoortgang

Het generieke model is tevens aan een toetsing voor de afzonderlijke sectoren natuur & techniek, gezondheid, economie, recht, gedrag & maatschappij en taal & cultuur onderworpen. In figuur 5 wordt het model weergegeven op basis waarvan sectorale vergelijking van determinanten van studievoortgang mogelijk is. Aangezien de opleiding het selectie criterium vormt voor de geconstrueerde deelbestanden maken de opleidingsfactoren geen deel uit van het te toetsen model.⁶

In de tabellen 7 en 8 worden de directe gestandaardiseerde invloeden op studeergedrag respectievelijk studievoortgang weergegeven.⁷

Uit de tabel blijkt dat intrinsieke motivatie en extrinsieke motivatie slechts in enkele gevallen significant samenhangen met het studeergedrag. De bevinding bij het generieke model, namelijk geen effect van motivatie op studeergedrag, wordt door de sectormodellen dan ook grotendeels ondersteund. Opvallend is met name dat rechtenstudenten die beschikken over



Figuur 5: Model ter verklaring van studievoortgang voor sectoren.

een hoge intrinsieke motivatie dit ook daadwerkelijk in het studeergedrag tonen.

Instroomkenmerken hangen in twee van de zes sectoren samen met het studeergedrag. In de sector natuur & techniek besteden studenten met meer exacte vakken in het eindexamenpakket, hogere cijfers voor de exacte vakken en jonger ingestroomde studenten meer tijd aan hun studie. Het omgekeerde effect wordt waargenomen in de sector recht.

Tabel 7: Directe effecten op studeergedrag naar sector (** significant op 5%; * significant op 10%).

	intrinsieke motivatie op studeergedrag	extrinsieke motivatie op studeergedrag	instroomkenmerken op studeergedrag
natuur en techniek	0.10	-0.20*	0.48**
gezondheid	0.02	0.05	-0.01
economie	-0.03	-0.06	0.19
recht	0.34**	-0.20	-0.37**
gedrag en maatschappij	0.04	0.18	0.01
taal en cultuur	0.15	-0.14	0.04

Het generieke negatieve effect van werken op studievoortgang wordt niet gevonden indien het databestand opgesplitst wordt in de zes onderscheiden sectoren. Het effect van de werkattitudes (intrinsieke motivatie voor het werk) op de studievoortgang is in de sector gedrag & maatschappij significant positief, terwijl het in de taal & cultuur-sector negatief is.

Tabel 8: Directe effecten op studievoortgang naar sector (** significant op 5%, * significant op 10%).

	werk op studievoortgang	werk-attitudes op studievoortgang	instroomkenmerken op studievoortgang	studeergedrag op studievoortgang
natuur en techniek	-0.11	-0.11	0.15	0.66**
gezondheid	-0.11	-0.05	0.47**	0.15
economie	-0.11	-0.18	0.52**	0.02
recht	-0.04	-0.05	0.74**	0.66**
gedrag en maatschappij	0.10	0.27**	0.50**	0.19
taal en cultuur	0.08	-0.29**	0.39**	0.51**

Opvallend is, dat de instroomkenmerken voor vijf van de zes sectoren een significant positieve directe invloed uitoefenen op studievoortgang. Dit impliceert dat jongere studenten, studenten met meer exacte vakken in het eindexamenpakket en hogere cijfers voor de exacte vakken, in alle sectoren met uitzondering van natuur & techniek, significant sneller studeren dan de studenten die deze eigenschappen in mindere mate bezitten⁵.

Het studeergedrag oefent in de sectoren natuur & techniek, taal & cultuur en recht een significant invloed uit op studievoortgang. Voor studenten uit deze sectoren geldt dat niet alleen de leeftijd en de bij de vooropleiding behaalde resultaten voor de exacte vakken van belang zijn voor studiesucces, maar dat tevens het aantal uren dat door de studenten aan hun studie besteed wordt een cruciale factor is. Een hogere studie-inzet leidt in deze drie sectoren tot een significant grotere studievoortgang.

Tot slot worden in tabel 9 de indirecte effecten op studievoortgang gepresenteerd.

Tabel 9: Indirecte effecten op studievoortgang naar sector (** significant op 5%, * significant op 10%).

	intrin op studievoortgang	extrin op studievoortgang	instroom op studievoortgang
natuur en techniek	0.06	-0.13*	0.31**
gezondheid	0.00	0.01	0.00
economie	0.00	0.00	0.00
recht	0.23**	-0.13	-0.24**
gedrag en maatschappij	0.01	0.03	0.00
taal en cultuur	0.08	-0.07	0.02

Opvallend is het positieve indirecte effect van instroomkenmerken via het studeergedrag op de studievoortgang in de sector natuur & techniek en het negatieve indirect hiervan in de sector recht. We stellen vast dat studenten die jonger zijn bij instroom in het Wetenschappelijk Onderwijs en een hogere mate van wiskundige aanleg hebben, in de sector natuur & techniek meer studievoortgang behalen doordat zij meer tijd aan hun studie besteden. Het directe effect van instroomkenmerken op studievoortgang is, zoals reeds uit tabel 8 bleek, niet significant. Het totaleffect (indirect + direct effect) is echter wel significant positief. In de sector recht daarentegen wordt door studenten die jonger instromen, hogere cijfers voor exacte vakken behaald hebben en meer exacte vakken in het eindexamenpakket hebben, minder gestudeerd (tabel 7), waardoor het indirecte effect op studievoortgang negatief is. Het directe effect van instroomkenmerken op studievoortgang is voor rechtenstudenten echter positief (tabel 8). Hetzelfde geldt voor het totale effect (indirecte + directe effect) waaruit geconcludeerd kan worden dat het effect van instroomkenmerken domineert. Wellicht is de rechtenstudie voor de jongere, relatief intelligente, meer exact opgeleide student met beduidend minder inspanning te volbrengen dan voor de andere studenten.

5. CONCLUSIES

In deze studie ter verklaring van studievoortgang van studenten in het Wetenschappelijk Onderwijs in Nederland spelen opleidingsfactoren, omgevingsfactoren en studentfactoren een centrale rol.

We stellen vast dat studenten die een opleiding volgen waarvoor een groter wiskundig inzicht vereist is en/of een opleiding volgen met een langere cursusduur, meer uren per week aan hun studie besteden.

Het indirecte effect van opleidingsfactoren op studievoortgang is positief, maar het directe effect is sterk negatief. Dit impliceert dat studenten die een opleiding volgen met een lange cursusduur en met een groot beroep op het wiskundig inzicht aanzienlijk minder studievoortgang behalen dan andere studenten. Zelfs het feit dat zij meer tijd aan hun studie besteden dan andere studenten (het indirecte positieve effect) kan niet voorkomen dat zij meer studie-vertaling oplopen. Gecombineerd met het feit dat studenten die exacte (technische) studies kiezen gemiddeld hogere eindexamencijfers behalen en meer exacte vakken kiezen kan uit het voorgaande worden geconcludeerd dat exacte (technische) studies onevenredig zwaar zijn in vergelijking met andere studies.

Dit betekent dat de eerste hypothese *"Naarmate het voor de opleiding vereiste wiskundig inzicht groter is, besteedt de student meer uren per week aan zijn studie, hetgeen een (indirect) positief effect uitoefent op de studievoortgang. Het directe effect op studievoortgang is negatief"* wordt bevestigd. Toegevoegd kan worden, dat het directe negatieve effect domineert en dat het totaaleffect van opleidingsfactoren op studievoortgang derhalve negatief is.

Intrinsieke motivatie, dit is de mate waarin de student inhoudelijk geïnteresseerd is in zijn opleiding en extrinsieke motivatie, dit is de mate waarin de student met name materieel (financieel) geïnteresseerd is in zijn opleiding, oefenen geen significante invloed uit op studeergedrag⁹. De tweede hypothese, *"Een grotere (intrinsieke dan wel extrinsieke) motivatie gaat gepaard met meer aan de studie bestede uren, hetgeen een positief effect uitoefent op de studievoortgang"* wordt daarmee verworpen.

De derde hypothese, waarin gesteld werd dat studenten die meer uren per week werken naast de studie minder studievoortgang behalen, wordt bevestigd. Werken hangt inderdaad negatief samen met studievoortgang. De werkhoudingen, waarmee aangegeven wordt of de student werkt om werkervaring op te doen, of het werk inhoudelijk interessant is en in hoeverre het werk aansluit bij de opleiding, oefenen een bescheiden positieve invloed uit op studievoortgang. Geconcludeerd kan worden dat een beleidsmaatregel die erop gericht is om studenten meer studiefinanciering te verstrekken, kan leiden tot een hoger studietempo, doordat studenten in mindere mate genoodzaakt zijn om betaald werk naast de studie te verrichten. Het feit dat studenten werken uit noodzaak wordt ondersteund door het onderzoek 'Studeren en werken in het Wetenschappelijk Onderwijs' (De Boom et al., 1998). Hieruit blijkt dat 69% van de studenten werkt om te voorzien in de kosten voor het levensonderhoud. De stelling 'Ik werk om minder of geen geld te hoeven lenen' is op 53% van de studenten van toepassing.

Toetsing van aparte deelmodellen voor mannen en vrouwen laat zien dat de in het analysemodel opgenomen factoren de studievoortgang van mannen beter voorspellen dan die van vrouwen. Het effect van studeergedrag op studievoortgang is in beide deelmodellen sterk positief maar is voor mannen groter (0.73) dan voor vrouwen (0.55). Dit houdt in dat extra studie-uren voor mannen meer studievoortgang opleveren dan voor vrouwen.

Toetsing van sectorale verschillen in modelverbanden laat zien dat instroomkenmerken in alle sectoren, met uitzondering van natuur & techniek, een positieve directe invloed uitoefenen op de studievoortgang. Dit betekent dat studenten met meer bèta-aanleg en studenten die jonger ingestroomd zijn in het Wetenschappelijk Onderwijs significant sneller studeren dan studenten die deze eigenschappen in mindere mate beheersen. Daarnaast is in de sectoren recht, taal & cultuur en natuur & techniek een positieve directe invloed van studeergedrag op studievoortgang waarneembaar.

Opvallend is het gegeven dat studenten met meer exacte vakken in het eindexamenpakket, hogere cijfers voor deze vakken en een lagere leeftijd bij instroom in de sector natuur & techniek meer tijd aan hun studie besteden terwijl voor studenten in de sector recht het omgekeerde effect geconstateerd wordt. Het gegeven dat het directe effect van instroomkenmerken op studievoortgang voor de laatste groep positief is, en dat dit eveneens geldt voor het totaaleffect leidt tot de conclusie dat het effect van instroomkenmerken het effect van studiezet domineert. Voor de sector natuur & techniek moet deze conclusie worden genuanceerd omdat hier het totaal-effect en het indirecte effect van instroomkenmerken op studievoortgang positief is, terwijl het directe effect niet significant is. Beta-inzicht en een jongere leeftijd bij instroom in het Wetenschappelijk Onderwijs leiden in deze sector derhalve niet zonder meer tot een grotere studievoortgang, maar slechts indirect, via een grotere studie-inspanning.

Tevens is nagegaan in hoeverre effecten van het Nederlandse onderwijsbeleid op de studievoortgang in het Wetenschappelijk Onderwijs zijn te traceren. Geconstateerd is dat prestatiebeursstudenten (eerstejaars studenten) een beduidend grotere studievoortgang laten zien dan tempobeursstudenten (ouderejaars studenten). Hypothese vier wordt dus geaccepteerd. Ter verklaring van dit verschijnsel zijn aparte modellen geconstrueerd voor de onderscheiden studiefinancieringsregimes. Het belangrijkste verschil tussen tempobeurs- en prestatiebeursstudenten betreft de invloed van werkgedrag op studievoortgang. Werken naast de studie leidt voor beide groepen studenten tot een geringere voortgang, dit effect is voor de prestatiebeursgroep echter significant groter dan voor de tempobeursgroep (10% onbetrouwbaarheidsniveau). Een mogelijke verklaring voor dit feit is dat prestatiebeursstudenten vaker jongerejaars zijn dan tempobeursstudenten en dat zij nog minder goed geïntegreerd zijn in de universitaire samenleving. Het is ook denkbaar dat verschillen in beursregimes ertoe bijdragen dat de studievoortgang van prestatiebeursstudenten anders beïnvloed wordt dan van tempobeursstudenten. Vervolgmetingen die in de komende jaren worden verricht, kunnen antwoord geven op de vraag of de invloed van het werkgedrag op de studievoortgang van prestatiebeursstudenten anders verloopt dan die van tempobeursstudenten, of dat het werkgedrag van jongerejaars studenten een andere invloed uitoefent op de studievoortgang dan het werkgedrag van ouderejaarsstudenten, dan wel dat het een combinatie betreft van beide factoren.

NOTEN

1. Prestatiebeursstudenten ontvangen bij een vierjarige opleiding gedurende vier jaar beurs. Bij tempobeursstudenten die een vierjarige opleiding volgen wordt de beurs gedurende vijf jaar verstrekt. Tempobeursstudenten dienen jaarlijks minimaal 21 studiepunten te behalen. Prestatiebeursstudenten dienen in het eerste studiejaar minimaal 21 studiepunten te behalen en dienen binnen zes jaar minimaal 168 studiepunten te hebben behaald.

2. Een sector is een verzamelnaam voor studies die betrekking hebben op hetzelfde vakgebied. De sectoren die in het RISBO-onderzoek onderscheiden worden zijn: natuur & techniek (n&t), gezondheid, economie, recht, gedrag & maatschappij (g&m) en taal & cultuur (t&c). De sector natuur & techniek is een samenstelling van het HOOP-gebied natuur en het HOOP-gebied techniek. De overige sectoren komen overeen met de indeling in HOOP-gebieden.

Een cohort is een samenvoeging van diverse startjaren. In het RISBO-onderzoek worden drie cohorten onderscheiden. Cohort I bevat de eerste- en tweedejaars studenten, cohort II de derde- en vierdejaars studenten en cohort III de overige studenten (vijfdejaars en ouder).

3. Per cel, dit is een combinatie van de strata, wordt het gewicht als volgt berekend (De Boom et al., 1998):

$$\text{gewicht} = (\text{populatie}_{\text{groep}} / \text{populatie}_{\text{totaal}}) * (\text{steekproef}_{\text{totaal}} / \text{steekproef}_{\text{groep}})$$

4. Volgens het officiële curriculum dienen jaarlijks 42 studiepunten behaald te worden. De totale studielast van een vierjarige opleiding bedraagt derhalve 168 studiepunten en van een vijfjarige opleiding 210 studiepunten.

5. De waarde voor de CFI komt als volgt tot stand. $CFI = |((\chi_i^2 - df_i) - (\chi_h^2 - df_h)) / (\chi_i^2 - df_i)|$ waarbij $\chi_i^2 > df_i$ en $\chi_h^2 \geq df_h$. Hierbij staat i voor het onafhankelijke model en h voor het hypothetische model. Met df wordt het aantal vrijheidsgraden aangeduid. Een CFI-waarde tussen 0.90 en 1.00 geeft aan dat het model een acceptabele fit heeft (Bentler, 1995). Een lage standardized root mean square residuel (SRMR), van ongeveer 0.10 of lager, duidt op een acceptabele fit.

6. De comparatieve fit index van het verkregen model (CFI) bedraagt 0.90, de standardised root mean square residuel (SRMR) is 0.054.

7. De verklaarde variantie van de afhankelijke variabele studievoortgang loopt uiteen van 28% in de sector gezondheid tot 59% in de sector natuur & techniek en zelfs 67% in de sector recht. Dit houdt in dat de onderscheiden determinanten in de sector recht (en ook natuur & techniek) een grotere rol spelen bij het verklaren van studievoortgang dan bij de overige sectoren.

8. Mogelijk wordt de uitzonderingspositie van de sector natuur & techniek verklaard door het gegeven dat de variantie in instroomkenmerken hier relatief laag is. Dit geldt met name voor de variabele 'leeftijd bij instroom'.

9. Ook De Jong e.a. (1997) vinden slechts een zwak effect van intrinsieke en extrinsieke motivatie. Prins (1997) constateert dat studenten met een hogere intrinsieke motivatie minder vaak uitvallen.

LITERATUUR

- Bean, J.P. en B.S. Metzner (1985). 'A conceptual model of nontraditional undergraduate student attrition'. *Review of educational research*, vol. 55, no.4 (winter), p.485-540.
- Bentler, P.M. (1995). *EQS Structural Equations Program Manual*. Encino, CA: Multivariate Software, Inc.
- Bijleveld, R. J. (1993). *Numeriek rendement en studiestaking*. Enschede: Universiteit Twente: CSHOB: proefschrift.
- Boom, J. de, W.H.A. Hofman, P. van Wensveen en M.N. van den Berg (1998). *Studeren en werken in het Wetenschappelijk Onderwijs: een onderzoek naar de relatie tussen werken, lenen en studievoortgang*. Rotterdam: RISBO Contractresearch BV/ Erasmus Universiteit Groep.
- Dyck, M. van, J. van den Elshout en K. Hoogeveen (1985). *Van voortgezet onderwijs naar HBO: problemen, oorzaken, oplossingen?* Den Haag: SVO.
- Jansen, E.P.W.A. (1996). *Curriculumorganisatie en studievoortgang*. Groningen: GION, Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling, afdeling COWOG, Rijksuniversiteit Groningen: proefschrift.
- Jong, U. de, J. Roeleveld, H.D. Webbink en A.E. Verbeek (1997). *Het Amsterdamse studieloopbaanmodel*. Den Haag: N.V. Sdu.
- Jong, U. de, M. van Leeuwen, J. Roeleveld en D. Webbink (1998). *Deelname aan Hoger Onderwijs deel 1. Studiekeuze en motieven van eerstejaars 1995/96*. Amsterdam: SCO/SEO.
- Jöreskog, K.G. en D. Sörbom (1993). *Lisrel®8: Structural Equation Modeling with the SIMPLIS™ Command Language*. Chicago: Scientific Software International, Inc.
- Pascarella, E.T. en P.T. Terenzini (1983). 'Predicting voluntary freshman year persistence/ withdrawal behavior in a residential university: A path analytic validation of Tinto's model'. *Journal of educational psychology*, 75, 2, p. 215-226.
- Prins, J.B.A. (1997). *Studieuitval in het wetenschappelijk onderwijs*. Nijmegen: Instituut voor onderwijskundige dienstverlening (IOWO): proefschrift.
- Spady, W. (1970). 'Dropouts from higher education: An interdisciplinary review and synthesis'. *Interchange*, vol. 1, p. 64-85.
- Tinto, V. (1987). *Leaving college: rethinking the causes and cures of student attrition*. Chicago: The University of Chicago Press.
- Verbeek, F. en U. de Jong (1995). *Het Risico van studeren en werken: Een secundaire analyse over de ontwikkeling van het werkstudentchap in de jaren tachtig en negentig*. Amsterdam: Stichting Centrum voor Onderwijsonderzoek van de Universiteit van Amsterdam.
- Verrijt, A.M.H. (1996). *Algemene studentenenquête 1996*. Nijmegen: Instituut voor Onderwijskundige Dienstverlening (IOWO).

- Vos, K. de en P. Fonteijn (1997). *Actualisering leengedrag en schuldopbouw in de WSF*. Tilburg: Economisch Instituut Tilburg.
- Webb, M.W. (1990). *Development and testing of a theoretical model for predicting student degree persistence at four-year commuter colleges*. Boston: Paper presented at the Annual meeting of the American Educational Research Association.

BIJLAGE I

Intrinsieke studiemotivatie:

- Ik ben geïnteresseerd in mijn opleiding
- Door mijn opleiding kan ik mijn capaciteiten benutten
- Ik kan me ontplooiën in het studeren
- Ik heb behoefte aan kennisuitbreiding

Extrinsieke studiemotivatie:

- Door mijn opleiding kan ik meer aanzien verwerven
- Door mijn opleiding heb ik meer kans op een baan
- Ik heb meer kans op een hoog inkomen als ik mijn opleiding afgerond heb
- Door mijn opleiding kan ik een hogere positie bereiken

Werk-attitudes:

- Ik werk om werkervaring op te doen
- Het werk is inhoudelijk interessant
- Het werk sluit aan bij mijn opleiding

De stabiliteit en consistentie van differentiële schooleffecten

R.J. Bosker en H. Luyten

Faculteit der Toegepaste Onderwijskunde, Universiteit Twente

ABSTRACT

In this paper research is reported on the stability and consistency of differential school effects. Using a data set of 150 randomly sampled primary schools in The Netherlands it is investigated whether the school specific relation between socio-economic status and achievement is stable across grades and consistent across subject areas. The results indicate that schools are very stable with regard to their effects on disadvantaged pupils, and less so with respect to other pupils. As for the consistency question, schools turn out to be medium consistent effective across the subject areas. It is contended that the school or teacher level 'causes' of school effectiveness for underprivileged pupils are to be found outside the well-known educational effectiveness enhancing factors.

1. INLEIDING

Het onderzoek naar schooleffectiviteit heeft zijn oorsprong gevonden in het ongelijkheids-onderzoek. Coleman, zijn eigen werk samenvattend, schrijft: "What difference does the school a child goes to make in the child's achievement? How much do schools overcome the inequalities with which children come to school? One might wonder why educational researchers had not asked these questions long ago, why they had not always focused on outputs of schools in assessing their quality" (Coleman, 1990, p. 2). Coleman laat zien hoe belangrijk scholen in Amerika zijn voor leerlingen in achterstandssituaties. In leerjaar 6 kan het verschil in kwaliteit van de school die een leerling bezoekt 13% van de variantie in prestaties van leerlingen van Afrikaanse origine verklaren, en 17% voor leerlingen van Europese oorsprong. In leerjaar 12 zijn deze cijfers respectievelijk 29% en 22%. Uiteindelijk is dus met name voor de zwarte leerlingen de school die bezocht wordt uitermate belangrijk. Bosker (1990) constateert dat ook in Nederland scholen met name van belang zijn voor leerlingen uit achterstandssituaties, zij het dat niet aangegeven kan worden welke school- en instructiefactoren in dezen van belang zijn. Nuttall et al. (1989) melden een negatieve correlatie tussen de mate waarin Caribische leerlingen op Britse scholen voor voortgezet onderwijs achterblijven op autochtone leerlingen en de 'overall' kwaliteit van de school. Met andere woorden: de kwaliteit neemt toe, naarmate de achterstand van de Caribische leerlingen kleiner wordt. In Kref en De Leeuw (1991) blijkt echter dat schooleffecten niet verschillen voor verschillende subgroepen leerlingen (laag versus hoog IQ gekruist met laag versus hoog sociaal-economische status). De auteurs wijten dit echter aan mogelijke statistische schattingsproblemen. Sammons et al. (1993) tenslotte betogen op grond van secundaire analyses op Engelse data dat er überhaupt geen differentiële schooleffecten zijn, anders dan voor hoog- en laagpresteerders.

Niet zelden wordt verondersteld dat in het onderwijs zogenaamde Mattheuseffecten werkzaam zijn. Soms wordt ook wel gesproken van de 'fan-spread' (waaier) hypothese. Bedoeld wordt dat verschillen in aanvangsniveau tussen leerlingen gedurende hun schoolloopbaan steeds groter worden (Walberg & Tsai, 1983; Stanovich, 1986; Shaywitz et al., 1995). De

term Mattheuseffect is geïntroduceerd door Merton (1968) en had oorspronkelijk betrekking op het carrièreverloop van wetenschappelijk personeel. De term is ontleend aan de bijbel, om precies te zijn aan Mattheus 25, vers 29: "Want aan een ieder die heeft zal gegeven worden en hij zal overvloedig hebben. Maar wie niet heeft, ook wat hij heeft, zal hem ontnomen worden." Een artikel van Walberg en Tsai uit 1983 heeft de term brede bekendheid bezorgd binnen de onderwijskunde. Het onderzoek waarover zij rapporteren heeft echter betrekking op een beperkte interpretatie van het Mattheuseffect, namelijk de invloed van opleiding en sociaal-economische achtergrond op het kennisniveau van jonge volwassenen (tussen de 25 en 35 jaar). De ontwikkeling van het kennisniveau over een langere periode komt in hun analyse niet aan bod. In het theoretische kader met betrekking tot de verwerving van leesvaardigheid dat Stanovich (1986) schetst is juist wel een centrale plaats ingeruimd voor de ontwikkelingen op lange termijn. Daarbij zijn twee begrippen van cruciaal belang: wederzijdse beïnvloeding ('reciprocal relationships') en de invloed van de omgeving ('organism-environment correlation'). Een voorbeeld van het eerste is de relatie tussen woordherkenning en leesvaardigheid. Om goed te kunnen lezen moet een leerling de woorden in een tekst kunnen herkennen. Hoe beter leerlingen kunnen lezen, des te gemakkelijker kunnen ze de betekenis van een onbekend woord afleiden uit de context. Hierdoor verbetert hun woordkennis, waardoor het lezen gemakkelijker wordt, zodat ze beter de betekenis van onbekende woorden kunnen afleiden, enz. Een voorbeeld van de invloed van de omgeving is dat een leerkracht minder hoge eisen stelt aan minder goed presterende leerlingen, met als gevolg relatief weinig leer-voordingen.

Het onderzoek waarover in dit artikel wordt gerapporteerd heeft betrekking op de beperkte interpretatie van het Mattheuseffect. Dit wordt aanwezig geacht als er een samenhang is tussen sociaal-economische status en leerprestaties, ofwel als een leerling uit de hogere sociale strata ('hij die heeft') hogere prestaties behaalt ('zal gegeven worden') dan een leerling uit de lagere sociale strata ('hij die niet heeft'). Wanneer dan de rol van de school in dezen voorwerp van onderzoek is, dan wordt deze beperkte invulling ook vaak aangeduid met de term "differentiële schooleffecten". De benaming "Mattheuseffect" wordt veelal gereserveerd voor de uitgebreide interpretatie waarin ook het lange termijn perspectief is verwerkt.

Bosker (1990) laat zien dat in het Nederlandse onderwijs de achterstand van lager milieu leerlingen op hoger milieu leerlingen op zesjarige leeftijd ongeveer één jaar bedraagt, maar dat dit verschil op twaalfjarige leeftijd twee en op achttienjarige leeftijd zelfs drie jaar bedraagt. We zien dus dat de achterstand van lager milieu leerlingen in de loop van hun schoolloopbaan steeds groter wordt. Daarmee dringt zich de vraag op of sommige scholen systematisch grotere 'waaiers' hebben dan andere scholen: is het verschil in prestaties tussen lager en hoger milieu leerlingen iets dat consequent aanwezig is in de diverse jaargroepen en in de diverse kennisdomeinen? De eerste vraag zou men kunnen aanduiden als die naar de stabiliteit van differentiële schooleffecten over de leerjaren, en de tweede als de consistentie van deze effecten over de kennisdomeinen.

Het onderzoek naar differentiële schooleffecten is voor het schooleffectiviteitsonderzoek en het ongelijkheidsonderzoek niet zonder belang. De bewijskracht van de effectiviteitsbevorderende school- en instructiefactoren is niet overweldigend (zie bijvoorbeeld: Creemers & Osinga, 1995), en het lijkt erop dat dit a fortiori geldt voor de achterstandsleerlingen: factoren die specifiek voor deze groep leerlingen effectief zijn (zodat ze hun achterstand kunnen inlopen) worden zelden gevonden (bijv. Bosker, 1990; Brandsma, 1993; Mortimore et al., 1988; Weide, 1995).

De meest bekende onderwijs-effectiviteitsmodellen (Scheerens, 1990; Stringfield & Slavin, 1992; Creemers, 1994) geven alleen in algemene zin aan welke factoren een positieve uitwerking hebben op de leerprestaties. Welke factoren specifiek effectief zijn voor achterstandsleerlingen blijft buiten beschouwing. Het onderwijsleerproces in de klas wordt in deze modellen beschouwd als de basis voor onderwijs-effectiviteit. Binnen scholen kunnen wat dit betreft echter grote verschillen bestaan. De invloed van variabelen op schoolniveau en hoger

(bijvoorbeeld het onderwijsbeleid op plaatselijk of nationaal niveau) is indirect. Schoolfactoren kunnen de effectiviteit van het onderwijs in de klas stimuleren of belemmeren. Zo kan een gericht schoolbeleid ertoe bijdragen dat de vorderingen van de leerlingen regelmatig worden bijgehouden. De (plaatselijke) overheid kan dergelijke maatregelen op schoolniveau bevorderen door middel van financiële prikkels.

Kunnen we de stabiliteit van differentiële schooleffecten demonstreren dan impliceert dit dat naar schoolfactoren buiten de effectiviteitsmodellen van Scheerens (1990), Stringfield en Slavin (1992) en Creemers (1994) gezocht moet worden om de kansen van achterstandsleerlingen in het onderwijs te vergroten. Een mogelijk aanknopingspunt is Stringfields conceptie van scholen als 'high reliability organizations' (1995). Het belangrijkste kenmerk van dergelijke organisaties is het besef dat een falen rampzalige gevolgen kan hebben. Een 'high reliability organization' (bijvoorbeeld een kerncentrale of luchtverkeersleiding) vertoont enkele overeenkomsten met de klassieke bureaucratie (duidelijke hiërarchie en veelvuldig gebruik van standaard procedures), maar wordt daarnaast gekenmerkt door een sterk besef van haar primaire doelstelling. Procedures en gezagsverhoudingen worden zodoende nooit een doel op zich. Volgens Stringfield kan de samenleving het zich tegenwoordig niet meer veroorloven om aanzienlijke aantallen leerlingen onvoldoende toe te rusten. Een grote inspanning vroeg in de schoolloopbaan zal ruimschoots opwegen tegen de voordelen op de lange termijn. Stringfield verwijst in dit verband naar het "Success for All" programma, dat is ontworpen voor scholen met veel achterstandsleerlingen in de Amerikaanse binnensteden (Slavin, 1994; 1996). De basisgedachte achter het programma is dat elke leerling koste wat het kost moet slagen in de onderbouw van het primair onderwijs. Cruciaal is daarbij dat leervorderingen met grote regelmaat worden bijgehouden en dat leerlingen snel extra aandacht krijgen als ze dreigen achter te blijven.

Uit onderzoek naar de verhouding tussen "gewone" leerkracht- en schooleffecten komt het beeld naar voren dat leerkrachteffecten meer gewicht in de schaal leggen dan schooleffecten (Scheerens & Bosker, 1997). Met name de stabiliteit over leerjaren blijkt gering. Als een school uitstekende resultaten behaalt met de leerlingen in bijvoorbeeld jaargroep 6, dan zegt dit weinig over de resultaten in andere jaargroepen. Blijken ook de differentiële effecten instabiel te zijn dan moet kennelijk gezocht worden naar factoren op leerkrachtniveau die de ongelijkheid in onderwijskansen kunnen verklaren. De verwachtingen en streefdoelen van leerkrachten kunnen hierbij een belangrijke rol spelen. Gebleken is dat leerkrachten minder hoge verwachtingen hebben van en minder hoge eisen stellen aan leerlingen uit een lager milieu, ook bij een gelijke intelligentie (Van der Hoeven-Van Doornum, 1990). Overigens kunnen verwachtingen en streefdoelen nog steeds een rol spelen als de effecten stabiel blijken. Voorwaarde is dan wel dat er in dit opzicht binnen scholen grote overeenkomsten bestaan tussen de leerkrachten. De implementatie van een programma als "Success for all" vereist immers ook een consensus onder de leerkrachten binnen een school met betrekking tot verwachtingen en streefdoelen (cf. Van der Hoeven-Van Doornum, 1994).

Zijn de differentiële effecten bovendien niet erg consistent dan dient zelfs gezocht te worden naar domeinspecifieke nieuwe instructiefactoren, of eventueel naar domeinspecifieke leerkracht-leerling interacties.

2. CRITERIUMMATEN IN SCHOOLEFFECTIVITEITSONDERZOEK

In de praktijk van het schooleffectiviteitsonderzoek komt men, wanneer het blikveld gericht wordt op leerdoelen de cognitieve korte termijn betreffend (d.w.z. vastgesteld aan het eind van de onderzochte instructieperiode), vier verschillende operationele definities van het begrip schooleffect tegen:

1. het effect van de school is de gemiddelde ongecorrigeerde prestatiescore van de leerlingen. Het gaat met andere woorden om een *bruto prestatieniveau*. De legitimering van deze

- specifieke operationalisering kan men ontlenen aan de discussie over eindtermen c.q. een minimum beheersingsniveau. Hierbij gaat het om het door de leerlingen behaalde cognitieve peil ongeacht hun achtergrond respectievelijk initiële vaardigheden.
2. het effect van een school is de gemiddelde mate van 'overachievement' van de leerlingen. Bij deze operationalisering wordt een voorspelling gedaan van de leerprestaties van leerlingen. Deze voorspelling is gebaseerd op achtergrondkenmerken van de leerling als milieu van herkomst, sexe en nationaliteit en indicatoren voor de begaafdheid van de leerling (de facto I.Q.-testscores). De gemiddelde mate van afwijking van deze voorspelling vormt dan het schooleffect.
 3. het effect van de school is de gemiddelde leerwinst die leerlingen in een bepaalde tijdsperiode boeken. Hier staat de *bruto leerwinst* dus centraal. In de praktijk werkt men niet met verschilscores, vanwege hun intrinsieke onbetrouwbaarheid, maar met het verschil tussen de op basis van initiële prestaties voorspelde en de feitelijk geobserveerde prestatiescore. Het probleem bij deze operationalisering is dat er een zekere periode afgebakend moet worden. De keuze valt veelal op een schooljaar.
 4. het effect van een school op haar leerlingen is de gemiddelde leerwinst die deze leerlingen boeken, althans voor zover deze leerwinst niet terug te voeren is op ascriptieve leerlingkenmerken en de begaafdheid van de leerling. Het gaat met andere woorden om de *netto leerwinst* die de leerlingen gemiddeld boeken.

Elke definitie heeft zijn voors en tegens. Formeel gesproken is de enige juiste definitie van een schooleffect de leerwinst die leerlingen in acht jaren basisonderwijs boeken. Het construeren van instrumenten die het prestatieniveau bij entree in en bij het verlaten van de school kunnen vaststellen, is echter ten principale onmogelijk. Gewerkt wordt echter aan een benadering van dit principe middels het vaststellen van groeicurves van leerlingen met behulp van een volgens item respons theorie geconstrueerd leerlingvolgsysteem (Moelants et al., 1990; van den Bergh & Kuhlemeier, 1991; Voorbosker & Scheerens, 1995; Raudenbush, 1989; van den Bergh & Kuhlemeier, 1991). Vooralsnog zal in schooleffectiviteitsonderzoek echter in hoofdzaak gewerkt moeten worden met de vier genoemde maten. Bij onderzoek naar differentiële schooleffectiviteit (verschillende schooleffecten voor verschillende groepen leerlingen) ligt het voor de hand alleen te werken met de 'overachievement' en de netto leerwinst maat, aangezien er immers vanuit wordt gegaan dat een nader onderscheid naar subgroepen van leerlingen, in casu sociale strata, zinvol is.

Naast deze verschillende operationaliseringën moet bovendien nog het onderscheid naar kennisdomeinen worden gemaakt. Zowel in het primair als het secundair onderwijs bestaat er een voorkeur voor onderzoek naar effecten op het gebied van reken- respectievelijk wiskunde prestaties enerzijds en de (moeder)taalprestaties anderzijds. De effecten van basisscholen op rekenprestaties blijken over het algemeen twee keer zo groot te zijn als de effecten op taalprestaties. Voor rekenprestaties geldt kennelijk sterker dan voor taalprestaties dat het een vaardigheid is die leerlingen via het onderwijs verwerven, terwijl taalprestaties sterker door de thuissituatie worden beïnvloed. Voor taalvaardigheid geldt echter weer dat het voorwaardelijk is voor de vorderingen in de overige schoolvakken.

In het onderhavige onderzoek naar de consistentie en stabiliteit van differentiële schooleffecten zullen vier effectmaten naast elkaar gehanteerd worden: twee kennisdomeinen (rekenen en taal) en twee verschillende manieren van operationalisering ('overachievement' en 'netto leerwinst').

3. VRAAGSTELLINGEN

- Op grond van bovenstaande overwegingen formuleren we de volgende onderzoeksvragen:
- 1 a. In welke mate zijn differentiële effecten gesitueerd op school- dan wel leerjaarniveau?
 - b. Is de effectiviteit van een school hoger naarmate het differentiële effect kleiner is?
 - c. In hoeverre verschilt het antwoord op deze vragen voor reken- en taalprestaties en voor 'overachievement' en netto leerwinst?

Deze eerste vraag richt zich dus met name op binnenschoolse verschillen tussen leerlingen met een verschillende sociaal-economische achtergrond, waarbij de b-vraag dan betrekking heeft op de geuite veronderstelling dat die verschillen kleiner zullen zijn als in het algemeen de leerlingen op een school goede prestaties behalen. De c-vraag vloeit voort uit hetgeen hiervoor gezegd is over de diverse operationalisering van het begrip 'schooleffect', en vanwege het empirisch gegeven dat het effect van scholen op rekenprestaties aanmerkelijk groter is dan op taalprestaties en dus mogelijk ook het differentiële effect.

2 a. *In hoeverre komen jaargroepen van een zelfde basisschool overeen in hun effectiviteit voor laag, midden en hoog milieu leerlingen?*

b. *In hoeverre verschilt het antwoord op deze vraag voor reken- en taalprestaties en voor 'overachievement' en netto leerwinst?*

Deze tweede vraag richt zich niet op binnenschoolse (relatieve) verschillen tussen leerlingen uit de diverse sociale lagen, maar op prestatieniveaus per sociale laag. De vraag spitst zich hier dus toe op het effect van scholen voor bepaalde sociale strata en de mate waarin dit stabiel is over de leerjaren heen.

3 a. *In hoeverre komen binnen een school de reken- en taalprestaties overeen van laag, midden en hoog milieu leerlingen?*

b. *In hoeverre verschilt het antwoord op deze vraag voor leerjaar 6 en 8 en voor 'overachievement' en netto leerwinst?*

De laatste vraag tenslotte richt zich op de samenhang tussen de vakken taal en rekenen voor wat betreft de effectiviteit van een school voor leerlingen uit een bepaalde sociale laag. Ook nu gaat het dus niet meer om relatieve verschillen binnen scholen tussen de leerlingen uit de diverse sociaal-economische milieus, maar meer om de vraag of scholen die het wat betreft rekenen goed doen voor bijvoorbeeld de leerlingen uit de lagere sociaal-economische milieus, dat ook doen voor taal.

4. DE ONDERZOEKSOPZET

4.1 Steekproef en variabelen

Een a-selecte steekproef van 212 scholen van alle Nederlandse scholen voor basisonderwijs vormt de onderzoeksgroep. Leerlingen uit de jaargroepen 6 en 8 van deze scholen kregen een voor- (aan het eind van groep 5 respectievelijk groep 7) en een jaar later een natoets (aan het eind van groep 6 respectievelijk groep 8) voor taal en rekenen. Bovendien werd van deze leerlingen bepaald:

- de verbale en non-verbale intelligentie
- het geslacht
- de sociaal-economische status
- de nationaliteit

Deze kenmerken hebben samen met de scores op de voortoets als controle-variabelen dienst gedaan. De gegevens stammen uit het schooljaar 1987-1988. Voor het meten van de taal- en rekenprestaties zijn (onderdelen van) bestaande CITO-toetsen afgenomen. Voor het meten van taalprestaties zijn toetsonderdelen afgenomen over taalgebruik, begrijpend lezen en spelling. De spellingstoets voor de leerlingen in groep 5/6 is niet afkomstig uit een CITO-toets maar uit Praxis 35 (Van den Heuvel, 1983). De verbale en performale intelligentie is in groep 7 gemeten door subtests van de ISI (Interesse, Schoolvorderingen, Intelligentie) af te nemen (Van Boxtel, Snijders & Welten, 1982). De performale intelligentie betreft onderdelen die de ruimtelijke factor van het intelligentiedeel van deze test meten, namelijk "geknipte figuren", "draaien" en "soortbegrip figuren". Het verbale deel betreft "synoniemen", "soortbegrip woorden" en "tegenstellingen". In groep 5 is de verbale en performale intelligentie gemeten door het afnemen van verwante subtests van de PMA 2-4 (Kema, 1980). Met betrekking tot de

sociaal-economische status en nationaliteit merken we op dat een indirecte meting heeft plaats gevonden. De leerkrachten is gevraagd om voor elke leerling een inschatting te maken van het beroeps- en opleidingsniveau van de ouders (in vijf categorieën). Daarnaast is hen gevraagd of een leerling de Nederlandse nationaliteit bezit. Dat bleek voor 95 procent van de leerlingen in deze onderzoeksgroep het geval te zijn. De steekproef kan representatief geacht worden voor Nederland, ondanks een kleine ondervertegenwoordiging van scholen uit de regio Utrecht, Nijmegen, Veluwe, en van Protestants Christelijke scholen (zie Brandsma, 1993, p. 56; Knuver, 1993, pp. 58-63).

In deze groep bleken er 150 scholen aanwezig waarvan van meerdere leerlingen uit de jaargroepen 6 en 8 alle gegevens bekend waren wat betreft de scores op de natoets voor taal en rekenen alsmede voor de hiervoor genoemde controle-variabelen. Op deze groep zal de vraag naar de consistentie van differentiële schooleffecten over jaargroepen heen onderzocht worden. Alle variabelen zijn per jaargroep getransformeerd naar z-scores (met gemiddelde 0 en variantie 1). De reden dat zodoende een niet gering deel van het gegevensbestand ongebruikt is gebleven is een gevolg van het feit dat de beschikbare programmatuur om multiniveau-analyses uit te voeren geen mogelijkheden biedt om berekeningen uit te voeren op gegevens met ontbrekende waarden.

4.2 Een drie-niveau model voor onderzoek naar consistentie en stabiliteit

Uitgangspunt voor de statistische analyse is, dat er een drietraps steekproef is getrokken (scholen, daarbinnen klassen en daarbinnen leerlingen). Weliswaar zijn de tweede en derde trap (leerjaren respectievelijk leerlingen) niet echt a-select getrokken, maar dit kan dan als een steekproef in de tijd beschouwd worden: het onderzoek had ook twee jaar later uitgevoerd kunnen worden bij andere lichtingen en/of bij andere jaargroepen. Als gevolg van de drietraps steekproeftrekking zijn de onderzoekseenheden statistisch gezien niet onafhankelijk van elkaar; er is sprake van een natuurlijke clustering van de data (c.f. Aitkin et al., 1981). Leerlingen die in dezelfde klas zitten lijken op elkaar, aangezien ze hun leerkracht gemeenschappelijk hebben; leerlingen die op dezelfde school zitten maar in verschillende jaargroepen lijken op elkaar omdat ze hun schoolomgeving delen. Juist deze afhankelijkheid van de waarnemingen vormt de basis voor het te hanteren statistisch model (c.f. Aitkin & Longford, 1986; Goldstein, 1987; Raudenbush & Bryk, 1986). Dit model kan als volgt uitgeschreven worden, waarbij we direct de relevante predictoren opnemen (een model zonder predictoren wordt kortweg aangeduid met 'nulmodel', en geeft een eerste inzicht in de verdeling van de variantie over de diverse niveaus):

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk} \text{milieu}_{ijk} + \beta_{2} \text{sexe}_{ijk} + \beta_{3} \text{nationaliteit}_{ijk} + \beta_{4} \text{IQverb}_{ijk} + \beta_{5} \text{IQperf}_{ijk} + \beta_{6} \text{voortoets}_{ijk} + R_{ijk} \quad (1)$$

Hierin is

- Y_{ijk} : de toetsscore van leerling i in groep j van school k
- β_{0jk} : het groepspecifieke intercept
- β_{1jk} : de groepspecifieke coëfficiënt voor de regressie van y op milieu
- milieu_{ijk} : het sociaal-economisch milieu van herkomst van leerling i in groep j van school k
- $\beta_2, \beta_3, \text{etc.}$: de coëfficiënt voor de regressie van Y op sexe , nationaliteit , etc.
- sexe_{ijk} : het geslacht van leerling i in groep j in school k

- nationaliteit_{ijk}* : de nationaliteit van leerling *i* in groep *j* in school *k*
IQverb_{ijk} : de score op het verbale deel van de ISI IQ-test van leerling *i* in groep *j* in school *k*
IQperf_{ijk} : de score op het non-verbale deel van de ISI IQ-test van leerling *i* in groep *j* in school *k*
voortoets_{ijk} : de score op de voortoets van leerling *i* in groep *j* van school *k*
R_{ijk} : het residu op leerlingniveau met variantie σ^2

Voorts geldt voor de groepspecifieke intercepten en regressiecoëfficiënten:

$$\beta_{0jk} = \gamma_{00k} + U_{0jk} \quad (2a)$$

$$\beta_{1jk} = \gamma_{10k} + U_{1jk} \quad (2b)$$

waarbij

γ_{00k} : het schoolspecifieke intercept

γ_{10k} : de schoolspecifieke regressie van *Y* op milieu

U_{0jk} : het residu op groepsniveau met variantie τ_{00}^2

U_{1jk} : de groepspecifieke afwijking van de schoolspecifieke regressie van *Y* op milieu met variantie τ_{11}^2

De covariantie tussen U_{0jk} en U_{1jk} (τ_{01}) is in het algemeen ongelijk aan 0. Als deze covariantie positief is, betekent dat, dat groepen waar leerlingen gemiddeld relatief hoge leerprestaties behalen ook een relatief grote afstand tussen lager en hoger milieu leerlingen hebben in hun uiteindelijke leerprestaties. Een negatieve covariantie duidt op het samengaan van kwaliteit met relatief minder ongelijkheid tussen hoger en lager milieu leerlingen.

Het derde niveau van het statistisch model kan als volgt uitgeschreven worden

$$\gamma_{00k} = \delta_{000} + V_{00k} \quad (3a)$$

$$\gamma_{10k} = \delta_{100} + V_{10k} \quad (3b)$$

In (3a en 3b) worden ook de schoolspecifieke intercepten en regressies opgesplitst, en wel in een populatiedeel (δ_{000} respectievelijk δ_{100}) en schoolspecifieke afwijkingen (V_{00k} respectievelijk V_{10k}). Deze afwijkingen hebben variantie φ_{00}^2 respectievelijk φ_{11}^2 . Net als op het groepsniveau het geval is, kunnen V_{00k} en V_{10k} onderling covariëren (met φ_{01}). De geschatte variantiecomponenten op het tweede en derde niveau leveren de basis voor het antwoord op de eerste vraag. De relatieve omvang van φ_{00}^2 (relatief ten opzicht van de som $\varphi_{11}^2 + \tau_{11}^2$) geeft aan of het differentiële effect op schoolniveau is gelegen (φ_{11}^2 is verhoudingsgewijs groot) of op leerjaarniveau (φ_{11}^2 is verhoudingsgewijs klein). De geschatte covariantiecomponent φ_{01} kan gebruikt worden om de correlatie te schatten tussen het differentiële schooleffect en het gemiddelde effect van een school.

Om nu na te kunnen gaan of de schooleffecten voor lager milieu leerlingen in jaargroep 6 samenhangen met die voor dezelfde groep leerlingen in jaargroep 8, worden Empirisch Bayesiaanse schattingsprincipes geïntroduceerd. Deze stap is voor de analyses een gewichtige.

Zowel voor de schatting van β_{0jk} als γ_{00k} alsmede voor de schatting van zowel β_{1jk} als γ_{10k} wordt gebruik gemaakt van alle aanwezige informatie. Het principe hierbij is, dat alle jaargroepen en alle scholen uit dezelfde populatie afkomstig zijn. Derhalve kunnen we de schattingen voor de ene klas of de ene school optimaliseren door gebruik te maken van de gegevens van de andere klassen of de andere scholen. In het algemeen geldt voor een klas j

$$\beta_j^* = w_j \beta_j + (1 - w_j) \beta^* \quad (4)$$

In (4) wordt β als nieuwe term geïntroduceerd. Hiermee kan zowel β_j als β_0 worden aangeduid. Mutatis mutandis geldt hetzelfde voor alle parameters in (4). β^* kan dus zowel betrekking hebben op de regressie van Y op het milieu van herkomst als op het intercept. De gewichten w_j worden gedefinieerd als

$$w_j = \tau_{00}^2 / (\tau_{00}^2 + \sigma^2 / \sum_i X_i) \quad (5)$$

Wanneer (5) wordt toegepast op de klasspecifieke intercepten hebben alle leerlingen op variabele X_i de score 1 (dit wordt dus een constante; $\sum_i X_i$ wordt in dat geval dus n_j , ofwel het aantal leerlingen in klas j). Passen we (5) toe op de regressiecoëfficiënten, dan staat X_i voor het milieu van herkomst van leerling i . Uit (4) en (5) is te lezen dat regressiecoëfficiënten en interceptwaarden voor een bepaalde klas meer door het algemeen gemiddelde van de coëfficiënten worden bepaald naarmate de informatie van deze klas onbetrouwbaarder is (i.e. weinig leerlingen en/of weinig variatie in het milieu van herkomst). Naar analogie met het bovenstaande kunnen we ook schoolspecifieke interceptwaarden en regressiecoëfficiënten schatten.

Hoe kan nu bepaald worden of school-leerling interacties stabiel zijn over de leerjaren en consistent over de vakgebieden heen? Met andere woorden: hoe weten we nu of de schooleffecten voor lager milieu leerlingen uit jaargroep 6 overeenstemmen met die voor de dezelfde groep leerlingen uit jaargroep 8? Daartoe passen we (1) tot en met (5) toe. Vervolgens wordt voor drie denkbeeldige leerlingen (-1.28 op de schaal voor milieu van herkomst, d.w.z. op het tiende percentiel; 0 op deze schaal, d.w.z. op het vijftigste percentiel; +1.28 op deze schaal, d.w.z. op het negentigste percentiel) de prestatie voorspeld. Ofwel voor jaargroep 6:

$$Y_{i6k} = (\beta_{06k} + \gamma_{00k} + \delta_{000}) + (\beta_{16k} + \gamma_{10k} + \delta_{100}) \text{milieu}_{i6k} \quad (6)$$

waarin de variable milieu achtereenvolgens de waarden -1.28, 0 en +1.28 aanneemt.

Hetzelfde doen we vervolgens voor de leerlingen van jaargroep 8. We kunnen nu de correlaties schatten tussen de voorspelde prestatiescores voor twee lager milieu leerlingen, respectievelijk twee midden milieu leerlingen, respectievelijk twee hoger milieu leerlingen van dezelfde school. Per leerjaar kan eveneens de correlatie tussen de voorspelde scores voor rekenen en taal per groep leerlingen worden geschat. Deze geschatte correlaties vormen de basis om een antwoord te kunnen formuleren op de tweede en derde onderzoeksvraag.

5. RESULTATEN

5.1 Differentiële effecten op school- en leerjarniveau

Alvorens de eerste onderzoeksvraag te beantwoorden presenteren we in tabel 1 de nulmodellen voor rekenen en taal. De nulmodellen laten zien hoeveel variatie in de bruto leerprestaties voor rekenen en taal is gesitueerd op respectievelijk het leerling-, leerjaar- en schoolniveau, waarbij het begrip 'bruto' erop wijst dat er geen predictoren in het regressiemodel zijn opgenomen. Uit tabel 1 valt op te maken dat met name op het leerjarniveau de uitkomsten voor rekenen en taal verschillend zijn. Voor rekenen is de leerjaarvariantie bijna twee maal

zo groot als voor taal. Voor beide vakken geldt dat de variantie op schoolniveau groter is dan die op leerjaarniveau.

Tabel 1: Nulmodellen - variantiecomponenten voor rekenen en taal

	Rekenen	Taal
Variantie op leerlingniveau	.751	.835
Variantie op leerjaarniveau	.109	.056
Variantie op schoolniveau	.148	.125

150 scholen, 300 leerjaren (per school, groep 6 en 8), 4658 leerlingen

De uitkomsten van de analyses die een antwoord geven op de vraag naar de situering van differentiële effecten op school- en leerjaarniveau staan vermeld in tabel 2. Helemaal onderaan in de tabel staan de hoeveelheden verklaarde variantie vermeld. Daarbij dient te worden bedacht dat ze betrekking hebben op multiniveau modellen met 'random slopes'. Het effect van milieu op leerprestaties verschilt tussen scholen en binnen scholen weer tussen leerjaren. De hoeveelheid verklaarde variantie verschilt zodoende eveneens tussen scholen en leerjaren. De vermelde hoeveelheid verklaarde variantie geldt voor de leerlingen met een gemiddelde score op de variabele milieu¹.

Tabel 2: Regressiegewichten en geschatte varianties voor reken- en taalprestaties

	Rekenen		Taal	
	over- achievement	netto leerwinst	over- achievement	netto leerwinst
<i>Predictoren:</i>				
Sexe	-.085	-.080	.113	.048
Nationaliteit				
Milieu	.151	.103	.181	.100
IQ verbaal	.237	.152	.415	.164
IQ nonverbaal	.303	.179	.184	.072
Voortoets	---	.397	---	.576
<i>Variante-componenten:</i>				
Leerlingniveau	.515	.425	.537	.349
Leerjaarniveau	.068	.058	.036	.037
Schoolniveau	.093	.055	.184	.021
<i>Random effecten leerjaarniveau</i>				
Milieu	.00817	.00576	.00010	.00034
<i>Covariantie leerjaarniveau</i>				
Milieu*intercept	-.00766	-.00646	-.00187	-.00355
<i>Random effecten schoolniveau</i>				
Milieu	.00152	.00036	.00100	.00058
<i>Covariantie schoolniveau</i>				
Milieu*intercept	-.00086	.00166	.00017	.00138
<i>Verklaarde variantie</i>	33%	47%	26%	61%

150 scholen, 300 leerjaren (per school, groep 6 en 8), 4658 leerlingen

In de tabel staan alleen effecten vermeld die significant zijn voor ($\alpha < .05$ bij een tweezijdige toetsing). Met uitzondering van nationaliteit is voor alle controle-variabelen zowel bij taal als rekenen een significant effect gevonden. Als men controleert voor intelligentie en sociaal-economische status is het verschil tussen leerlingen met een Nederlandse nationaliteit en die met een andere nationaliteit niet significant. Deze variabele is in de multiniveau modellen verder buiten beschouwing gebleven. De coëfficiënten die achter de controle-variabelen

vermeld staan kan men lezen als een gewone regressiecoëfficiënt. Het milieu-effect varieert van 0.10 (netto leerwinst taal) tot 0.18 ('overachievement' met betrekking tot taal). Die 0.10, bijvoorbeeld, wijst erop dat het geschatte verschil tussen een hoger en lager milieu leerling 0.26 ($=0.10 \cdot (2 \cdot 1.28)$) bedraagt waarbij verschillen in intelligentie en sexe reeds zijn uitgepartialiseerd. Het schooleffect nu, kunnen we interpreteren als het effect van de school op de prestaties van de leerling, althans voor zover verschillen in prestaties niet meer zijn toe te schrijven aan de scores op de controle-variabelen. Onder de rubriek 'variantiecomponenten' staat aangegeven hoe de residuele variantie in prestaties verdeeld is over de drie niveaus: leerling, leerjaar en school. Met name relevant voor onze vraagstelling zijn de volgende vier rubrieken uit de tabel. Onder de rubrieken 'random effecten' staat vermeld in welke mate de regressie van prestatie op het milieu van herkomst van leerjaar tot leerjaar respectievelijk van school tot school verschilt. De interpretatie van het random effect op schoolniveau in de kolom 'rekenen, overachievement' is als volgt: het effect van milieu van herkomst op de rekenprestaties (althans na uitzuivering van sexe- en I.Q.-effecten op rekenprestaties) heeft een variantie van .00152 tussen scholen (zie tabel 2). Ofwel, de standaarddeviatie voor deze regressiecoëfficiënt is $\sqrt{.00152} = .039$. Met andere woorden, de tien procent scholen met de grootste ongelijkheid tussen lager en hoger milieu leerlingen hebben een regressiecoëfficiënt van $.151 + (1.28 \cdot .039) = .201$ of groter (hetgeen correspondeert met een geschat verschil van 0.51 of meer tussen leerlingen uit een hoog en laag milieu van herkomst), en de tien procent scholen met de minste ongelijkheid hebben een regressiecoëfficiënt van $.151 - (1.28 \cdot .039) = .101$ of kleiner (en hierbij is het geschatte verschil tussen leerlingen uit een hoog en laag milieu van herkomst 0.26 of minder). Onder de rubrieken 'covarianties' staan de parameterschattingen vermeld van de covariantie tussen deze leerjaar- respectievelijk schoolspecifieke regressiecoëfficiënt voor het milieu van herkomst en het schooleffect voor de gemiddelde leerling. Zo zien we dat de regressie van rekenen op milieu van herkomst negatief samenhangt ($-.00086$) met het gemiddelde schooleffect. Ofwel: op scholen met een kleine afstand tussen lager en hoger milieu leerlingen (d.w.z. met een kleine regressiecoëfficiënt) worden hogere rekenprestaties behaald dan op scholen met een grote afstand tussen deze twee groepen leerlingen, althans indien men niet controleert voor de score op de voortoets. Men kan deze covariantie ook uitdrukken in termen van een correlatie. Deze bedraagt dan $-.00086 / (\sqrt{.093} \cdot \sqrt{.00152}) = -.07$.

Een eerste belangrijke conclusie met betrekking tot de consistentievragen is, dat de regressie van rekenprestaties op het milieu van herkomst sterk tussen leerjaren verschilt, en relatief weinig tussen scholen. Voor 'overachievement' rekenen ligt de verhouding school:leerjaar op 1:5 ($.00152 / .00817$). Voor netto leerwinst rekenen ligt de verhouding nog veel hoger, namelijk op 1:16. Voor taal daarentegen is het geheel andersom: hier ligt voor 'overachievement' de verhouding school:leerjaar op 10:1 en voor netto leerwinst op 2:1. Met andere woorden: voor rekenen verschilt de afstand tussen hoger en lager milieu leerlingen op een zekere school sterk tussen leerjaren, terwijl voor taal de afstand tussen hoger en lager milieu leerlingen op een zekere school voor de jaargroepen 6 en 8 weinig verschilt.

Een tweede belangrijke conclusie is dat er op leerjaarniveau voor taal en rekenen zowel voor overachievement als voor netto leerwinst een negatieve covariantie is gevonden tussen de leerjaarspecifieke regressiecoëfficiënt voor het milieu van herkomst en het leerjaareffect voor de gemiddelde leerling. Binnen scholen worden dus de beste resultaten behaald in die jaargroepen waar het effect van milieu beperkt is. Op schoolniveau wijzen de uitkomsten van de analyses eerder in de tegengestelde richting. Drie van de vier covarianties zijn positief. Dit is met name een belangrijk gegeven voor de taalprestaties, omdat daar de random effecten op schoolniveau groter zijn dan die op het leerjaarniveau.

5.2 Stabiliteit en consistentie van leerprestaties voor laag, midden en hoog milieu leerlingen

Aangezien de klas- en schoolspecifieke regressiecoëfficiënten correleren met het gemiddeld schooleffect, berekenen we de correlatie tussen de per jaargroep geschatte schooleffecten voor

lager milieu leerlingen. Dit herhalen we vervolgens voor de leerlingen uit het midden milieu, en tenslotte ook nog eens voor de leerlingen uit het hoger milieu. De correlaties kunnen afgelezen worden uit tabel 3.

Tabel 3: *Correlaties overachievment en netto leerwinst uitgesplitst naar milieu*

	Milieu					
	Laag		Midden		Hoog	
	'over- -achievment'	netto leerwinst	'over- -achievment'	netto leerwinst	'over- -achievment'	netto leerwinst
Taal in groep 6 en 8	.99	.98	.65	.47	.68	.56
Rekenen in groep 6 en 8	.98	.98	.67	.58	.72	.65
Taal en rekenen in groep 6	.64	.51	.66	.53	.66	.54
Taal en rekenen in groep 8	.67	.56	.77	.71	.78	.72

150 scholen, 300 leerjaren (per school, groep 6 en 8), 4658 leerlingen

De correlaties die hierboven vermeld staan kan men als volgt interpreteren: er is een rangordening van alle 150 scholen gemaakt voor de rekenprestaties van lager milieu leerlingen in jaargroep 6. Hetzelfde is vervolgens gebeurd voor jaargroep 8. We hebben dus als het ware een dubbele rangordening van de 150 scholen: eerst voor jaargroep 6 en vervolgens voor jaargroep 8. De weergegeven correlaties zijn de rangordecorrelaties tussen deze twee rangordes. Deze geven dus het antwoord op de vraag of een school die het relatief 'goed' doet met betrekking tot de taalprestaties van de lager milieu leerlingen in jaargroep 6 (althans voorzover verschillen in deze rekenprestaties niet meer zijn toe te schrijven aan de controle-variabelen) het eveneens 'goed' doet met betrekking tot de rekenprestaties van lager milieu leerlingen uit jaargroep 8. Voor lager milieu leerlingen blijken de schooleffecten zowel voor rekenen als taal (met correlaties van .98 en .99), uitermate consistent te zijn. Anders gezegd: scholen die het 'goed' doen voor lager milieu leerlingen in jaargroep 6, doen het ook 'goed' voor lager milieu leerlingen in jaargroep 8. Voor hoger en midden milieu leerlingen liggen deze correlaties aanmerkelijk lager. Dat jaargroepen van eenzelfde school dus niet altijd even effectief zijn, vloeit dus voort uit de differentiele effectiviteit voor de midden en hoger milieu leerlingen.

Het antwoord op de vraag of de schooleffecten voor het vakgebied taal en rekenen samenhangen onderscheiden naar de drie milieugroepen kan eveneens ontleend worden aan de correlaties vermeld in tabel 3. We zien dan dat de correlaties tussen de geschatte schooleffecten voor overachievment rekenen en taal in jaargroep 6 voor alle drie de milieugroepen ongeveer dezelfde sterkte hebben (.64 voor lager milieu leerlingen, .66 voor midden zowel als voor hoger milieu leerlingen). Voor netto leerwinst zijn de correlaties wat lager, maar ook hier zijn er nauwelijks verschillen tussen de drie milieu-groepen (respectievelijk .51, .53 en .54). In jaargroep 8 zijn de correlaties voor de midden en hoger milieu leerlingen wat sterker dan die voor de lager milieu leerlingen. Met betrekking tot overachievment ligt de correlatie voor de lager milieu leerlingen op .67 en voor de midden en hoger milieu leerlingen op respectievelijk .77 en .78. Voor netto leerwinst is het verschil tussen de lager milieu leerlingen en de overige leerlingen nog wat groter (.56 versus .71 en .72).

6. CONCLUSIES EN NABESCHOUWING

In deze bijdrage stond de vraag centraal in hoeverre differentiële effecten, in casu milieuspecifieke schooleffecten, stabiel zijn over de leerjaren heen, en consistent over de kennisdomeinen. Bovendien zijn twee typen schooleffect-maten gehanteerd. De eerste is een 'overachievement' maat, waarin wordt uitgedrukt in hoeverre prestatieverschillen tussen leerlingen, althans voor zover niet te herleiden op achtergrondkenmerken en algemene intellectuele vaardigheden, 'veroorzaakt' worden door de specifieke school die ze bezoeken, en/of door de specifieke jaargroep binnen die school. In de netto leerwinst maat zijn bovendien effecten van initiële op uiteindelijke prestatieverschillen uitgezuiverd. Onze uitkomsten kunnen als volgt worden samengevat:

1. Voor de beantwoording van de onderzoeksvragen maakt het weinig uit of men 'overachievement' dan wel netto leerwinst hanteert als effectmaat.
2. Voor rekenen zijn de differentiële effecten vooral gesitueerd op leerjaarniveau, voor taal vooral op schoolniveau.
3. Op leerjaarniveau is de effectiviteit hoger naarmate het differentiële effect kleiner is, op schoolniveau is het in drie van de vier gevallen andersom (uitzondering is 'overachievement' rekenen).
4. De stabiliteit van leerprestaties over leerjaren is zeer sterk voor lager milieu leerlingen maar veel minder voor midden en hoger milieu leerlingen.
5. In groep 8 zijn leerprestaties van de lager milieu leerlingen minder consistent tussen vakken dan de prestaties van de midden en hoger milieu leerlingen.

Bij deze conclusies dient te worden aangetekend dat de indeling naar milieu in dit onderzoek is gebaseerd op het oordeel van de leerkracht. Het kan niet worden uitgesloten dat dit de onderzoeksuitkomsten enigszins vertekent, ook al is uit eerder onderzoek gebleken dat leerkrachten in de meeste gevallen het opleidingsniveau van beide ouders correct inschatten (Van der Hoeven-Van Doornum (1990) rapporteert dat leerkrachten in 60 tot 70% van de gevallen het ouderlijk opleidingsniveau juist inschatten). Wellicht wordt het opleidings- en beroepsniveau van allochtone ouders relatief vaak te laag ingeschat.

Een tweede kanttekening betreft de onderzoeksgroep. De hier geanalyseerde gegevens hebben betrekking op de situatie in 1989. Herhaald onderzoek met gebruik making van meer recentere gegevens zal moeten uitwijzen of de hier vermelde bevindingen ook later nog repliceerbaar zijn.

Een opvallende uitkomst is dat op leerjaarniveau de differentiële effecten steeds samengaan met een hoge effectiviteit. De meest effectieve leerkrachten zijn kennelijk die leerkrachten bij wie het effect van milieu op de leerprestaties beperkt blijft. In tegenstelling tot rekenen geldt voor het vakgebied taal dat dit effect op leerjaarniveau teniet wordt gedaan door het random effect op schoolniveau. Al met al geldt voor taal dat de effectiviteit het hoogst is op scholen met een relatief sterk effect van milieu op de leerprestaties, terwijl dat voor rekenen juist niet het geval is. Een aannemelijke verklaring is dat rekenen meer dan taal een schoolse vaardigheid is. Veel meer dan rekenvaardigheden worden taalvaardigheden ook buiten school aangeleerd. De mogelijkheid voor scholen en leerkrachten om de achterstand van lager milieu leerlingen te beperken is voor het vakgebied taal derhalve kleiner dan voor rekenen.

Het meest opmerkelijke resultaat van onze analyses is de bijna perfecte stabiliteit over leerjaren van taal- en rekenprestaties voor de lager milieu leerlingen. Als de lager milieu leerlingen op een school het goed doen in groep 6, dan doen ze het ook goed in groep 8. Doen ze het slecht in groep 6, dan doen ze het ook slecht in groep 8. De instabiliteit over leerjaren komt nagenoeg geheel voor rekening van de midden en hoger milieu leerlingen.

Voor het vakgebied rekenen blijkt dat een zwak effect van milieu samengaat met een hoge effectiviteit, maar voor taal komt het er juist op neer dat een zwak effect van milieu samengaat met een lage effectiviteit.

Tenslotte is het opmerkelijk dat in groep 8 de leerprestaties van de lager milieu leerlingen minder consistent zijn over vakgebieden dan de prestaties van de midden en hoger milieu leerlingen. De auteurs zijn niet bekend met andere onderzoeksuitkomsten die betrekking hebben op verschillen in consistentie tussen lager en hoger milieu leerlingen. Wel is een geringere consistentie in leerprestaties eerder al eens naar voren gebracht als een mogelijke verklaring voor het gegeven dat leerkrachten minder hoge verwachtingen hebben van leerlingen uit een lager milieu (Luyten & Snijders, 1996; p. 17). Overigens is het verschil in consistentie alleen gevonden in groep 8 en bovendien is het aanzienlijk geringer dan het verschil in stabiliteit over leerjaren.

Het onderzoek naar schooleffectiviteit heeft zijn oorsprong gevonden in het ongelijke kansen onderzoek. Alhoewel de teneur van het effectiviteitsonderzoek lijkt dat we wel weten wat scholen in het algemeen effectief maakt, maar niet zozeer specifiek voor lager milieu leerlingen, is het des te opvallender dat scholen consequent meer of juist minder effectief zijn voor leerlingen uit de lagere sociale strata. Deze bevinding leidt ertoe dat nader doordacht dient te worden welke schoolfactoren hier werkzaam zijn. Voor het onderzoek naar ongelijke onderwijskansen is de vermelde bevinding van belang in die zin, dat een schoolfactor achter dit fenomeen schuil moet gaan. Deze bevinding lijkt op het eerste gezicht wellicht strijdig met de uitkomsten van onderzoek waaruit blijkt dat leerlingbeelden en streefdoelen van leerkrachten de mindere prestaties van leerlingen uit lagere milieus gedeeltelijk verklaren (Van der Hoeven-Van Doornum, 1994). Het meest waarschijnlijk is echter dat binnen scholen de overeenkomsten tussen leerkrachten met betrekking tot leerlingbeelden en streefdoelen erg groot zijn. Mogelijk vindt er juist voor de lager milieu leerlingen een transmissie plaats vindt van verwachtingen en streefniveaus voor juist deze groep leerlingen van toeleverende (bijvoorbeeld de groep 7 leerkracht) naar ontvangende leerkrachten (in dit voorbeeld de groep 8 leerkracht). Dit zou in vervolgonderzoek nadere aandacht verdienen. Waarschijnlijk bestaan er met betrekking tot het geloof in de mogelijkheden van lager milieu leerlingen aanzienlijke verschillen tussen scholen. Van de ene kant scholen die het predikaat "high reliability organization" verdienen. Deze worden in de allereerste plaats gekenmerkt door een sterk besef onder de leerkrachten dat alle leerlingen in voldoende mate kunnen en moeten worden toegerust. Van de andere kant scholen waar onder de leerkrachten weinig vertrouwen bestaat in de mogelijkheden op succes voor de lager milieu leerlingen. De goede resultaten die met Slavin's "Success for all" programma (Slavin, 1994; 1996) worden bereikt tonen aan dat juist scholen met kenmerken van "high reliability organizations" er in slagen om de verschillen in leerprestaties tussen leerlingen uit verschillende milieugroepen te beperken.

LITERATUUR

- Aitkin, M., Anderson, D., & Hinde, J. (1981). Statistical Modelling of Data on Teaching Styles. *The Journal of the Royal Statistical Society, Series A (General)* 144, 419-461.
- Aitkin, M., & Longford, N. (1986). Statistical Modelling Issues in School Effectiveness Studies. *The Journal of the Royal Statistical Society, Series A (General)* 149, Part 1, 1-43.
- Bergh, H. van den, & Kuhlemeier, H. (1991). *On the stability of school effectiveness and effects of social economic background (Occasional paper presented at ETS, Princeton)*. Utrecht/Arnhem.
- Bosker, R.J. (1990). *Extra kansen dankzij de school?* (dissertatie) Nijmegen: ITS/OoMO.
- Bosker, R.J. (1990). Theory development in schooleffectiveness research: in search for stability of effects. In P. van den Eeden, J. Hox & J. Hauer (Eds.). *Theory and model in multilevel research: convergence or divergence?* Amsterdam: SISWO.
- Bosker, R.J., & Scheerens, J. (1995). A self-evaluation procedure for schools using multilevel modelling. *Tijdschrift voor Onderwijsresearch*, 20, 2, 154-164.
- Boxtel, H.W. van, Snijders, J. Th. & Welten, V.J. (1982). *ISI: Interesse, Schoolvorderingen, Intelligentie*. Groningen: Wolters-Noordhoff.
- Brandsma, H.P. (1993). *Basisschoolkenmerken en de kwaliteit van het onderwijs*. Groningen: RION.

- Coleman, J.S. (1990). *Equality and achievement in education*. Boulder/San Fransico/London: Westview Press.
- Creemers, B.P.M. (1994). *The effective classroom*. London: Cassell.
- Creemers, B.P.M., & Osinga, N.J. (1995). *ICSEI Country Reports*. Leeuwarden: GCO.
- Creemers, B.P.M., & Scheerens, J. (1994). Developments in the educational effectiveness research programme. *International Journal of Educational Research*, 21, 2, 125-138.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. Londen: Charles Griffin & Co.
- Hoeven-van Doornum, A.A. van der (1990). *Effecten van leerlingbeelden en streefniveaus op schoolloopbanen*. (dissertatie) Nijmegen: ITS/OOMO.
- Kema, G.N. (1980). *PMA 2-4, Nederlandse bewerking*. Amsterdam: Swets & Zeitlinger.
- Knuver, J.W.M. *De relatie tussen klas- en schoolkenmerken en het affectief functioneren van leerlingen*. Groningen: RION
- Kreft, G.G., & Leeuw, J. de, (1991). Model based ranking of schools. *International Journal of Educational Research*, 15, 1, 45-60.
- Luyten, H., & Sijnders, T.A.B. (1996). School effects and teacher effects in Dutch elementary education. *Educational Research and Evaluation*, 2, 1, 1-24.
- Merton, R.K. (1968). The Matthew effect in science. *Science*, 175, 56-63.
- Moelants, F., Mommers, C., & Oud, H. (1990). Leerlingvolgsystemen verklaard en vergeleken. *School en Begeleiding*, 26, 19-28.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School Matters: The Junior Years*. Wells: Open Books.
- Nuttall, D.L., Goldstein, H., Prosser, R., & Rasbash, J. (1989). Differential school effectiveness. *International Journal of Educational Research*, 13, 7, 769-776.
- Raudenbush, S.W. (1989). The analysis of longitudinal, multilevel data. *International Journal of Educational Research*, 13, 7, 721-740.
- Raudenbush, S.W., & Bryk, A.S. (1986). A hierarchical model for studying school effects. *Sociology of Education*, 59, 1-17.
- Reezigt, G.J., & Weide, M.G. (1989). *Effecten van differentiatie. Resultaten survey-onderzoek; deelrapport IV*. Groningen: RION.
- Sammons, P., Nuttall, D., & Cuttance P. (1993). Differential School Effectiveness: results from a reanalysis of the Inner London Education Authority's Junior School Project Data. *British Educational Research Journal*, 19, 4, 381-405.
- Scheerens, J. (1990). School effectiveness and the development of process indicators of school functioning. In *School effectiveness and school improvement*, pp. 61-80. Lisse: Swets & Zeitlinger.
- Scheerens, J., & Bosker, R.J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Shaywitz, B.A., Holford, T.R., Holahan, J.M., Fletcher, J.M., Stuebing, K.K., Francis, D.J., & Shaywitz, S.E. (1995). A Matthew effect for IQ but not for reading: Results from a longitudinal study. *Reading Research Quarterly*, 30, 4, 894-906.
- Slavin, R.E. (1994). Quality, appropriateness, incentive, and time: a model of instructional effectiveness. *International Journal of Educational Research*, 21, 2, 141-155.
- Slavin, R.E. (1996). *Success for all*. Lisse: Swets & Zeitlinger.
- Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 4, 360-406.
- Stringfield, S.C. (1995). Attempting to enhance student learning through innovative programs: The case for schools evolving into high reliability organizations. *School Effectiveness and School Improvement*, 6, 67-96.
- Stringfield, S.C., & Slavin, R.E. (1992). A hierarchical longitudinal model for elementary school effects. In B.P.M. Creemers & G.J. Reezigt (Eds.), *Evaluation of effectiveness* (ICO Publication 2). Groningen: ICO.
- Walberg, H.J., & Tsai, S.L. (1983). Matthew effects in education. *American Educational Research Journal*, 20, 3, 359-373.
- Weide, M.G. (1995). *Effectief basisonderwijs voor allochtone leerlingen*. Groningen: RUG/GION.

NOOT

1. Om precies te zijn: voor leerlingen met een score van 0 op milieu. Omdat alle scores zijn getransformeerd naar z-scores, is het gemiddelde gelijk aan 0.

Notities en Commentaren

A MODEL BASED SCALE ANALYSIS OF A DUTCH VERSION OF BRADLEY AND BRYANT'S RHYME AND ALLITERATION TEST

Margo G.H. Jansen
University of Groningen

Abstract

The quality of an experimental version of the R&A test as a measurement instrument for phonological sensitivity in the context of reading research, was studied using latent trait theory based methods. From a classical theory point of view the complete set of 32 items was found to be adequate, with an overall reliability of .85. To obtain a homogeneous Rasch scale revisions were necessary. A subset of 18 items was found to fit the assumptions of the one-parameter model (OPLM) a generalisation of the Rasch model. Raw scores on both the reduced and complete set of items were found to correlate highly with the weighted subset score provided by OPLM.

Introduction

In the course of learning to read correctly and fluently, children have to learn to identify printed or written words. A pupil is considered dyslectic if, under adequate teaching conditions, the automatization of the word identification process develops very slowly or incomplete. Coltheart (1978) distinguished between two word identification procedures or strategies, a phonological and a lexical. The phonological procedure consists of a number of steps: the identification of graphemes or clusters of graphemes followed by the translation into the corresponding phonemes, which then have to be assembled into words. The lexical procedure involves the instant recognition of words. In the Netherlands, the reading instruction explicitly teaches decoding skills, so with young readers the phonological strategy will predominate. It seems also reasonable that the phonological strategy will be used for novel or unusual words (e.g. pseudo-words), and the lexical strategy more often for easy well-known (and 'real') words. Bryant and others (1985) have demonstrated a relation between reading performance and the awareness of phonological similarities between spoken words in English. Also, backward readers tended to perform badly at detecting and forming rhymes. Pre-school differences were shown to be predictive for later differences in learning to read. In a series of studies Wimmer and co-workers explored the relation between phonological awareness and reading. Wimmer, Landerl and Schneider (1994) raised the question if rhyme and word-onset awareness is of similar importance in learning to read in German as in English. In general the German writing system is more consistent with regard to grapheme phoneme correspondence, making a phonics approach to reading therefore more successful than in English. The same argument is applicable to learning to read in Dutch. To be able to study this issue, adequate measuring instruments are of utmost importance. To measure phonological awareness Bradley and Bryant developed a test for detecting rhyme oddities (1985). Wimmer et al. used a German adaptation of the same instrument. Van den Bos and Spelberg (1997) developed an experimental Dutch version.

Use of the test implies that we suppose that the extent to which the subjects possess the property of phonological awareness can be inferred from the observed behaviour on the test. The measurement of phonological awareness presents us with a problem, which it shares with the measurement of many other properties such as intelligence, neuroticism, language ability etc. Such a concept has a certain surplus value in the sense that it cannot be fully covered by manifest behaviour such as responses on test items however well chosen. For this reason

we refer to these properties as latent traits (Molenaar, 1995). In order to infer a person's position on the trait from her or his answers to the test items a statistical model for the measurement process is needed. Such models are provided by Item Response Theory. IRT models are based on assumptions such as unidimensionality of the measurement instrument, and statistical test for checking if these assumptions hold, are usually available.

In the present article we will describe an IRT based scale analysis of the Dutch version of the rhyme-and-alliteration (R&A) test. Since the R&A test will be used to make comparisons of subjects of different age groups, it is for instance important to know if the test measures the same (latent) ability or combination of abilities for different subpopulations, otherwise scores cannot be compared. IRT provides the tools for checking on the validity of these assumptions.

Method

Subjects

The sample consisted of 789 children in grade 1 through grade 6 of four primary schools. There were 396 boys and 393 girls. Although the subjects were not selected by a random selection procedure, we assume that the sample represents the normal Dutch primary school population.

The R&A Instrument

The R&A test we consider here is an experimental version based on the original rhyme and alliteration task developed Bradley & Bryant (1985). The test consists of 32 items, and is preceded by eight practice items. An R&A test item consists of four meaningful one-syllable words. The set of 32 items consists of four instances - with varying correct-rejection positions - of eight subcategories which can be distinguished in the three major linguistic categories of Onset Cluster, Rhyme, and End Cluster. In the present task, the category Onset Cluster means that in all words but one, the onset is the same. There are three subcategories: the number of letters in the onset clusters can be 1, 2, or 3; an example of an item from the first subcategory (Dutch words) is: *keer, koop, zaag, kiem*. The category Rhyme means that three out of four words of an item rhyme, and one does not. There are two subcategories: rhyme is disturbed by the 'middle sound' or by the 'end sound'. An example (in Dutch!) of an item from the first subcategory is: *zuur, muur, door, duur*, and an example from the second subcategory is: *dier, kies, hier, gier*. The category 'End Cluster' means that in all words but one, the end cluster is the same. There are three subcategories: the number of letters in the end clusters can be 1, 2, or 3. An example of an item from the second subcategory is: *bukt, want, lont, sint*. Items are audiotaped and presented to the children via a taperecorder. After having listened to an item twice, the child is asked to indicate which word "does not belong to the set on the basis of its sound." The test is scored by counting the number of correct responses.

Scale analysis

In the following paragraph we will discuss the results of the attempts at constructing a Rasch-homogeneous Scale for Rhyme-Oddities Detection. For this we have used the OPLM program. Among latent trait models the Rasch-model has a special position. For instance, if the Rasch-model holds, the usual scoring rule of counting the number of correct responses can be used without loss of information. However, the Rasch-model is based on assumptions that may be too restrictive.

OPLM is an item response model that combines the tractable mathematical properties of the Rasch-model with the greater flexibility of the two-parameter logistic model. In OPLM we have item difficulty parameters which have to be estimated and discrimination indices with imputed values. Ideally, the imputed values are hypotheses based on theoretical considerations, supplemented by empirical information. Usually we start assuming the Rasch-

model, which implies that all items have the same discrimination index with is fixed at a value of 1. Appropriate model tests are available, both on item and on test level. The item parameters can be estimated by conditional maximum likelihood methods, of which the advantages are well known in the psychometric literature.

Results

Preliminary analyses

Before turning to the main analysis, we will first present a few descriptive statistics. In table 1 we find the mean and standard deviation per grade of the scores on the total of 32 items. From this table we see that the mean increases regularly with age level. The percentage of correct responses varies between slightly larger than 40 percent in the third grade and more than 80 percent in eighth grade. In the next table we find the results of a reliability analysis, also performed on the set of 32 items. Of the 789 subjects 26 had been excluded because they had either zero or perfect total scores. Skipped items were counted as errors.

Table 1: Means and standaard deviations of the score on the 32 items of the R&A task per grade and overall test characteristics

Grade	Mean	Stddev	N
3	14.0	4.28	124
4	18.4	4.45	97
5	21.9	4.87	197
6	23.0	5.00	107
7	25.0	5.47	132
8	26.6	4.31	132

The internal consistency is, with a coefficient alpha of .85 fairly high. The same value was reported in an earlier study where the same data set was used. Lower alpha's of .65, .68, .75, .78, .85, and .79 were obtained for the separate grade levels (Dekker, 1996; Van den Bos and Spelberg, 1997). So, if we are restricting ourselves to using the classical test theory approach as our scaling model, we have no reason for a (further) revision of our test. A next logical step is to use the number correct score as the scale score. A factor analysis performed on the same data set (Dekker, 1996) showed that the test is not unidimensional in a factor analytic sense.

Analyses with OPLM

In the first analysis the complete R&A task was analyzed over the whole range of pupils of group 3 through group 8. Table 2 contains the main results. The discrimination indices (A) are equal to one. The third column of the table presents the estimated values of the item difficulties (B). The columns headed by respectively S, M, M2 and M3 contain fit-indices per item. The S statistic is defined at the item level and is based on the differences between the observed and expected proportions of a (correct) response in homogeneous score groups. The test has power against misspecifications of the discrimination indices. The M statistics are based on a rationale originally developed by Molenaar (Glas & Verhelst, 1995). The scores are partitioned in a high and a low group (eventually also a medium group), and the differences between the observed and the expected number correct are combined. A negative outcome indicates that the item discriminates better than average (the discrimination index is set to one for each item) and a positive value points to a low discriminating item. The three M's use different partitions of the total score. Large values suggest up- or downgrading of the discrimination indices in order to increase the fit (Verhelst, Glas & Verstralen, 1995). An index for the global fit, the R1c, can be found at the last row of the table. A more detailed description of fit indices can also be found in Eggen & Sanders (1993).

Table 2: Results of the OPLM analysis with the 32 rhyme and alliteration items

nr	A	B	SE(B)	S	DF	P	M	M2	M3
1	1	-.241	.087	4.644	5	.461	-.394	.369	.624
2	1	.453	.081	11.084	6	.086	2.432	1.579	2.338
3	1	-1.333	.113	12.777	5	.026	-2.297	-2.440	-2.204
4	1	-.002	.084	17.444	6	.008	.878	-.704	-.508
5	1	-1.170	.107	2.658	5	.753	.003	.070	-.456
6	1	.890	.080	19.929	6	.003	2.887	3.494	2.391
7	1	.168	.082	1.632	6	.950	.782	1.103	.760
8	1	.643	.080	36.897	6	.000	5.570	5.384	5.641
9	1	-.368	.089	3.330	5	.649	-.632	-.208	-.375
10	1	-.157	.086	6.724	5	.242	-1.582	-2.189	-2.116
11	1	.954	.080	9.901	6	.129	-.567	-.133	1.841
12	1	-.698	.095	9.242	5	.100	-.947	-2.213	-2.712
13	1	.299	.081	17.385	6	.008	-2.639	-2.930	-3.681
14	1	.604	.080	23.852	6	.001	2.347	1.044	2.220
15	1	.650	.080	90.455	6	.000	7.249	7.455	7.562
16	1	-2.661	.185	10.608	2	.005	-1.125	-1.955	-2.514
17	1	1.413	.081	6.833	6	.337	-.982	-1.661	.008
18	1	.993	.080	9.544	6	.145	-1.895	-2.125	-1.328
19	1	.339	.081	5.018	6	.541	1.328	.422	1.878
20	1	-.187	.086	6.446	5	.265	-.454	2.011	1.311
21	1	-.502	.091	5.753	5	.331	-.103	-1.704	-1.782
22	1	.112	.083	7.876	6	.247	.190	-.476	-.612
23	1	-1.206	.109	8.146	5	.148	-2.349	-2.919	-2.137
24	1	.806	.080	17.546	6	.007	-2.880	-3.658	-2.546
25	1	-1.088	.105	7.234	5	.204	-1.745	-1.613	-1.937
26	1	-.643	.094	5.008	5	.415	.820	-.371	-1.271
27	1	.838	.080	5.417	6	.492	-.204	.538	-.849
28	1	-1.134	.106	4.175	5	.525	-.861	-.883	-1.421
29	1	.182	.082	9.801	6	.133	-.448	-1.301	-1.939
30	1	1.058	.080	12.378	6	.054	-2.301	-1.022	-2.852
31	1	.154	.083	4.984	6	.546	-2.146	-.866	-1.560
32	1	.832	.080	20.375	6	.002	2.479	2.425	2.575

R1c = 295.79; df = 93; p = .00

N = 763

Mean = 21.4 S.D. = 6.04

Alpha = .849

From this it is clear that the set of 32 items does not fit the assumptions of the Rasch model (R1c = 295.79; with df = 93). Deleting the items with relatively large values (positive or negative) for all three of the item-level fit-indices (items 3, 5, 8, 13, 15, 23, 24 and 32) produces an increase of fit (R1c = 121.52; df = 69; p = .00), but not sufficiently. This also happens when we relax the assumption of the Rasch-model that all items have the discrimination index. Allowing values for the item discrimination parameters to differ over items increases the fit, but the "overall-fit" is still not sufficient. Therefore we concluded that our approach, selecting items purely on the basis of individual fit-indices, would not lead to the desired goal. To obtain Rasch homogeneous scales we could turn to the results of the factor analysis for a suggestion to split the complete set into possibly homogeneous subsets. We decided instead, to reduce the original set of 32 items to a smaller set, based on a procedure, where content oriented and statistical considerations were combined. All items were re-examined and items that we saw as possibly flawed, were deleted. An example is 'beek boom baas wees', where

the last word is the odd-one-out, since all the others have the same begin consonant (alliteration first sound). Note however that 'baas' and 'wees' both end on the same consonant 's', which may lead to confusion. The resulting smaller set of only 20 items was subjected to a following series of Rasch-analyses. Another two items that still showed a considerable amount of misfit were deleted. Then we performed another OPLM-analysis with imputed values for the discrimination parameters. The remaining set of 18 items showed a good fit globally. The main results are given in Table 3. This item set can be considered as forming a homogeneous scale. On item level we have a few items that, according to their individual M- of S-indices, still may benefit from a further adjustment of the discrimination-index values (for items 2 and 7 for instance the values for the A-index should be adjusted downwards toward 1).

Table 3: Results of an OPLM analysis of the 18 remaining rhyme items where the discrimination indices (A) are allowed to differ

nr	A	B	SE(B)	S	DF	P	M	M2	M3
1	2	.044	.044	6.842	6	.336	.559	.971	.185
2	2	.397	.041	9.600	6	.143	2.684	1.899	2.048
3	3	-.251	.041	2.806	4	.591	1.368	-.509	-1.424
4	2	.166	.042	7.418	6	.284	1.316	-.348	-.047
5	2	-.430	.054	1.532	5	.909	-.287	-.186	-.066
7	2	.253	.042	6.749	6	.345	2.052	1.308	1.787
8	1	.363	.074	13.343	6	.038	1.482	.318	-.465
9	2	-.020	.045	16.825	6	.010	-.059	.176	.619
10	2	.087	.043	5.072	6	.535	-.627	-1.095	-1.506
11	2	.651	.041	7.089	6	.313	.052	.546	.210
12	3	-.009	.035	3.477	4	.481	.976	.045	.427
13	2	.319	.041	4.307	6	.635	-1.754	-1.465	-1.278
15	1	.369	.074	10.186	6	.117	.813	.408	.519
16	3	-.762	.067	.756	2	.685	-.355	-1.310	-.965
20	2	.072	.043	7.073	6	.314	.638	1.297	1.360
23	2	-.448	.054	10.627	5	.059	-1.256	-2.497	-2.076
25	2	-.388	.053	6.454	5	.264	-1.514	-1.118	-.864
28	2	-.411	.053	4.974	5	.419	-.740	-1.435	-1.195

R1c = 59.546; df = 51; p = .1926

A Comparison of Scoring Rules

In the preceding paragraph we have shown that it is possible, though at the loss of a substantial number of items, to select a subset of items satisfying the assumptions of OPLM. This implies that we not also have a shorter test, but also that the scoring rule which makes the best use of the available information is a weighted instead of the simple sumscore. The weights are the discrimination parameter values. The next question we will address here, is how much of a difference it makes if we use the simple scoring rule instead of the optimal one. And secondly, how the sumscores on the reduced set compare with the simple sum scores on the original 32 items. The correlation between the weighted and unweighted sums of the 18 items, which form the OPLM scale, is .99, so from a practical point of view, using the weights makes no difference. The correlation between the unweighted sumscore on the original 32 items, and the other two sumscores is systematically lower, and on closer inspection, the relative differences between the full and the selected set of items seem most noticeable in the lower score ranges (this might be related to ceiling effects). Nevertheless, the linear regression function of sum18 on sum32, provides a good fit, with about 85 percent of explained variance.

Table 4: Correlations between weighted and unweighted sum scores of R&A-items

	SUM18	SUM32	WSUM18
SUM18	1.00	.92	.99
SUM32	.92	1.00	.92
WSUM18	.99	.92	1.00

Discussion

The reliability of the Dutch experimental version of the R&A test is, with a coefficient alpha of .85, fairly high. So, using a classical test theory approach the R&A is an acceptable instrument, needing no further adjustments (Eggen & Sanders, 1993), and unweighted sum-scores can be used to scale the subjects.

Analysing the data on the basis of the Rasch model, leads to different conclusions. First, it becomes clear that the set of 32 items does not form a Rasch scale. It is possible however, to find a subset of 18 items, fitting the less restrictive assumptions of OPLM. However, both the unweighted and the weighted sumscores on this shorter scale correlate highly with the unweighted sumscore on the complete set. We have therefore reasons to expect that it does not make much difference also, which of the scale scores is used for relating phonological sensitivity to other variables of interest. So, in our case, the additional benefits of using IRT scaling methods instead of classical ones are not large.

References

- Bradley, L., & Bryant, P. (1985). *Rhyme and reason in reading and spelling*. Ann Arbor: The University of Michigan Press.
- Brons, I. (1990). De constructie van een meetinstrument voor fonologisch bewustzijn. [The construction of a measurement instrument for phonological awareness]. In H. Veenker & P. van Geert (Eds.), *Eindrapport van het project taalverwerving en leesstrategieën bij dove kinderen [Final report on project 'language acquisition and reading strategies in deaf children']*. University of Groningen.
- Coltheart, M. (1978). *Lexical access in simple reading tasks*. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151 - 216). London: Academic Press.
- Dekker, K. (1996). Research with the Dutch rhyme-and-alliteration task. University of Groningen.
- Eggen, T.J.H.M., & Sanders, P.F. (1993). *Psychometrie in de Praktijk*. Arnhem: Cito, OPD.
- Glas, C.A.W., & Verhelst, N., (1995). *Testing the Rasch model*. In: Fisher, G.H. & Molenaar, I.W.(Eds.), *Rasch models: Foundations, recent developments and applications* (pp.3-14). New York, Springer-Verlag.
- Molenaar, I.W.(1995). *Some background for item response theory and the Raschmodel*. In: Fisher, G.H. & Molenaar, I.W.(Eds.), *Rasch models: Foundations, recent developments and applications* (pp.3-14). New York, Springer-Verlag.
- Van den Bos, K.P. & Lutje Spelberg, H.C. (1997). *Measuring word identification skills and related variables in Dutch children*. In: C.K. Leong & R. Malatesha Joshi(Eds.), *Cross-language studies of learning to read and spell: Phonologic and orthographic processing* (pp.271-288). Dordrecht: Kluwer Academic Publishers.
- Verhelst, N., Glas, C.A.W., & Verstralen, H.H.F.M.(1995). *One-parameter Logistic Model OPLM*. Arnhem: Cito, OPD.

Dr. Margo G.H. Jansen
 Department of Education
 University of Groningen
 Grote Rozenstraat 38
 9712 TJ Groningen.

Boekbesprekingen

Sabine Severiens

Gender and Learning. Amsterdam: Proefschrift, 1997.

ISBN nr 90-9010996-X

In dit proefschrift wordt onderzoek naar leerstijlen, zoals dat in de onderwijskunde en onderwijspsychologie plaatsvindt, verbonden met de vraag naar sekseverschillen, waarbij vooral vanuit het theoretisch kader van vrouwenstudies gewerkt wordt. Ik gebruik in het onderstaande het woord sekse als vertaling van het door Severiens gebruikte gender.

Ihoud

In het eerste hoofdstuk zet Severiens haar onderzoek in twee tradities. Ten eerste die van de toegenomen belangstelling voor het leren zelf, de processen waarlangs leren plaatsvindt en de strategieën die lerenden daarbij gebruiken. Daarbij is ook aandacht voor individuele verschillen, en voor verschillen tussen groepen studenten die bijvoorbeeld een bepaalde belangstelling, motivatie of ervaring gemeenschappelijk hebben. Het lijkt dus voor de hand liggend om ook eens te kijken naar mannen en vrouwen. En daarmee komt de tweede onderzoekslijn in zicht, die van *gender and education*. Traditioneel heeft deze onderzoeksrichting zich altijd meer bezig gehouden met de uitkomsten van het leerproces, zoals de achterstand van meisjes op het gebied van wiskunde, dan met het leerproces zelf. Het proefschrift wil deze lacune opvullen, de centrale vraagstelling ervan betreft sekseverschillen in leren.

Met betrekking tot leren maakt Severiens meteen een duidelijke keuze voor Vermunt's (1992) theorie over leerstijlen, mede omdat daarin leerstijlen als afhankelijk van de context beschouwd worden, wat de mogelijkheid geeft om de rol van het onderwijs bij de totstandkoming van sekseverschillen te onderzoeken. Wat betreft theorievorming over sekse kiest ze voor Baxter Magolda's (1992) theorie over *ways of knowing*.

Het tweede hoofdstuk geeft een beknopt overzicht van de na 1980 verschenen literatuur op het gebied van sekseverschillen in leerstijlen. Gedeeltelijk heeft dit de vorm van een verhalend overzicht, gedeeltelijk die van een meta-analyse van twee soorten onderzoek die met veelgebruikte standaardvragenlijsten voor het meten van leerstijlen hebben gewerkt. De theorieën op dit gebied worden in drie soorten ingedeeld: theorieën die leerstijlen beschouwen als een stabiel persoonlijkheidskenmerk, theorieën die stellen dat de omgeving er wel enige invloed op heeft, en tenslotte theorieën die stellen dat leerstijlen voornamelijk door omgevingskenmerken bepaald worden. Uit de twee overzichtsstudies tezamen blijkt dat bij onderzoeken die gebruik maken van een individueel stabiel leerstijlconcept geen sekseverschillen gevonden worden; en dat mannen meer dan vrouwen de voorkeur geven aan abstract conceptualiseren als leerstijl en meer extrinsiek gemotiveerd zijn. De resultaten zijn nogal heterogeen, er komt geen sterk beeld naar voren van een mannelijke en een vrouwelijke leerstijl.

Het derde hoofdstuk beschrijft een empirisch onderzoek in het volwassenenonderwijs. Severiens meent dat de heterogeniteit in het al dan niet vinden van sekseverschillen mede te wijten kan zijn aan een te nauwe, biologische definitie van het begrip sekse als man/vrouw. Zij neemt daarom sekse-identiteit mee als alternatieve conceptualisering van sekse. Het onderzoek bestond uit een enquête onder 432 lerenden uit het volwassenenonderwijs, met twee vragenlijsten, Vermunt's ILS om leerstijlen te meten en een instrument om sekse-identiteit te meten. De opdracht was om de vragenlijst met het oog op een bepaald vak in te vullen. Op grond van een factoranalyse worden vier leerstijlen onderscheiden, waarvan er drie overeenkomen met wat Vermunt had gevonden (betekenisgericht, reproductiegericht, en ongericht leren), en een vierde die de *jezelf bewijzen* stijl wordt genoemd. De docent blijkt geen invloed op leerstijl te hebben, maar het vak wel. Sekse-identiteit blijkt een betere voorspeller van leerstijl dan sekse.

Hoofdstuk vier beschrijft een longitudinale studie met 190 respondenten binnen het volwassenenonderwijs. De onderzoeksvraag is of vrouwen en mannen gedurende een schooljaar in verschillende mate stabiel zijn in hun leerstijl, waarbij ook weer de contextvariabelen leerkracht en vak betrokken worden. Ook hier werd de ILS gebruikt met dezelfde instructie als hierboven. Omdat aan het eind van het schooljaar de factorstructuur van de ILS anders blijkt te zijn dan aan het begin worden niet de scores op factoren (leerstijlen) vergeleken maar op de afzonderlijke schalen. Vrouwen en mannen veranderen op een verschillende manier op 5 van de 18 subschalen. Er wordt geen effect van leerkracht gevonden.

Paste het onderzoek tot nu toe nog naadloos binnen de mainstream onderwijskunde, vanaf hoofdstuk 5 komt een alternatief perspectief aan de orde in de vorm van het Epistemological Reflection Model van Baxter Magolda. Hierin worden vier stadia van weten onderscheiden, en daarbinnen steeds twee redeneerpatronen, connectionistisch en autonoom genoemd. Het eerste zou meer kenmerkend zijn voor vrouwen, het tweede voor mannen. Bij een onderzoek met dezelfde methode die Baxter Magolda gebruikte, interviews, onder 53 studenten uit het volwassenenonderwijs, bleken dergelijke sekseverschillen inderdaad op te treden. In een onderzoek door middel van een vragenlijst, bij 432 studenten uit het volwassenenonderwijs, bleven de veronderstelde sekseverschillen echter uit.

In het zesde hoofdstuk worden de theorieën van Vermunt en van Baxter Magolda eerst op conceptueel niveau vergeleken. Dan blijkt dat het begrip *leerconceptie* uit de leerstijlentheorie overeenkomsten vertoont met het begrip *manieren van weten*. In interviews met 53 studenten uit het volwassenenonderwijs werd een dergelijke overeenkomst echter niet gevonden.

In het laatste hoofdstuk worden de resultaten van de voorgaande hoofdstukken nog eens op een rijtje gezet, de gebruikte theorieën kritisch besproken, en aanbevelingen voor verder onderzoek gedaan.

Commentaar

De hoofdstukken 2 tot en met 6 van het proefschrift zijn als afzonderlijke artikelen geschreven. Dit heeft het voordeel dat ze beknopt en zakelijk geschreven zijn. Het nadeel is echter dat er wel eens wat overlap in de inleidingen zit. Een ander nadeel is dat daardoor nergens beschreven wordt dat alle onderzoeken bij dezelfde groep studenten gedaan zijn. Het zal de oplettende lezer al opgevallen zijn dat de vragenlijststudies en de interviewstudies onderling dezelfde aantallen respondenten hebben. Bij eerste lezing was mij dit niet opgevallen, die exacte aantallen vormen immers maar een detail in de grote hoeveelheid informatie. Nergens in het hele proefschrift wordt hier iets over gezegd. Navraag bij de auteur leerde echter dat het hier dezelfde steekproef betrof. De 53 respondenten van de interviewstudies vormen bovendien een subgroep uit de 432 van de vragenlijststudies, evenals de 190 van de longitudinale studie; en de laatste twee subgroepen overlappen elkaar. Kortom, eigenlijk betreft het hier één onderzoek met verschillende onderdelen waarover in afzonderlijke hoofdstukken c.q. artikelen verslag gedaan wordt.

Dat is om verschillende redenen een dubieuze praktijk. Ten eerste vanuit een statistisch oogpunt, omdat het leidt tot inflatie van resultaten. In de in hoofdstuk drie en vier beschreven onderzoeken worden bijvoorbeeld niet de verwachte sekseverschillen in leerstijlen gevonden, en geen effect van docent, en dat wordt beide keren als belangrijke resultaten gepresenteerd. Maar het betreft hier dus een onderzoek bij dezelfde respondenten. Geen wonder dat er dan twee keer hetzelfde resultaat gevonden wordt. Ten tweede wordt door het afzonderlijk publiceren van artikelen, en door in hoofdstuk 7 te spreken van *separate studies* (in de Nederlandse samenvatting *de verschillende studies*), de indruk gewekt dat het allemaal afzonderlijke onderzoeken betreft. Dat is een onjuiste voorstelling van zaken. Op deze manier publiceren (*piecemeal publication*) wordt overigens sterk ontraden door de gezaghebbende American Psychological Association, in de publicatiehandleiding die de meeste Nederlandse tijdschriften (inclusief het TOR) ook hanteren, en als het al gebeurt dan is het in elk geval

voorgescreven om expliciet aan te geven dat er andere verslagen gebaseerd op dezelfde data-verzameling zijn verschenen (APA, 1994, p. 296-297). Tenslotte is het in theoretisch opzicht jammer, want bij het in hoofdstuk 5 beschreven onderzoek dringt zich de vraag op of hier niet ook sekse-identiteit een rol zou kunnen spelen. Nu dus blijkt dat de onderzoekster van deze respondenten de sekse-identiteitscores ter beschikking had vind ik het eigenlijk onvergeeflijk dat deze niet in de analyse betrokken zijn.

Afgezien van dit probleem heeft het proefschrift een aantal aantrekkelijke aspecten. De combinatie van een verhalend en een meta-analytisch literatuuroverzicht is erg degelijk. De toetsing van twee leerstijlen-theorieën op een grote groep niet-standaard studenten levert interessante inzichten op. Toch laat het als validering van de LSI en de daarin onderscheiden leerstijlen te wensen over; een confirmatieve factoranalyse was dan de aangewezen methode geweest in plaats van de nu gebruikte principale componenten analyse. Het advies om voortaan de leerstijlen maar te laten vallen en alleen op schaalniveau naar de resultaten te kijken lijkt me dan ook wat voorbarig.

Het bestuderen van sekse wordt tamelijk eenzijdig opgevat als het zoeken naar sekseverschillen. Zelfs als in hoofdstuk 3 meteen al gebleken is dat de psychologische variabele sekse-identiteit meer variantie verklaart dan biologische sekse, worden in alle volgende hoofdstukken weer onvermoeibaar pogingen gedaan om sekseverschillen op te sporen, terwijl sekse-identiteit buiten beeld blijft. Het zou voor een psycholoog toch interessanter moeten zijn om de relatie tussen leerstijlen en sekse-identiteit, een variabele die mede voorgaande socialisatieprocessen vertegenwoordigt, te onderzoeken dan altijd maar weer naar verschillen tussen mannen en vrouwen te kijken?

Literatuurlijst

- American Psychological Association. *Publication Manual of the American Psychological Association* (4th ed.). Washington, DC: Author.
- Baxter Magola, B. M. (1992). *Knowing and reasoning in college. Gender related patterns in students' intellectual development*. San Francisco: Jossey-Bass.
- Vermunt, J. D. H. M. (1992). *Leerstijlen en sturen van leerprocessen in hoger onderwijs. Naar procesgerichte instructie in zelfstandig denken*. Amsterdam/Lisse: Swets & Zeitlinger.

Prof. dr. Tineke M. Willemsen
 Departement Vrouwenstudies, Katholieke Universiteit Brabant
 T.M.Willemsen@kub.nl

Doolaard, S., Cremers-Van Wees, L.M.C.M. & Bosker, R.J. (1999)
Basisvorming in 1996; beschrijving en vergelijking met de periode voor invoering.
 Enschede: OCTO, Universiteit Twente. ISBN 90-365-1357-x.

In 1993 werd de basisvorming in de onderbouw van alle Nederlandse middelbare scholen ingevoerd. De bedoeling was dat er nogal wat zaken zouden veranderen. Vakinhouden werden gereviseerd waarbij *vaardigheden*, *toepassingen* en *samenhang* de steekwoorden waren. De vakken Techniek en Verzorging werden aan het verplichte curriculum toegevoegd. Inlijving van lager beroepsonderwijs bij het algemeen vormend onderwijs werd aangemoedigd zodat leerlingen binnen één school makkelijker van schooltype-niveau zouden kunnen wisselen. Niet langer na groep 8 van de basisschool, maar pas na twee of drie jaar voortgezet onderwijs zouden leerlingen de basisvorming *afsluiten* en een advies voor een vervolgopleiding moeten krijgen.

Na 6 jaar is het tijd voor evaluatie en in september 1999 werd een onderzoeksrapport van de Universiteit Twente gepresenteerd: *Basisvorming in 1996; beschrijving en vergelijking*

met de periode voor invoering van S. Doolaard ea. De presentatie hiervan werd in de kranten verslagen met krantenkoppen als "Basisvorming blijkt een flop" (De Telegraaf, 3 september 1999). Aldus konden we uit de kranten kennis nemen van harde conclusies. Dergelijke krantenartikelen zijn vaak karikaturale uittreksels van een persbericht en geven niet de inhoud van het werkelijke rapport weer. Is het rapport van de Twentse onderzoekers echt zo negatief?

Het genoemde rapport baseert zich op de longitudinale VOCL-studie (Voortgezet Onderwijs Cohort Leerlingen). Binnen het VOCL worden onderwijsgegevens verzameld rondom leerlingen in hun eerste en in hun derde leerjaar. Het eerste VOCL-cohort bestaat uit leerlingen die in 1989 in het voortgezet onderwijs van start gingen. Het volgende cohort is van 1993. Hierdoor beschikt men dus over gegevens van één cohort van vóór en een van ná de invoering van de basisvorming. In het hier besproken onderzoeksrapport van Doolaard ea. is gebruikt gemaakt van de VOCL-data uit vragenlijsten aan schoolleiders, vaksectieleiders en leraren (Nederlands en wiskunde) verzameld in 1989/90 (leerjaar 1), 1991/92 (leerjaar 3), 1993/94 (leerjaar 1) en 1995/96 (leerjaar 3). Recentere gegevens, zoals die van het derde VOCL cohort werden niet gebruikt.

Uit de enorme hoeveelheid data maakte de Twentse onderzoeksgroep een selectie aan de hand van onderstaande onderzoeksvragen. De eerste vraag luidde: welke variabelen van de VOCL-studie zijn bruikbaar voor een beschrijving van de stand van zaken in de basisvorming? De tweede onderzoeksvraag is een logische voortzetting hierop: wat is de stand van zaken, afgeleid uit diezelfde variabelen?

Het rapport begint met een globale opsomming van de VOCL-instrumenten van de meting uit 1995/96 en geeft daarin aan welke selectie van variabelen gemaakt is. De geselecteerde variabelen volgen hierna in bladzijden-lange tabellen. De criteria voor de selectie van de variabelen blijven echter achterwege. Daarmee is de eerste onderzoeksvraag met een opsomming afgehandeld. Voor de lezer die geïnteresseerd is in de theoretische achtergronden van de basisvorming is dat teleurstellend, want door het ontbreken van een rationale voor de selectie wordt er geen samenhang aangebracht tussen enerzijds de gezochte variabelen en anderzijds onderwijskundige principes zoals die aan de basisvorming ten grondslag liggen (meta-cognitieve vaardigheden, leerling-activerende instructiemodellen, harmonisering van het onderwijsaanbod, enz).

De beantwoording van de tweede onderzoeksvraag neemt het grootste deel van het rapport in beslag. De eerste hoofdstukken beschrijven de stand van zaken anno 1995/96 op schoolniveau, vaksectieniveau (wiskunde en Nederlands) en klasniveau. Met middelen uit de beschrijvende statistiek (frequentieverdelingen, gemiddelden en spreidingsmaten) wordt een beeld geschetst. Waar met variantieanalyse significante verschillen optraden, zijn de gegevens gesplitst tussen (i)vbo en de andere schooltypen. Verschillen tussen de vakken Nederlands en wiskunde werden getoets met een t-toets. Afwijkend van de gangbare praktijk nemen de onderzoekers een 0,001% significantieniveau om te vermijden dat teveel verbanden op grond van toeval als significant worden aangemerkt. Dat hiermee de kans toeneemt op fouten van de tweede soort (dat een meetbaar verband niet wordt ontdekt) nemen ze voor lief.

De hoofdstukken, waarin de vergelijkingen tussen vóór en ná de invoering worden gemaakt, volgen hetzelfde niveau-systeem: schoolniveau, klasniveau, echter niet op vaksectie-niveau omdat daarover geen eerdere metingen gedaan zijn. De verschillen zijn hier cross-sectioneel en longitudinaal gemeten met t-toetsen. Een groot probleem blijkt echter, dat veel vraagstellingen in de loop der jaren werden gewijzigd, zodat er soms moeilijk te vergelijken viel.

Op schoolniveau wordt er in dit rapport gekeken naar vernieuwingsaspecten die op het management afkwamen, zoals bijvoorbeeld de schoolverbreding. Op de vraag of men in de periode sinds de invoering van de basisvorming uitbreidde met één of meerdere (i)vbo-afdelingen,

antwoord 70% van de scholen "nee". Dat lijkt mager, maar tegelijkertijd blijkt dat elke school er tussen 1990 en 1996 gemiddeld één schooltype bij kreeg. Bestond de helft van de scholen in 1990 nog uitsluitend uit één schooltype (alléén vwo, alléén havo, etc), in 1996 blijken ze veelal gefuseerd en is het aantal "smalle scholen" minder dan 20% van het totale aantal geworden. Tegelijkertijd er zijn meer brede scholengemeenschappen gekomen: scholen met vijf schooltypes (vwo-havo-mavo-vbo-ivbo) waren in 1990 nog een zeldzaamheid, in 1996 maken ze 17% van het scholenbestand uit.

De indeling van de klassen gaat meestal volgens een *dakpan* van twee gecombineerde schooltypes (bijvoorbeeld vwo/havo-klassen) in de eerste twee leerjaren. Tussen 1990 en 1996 lijkt de brugperiode te zijn verkort en het jaar waarin de definitieve plaatsing van de leerlingen plaatsvindt is vroeger geworden, met name voor de ivbo-leerlingen. Hun schooltype ligt grotendeels al in klas 1 vast, en voor vbo-mavo is dat ruim voor het einde van leerjaar 2, terwijl dit in 1990 gemiddeld pas aan het einde van leerjaar 3 was. De beslissing hierover wordt bepaald aan de hand van de instroomgegevens (Cito-score in de basisschool) en rapportcijfers, waarbij het docentencorps dit kan overrulen door motivatie en werkhouding mee te nemen in het besluit. Bij de keuze van het vakkenpakket is de school adviserend. Scholen laten leerlingen liever een wat makkelijker pakket kiezen om de slagings-kans te verhogen. Maar een leerling (of diens ouders) heeft op de meeste scholen het laatste woord in deze.

Het rapport vervolgt met nog meer beschrijvingen op schoolbeleidsniveau. De invoering van de basisvorming kreeg gestalte in de lessentabel, de leerlingbegeleiding en de invulling van de *vrije ruimte* (het niet-verplichte deel van het curriculum). Onderwijskundige vernieuwingsaspecten als "het verminderen van het traditioneel frontaal klassikaal lesgeven" maken op geen enkele school integraal deel uit van het beleid. Het heeft de aandacht, men oriënteert zich erop, men plant en overlegt - en daar is het in 1995/96 vooralsnog bij gebleven.

De onderzoekers hebben ook gekeken naar de regelgevingen op scholen. Wat daarin opvalt is de toename in regelgeving zoals voor het vaststellen van rapportcijfers, te laat komen en spijbelen. Er zijn weinig afspraken over didactische werkvormen en hoe te handelen bij spijbelen tijdens een proefwerk. De kwaliteit van het onderwijs meten schooldirecties af uit rapport- en examenresultaten. Op 70% van de scholen bezoekt de directie wel eens een les, maar dan meestal in het geval van beginnende en bij probleem-leraren en dito klassen. Wat betreft pedagogisch-didactisch handelen en vakinhoud zijn de leraren nog steeds autonoom.

De leerlingbegeleiding is persoonlijker geworden: er wordt meer tijd besteed aan individuele begeleiding en daarnaast is er leer-, studie- en beroepskeuzebegeleiding. Extra taallessen worden georganiseerd voor i-leerlingen en allochtonen. Het monitoren van de studievoortgang middels proefwerkcijfers gebeurt steeds meer geautomatiseerd (inmiddels op 90% van alle scholen). Daarnaast bestaan er toenemende registraties van leer- en gedragsproblemen. 20% van de scholen heeft een geautomatiseerd leerlingvolgsysteem, en opmerkelijk is dat men dit vooral gebruikt om het aantal zittenblijvers van het aankomende schooljaar te schatten.

De onderzoekers hebben een aantal schoolmanagement-aspecten onderzocht die niet direct verband houden met de invoering van de basisvorming, zoals de vragen over *spijbelmanagement*, automatisering, en ook hoe ziekte van een collega wordt opgevangen (antwoord: door een collega te laten invallen), hoeveel lesdagen er per jaar uitvallen (antwoord: gemiddeld 18 van de 194 lesdagen) en hoe rapport- en eindcijfers tot stand komen (op de meeste scholen telt het laatste trimester het zwaarst mee voor het eindcijfer). Op deze punten wreekt het zich, dat de auteurs niet hebben geëxpliciteerd wat zij als aspecten van de invoering van de basisvorming beschouwen. Spijbelmanagement of automatisering zijn nooit vernieuwingsdoelen van de basisvorming geweest. Je kunt hooguit vermoeden dat dit gevolgen zijn van de schaalvergroting. Maar beide aspecten kunnen ook gevolgen zijn van maatschappelijke ontwikkelingen zoals normvervaging respectievelijk een toegenomen computergebruik.

Op vaksectie- en klasniveau leren we uit het rapport dat leerlingen in klas I gemiddeld 4 lessen Nederlands en 4 lessen wiskunde krijgen, terwijl dit in de twee daarop volgende leerjaren 3 uur is voor beide vakken. In de leerstofgebieden wordt er naar schooltype gedifferentieerd en blijkt er een duidelijke scheiding tussen avo/vwo en (i)vbo. De abstractere onderwerpen komen in het eerste geval meer aan bod (bij Nederlands: fictie, taalgebruik en stijl, bij wiskunde: formules). De afsluiting van de basisvorming gebeurt op (i)vbo na 3¹/₂-4 jaar terwijl dit op het vwo binnen 3 jaar gebeurt. Meestal gebruikt men hiervoor de speciale basisvormingstoets van het Cito, maar ook wel zelfgemaakte toetsen.

De schoolmethode in leerjaar 3 bij Nederlands is gemiddeld ingevoerd in 1993, terwijl bij wiskunde het jaar van invoering varieert per schooltype: de ivbo-methode is gemiddeld uit 1993, de vbo-methode is uit 1992, de mavo-methode is uit 1990, de havo-methode is uit 1989 en de vwo-methode is uit 1987 (blz 69). Een verklaring voor deze verschillen ontbreekt in dit rapport evenals een antwoord op de vraag hoe deze cijfers rijmen met een andere tabel, waarin 83% van de schoolleiders aangeeft met de invoering van de basisvorming een nieuwe methode voor wiskunde te hebben aangeschaft (blz 38).

Rond de invoering van de basisvorming lezen we in dit rapport dat sectieleiders zich hebben geïnformeerd middels voorlichtingsdagen, bijscholing en oriëntatie op een nieuwe methode. Wederom 83% van de wiskunde-sectieleiders geeft aan een nieuwe methode te hebben aangeschaft en daarnaast heeft 42% aanvullend materiaal bij de bestaande methode aangeschaft. Het lijkt er daarmee toch wel op, dat de meeste secties zich vakinhoudelijk vernieuwd hebben.

Van de leraren wiskunde van klas 3 geven 95% aan dat ze de schoolmethode volgen, terwijl 70% van de leraren Nederlands dit doet. Men geeft zelfs aan dat men de methode strikter is gaan volgen door de invoering van de basisvorming. 50% van de ondervraagde leraren geeft alle leerlingen dezelfde leerstof en 25% reguleert tempoverschillen met herhalings- en/of verrijksstof. De overigen hebben een variatie aan differentiatie modellen. Bij wiskunde wordt er meer tijd besteed aan huiswerkbespreking, terwijl er bij Nederlands meer tijd is voor uitleg van nieuwe stof en het maken van huiswerk tijdens de les.

Nieuwe leerstof wordt in 80% van de lessen aangeboden met klassikale instructie en het geven van voorbeelden. Daarna volgt het klassikaal oefenen en vervolgens gaan leerlingen individueel verwerkingsopdrachten maken. Metacognitieve sturing als "een overzicht van de les geven" komt weinig voor (en zeker niet bij wiskunde). Groepswerk gebeurt meestal in tweetallen, volgens de opstelling waarin de leerlingen al in het lokaal zitten. Volgens de auteurs lijkt het erop, dat er met de invoering van de basisvorming minder instructie plaatsvindt en er meer tijd is voor oefening.

De toetsing gebeurt via proefwerk of schriftelijke overhoring, waarbij de leraren Nederlands strenger blijken te zijn: voor een voldoende moet een leerling gemiddeld 60% van de punten behalen, bij wiskunde is dat 55%. De leraren geven aan dat de algemene vaardigheidsdoelen uit de basisvorming (een onderzoekje doen, een standpunt verwoorden) niet gerealiseerd worden. Wat leerlingen meer zouden moeten leren (aldus de leraren): zich concentreren, fouten herstellen en toepassen van kennis in nieuwe situaties.

Zoals uit het voorgaande blijkt bestaat het rapport "*Basisvorming in 1996; beschrijving en vergelijking met de periode voor invoering*" van Doolaard ea. voor een groot deel uit cijfermatige gegevens. Al met al krijgt de lezer een beeld van het reilen en zeilen in de onderbouw van Nederlandse scholen voor voortgezet onderwijs in de eerste drie jaar na de invoering in 1993, maar een gedegen beschrijving van de basisvorming is het niet. De nadruk ligt op management aspecten, en dat is niet verwonderlijk aangezien de auteurs allen gespecialiseerd zijn in "onderwijsorganisatie en -management". Hoe het nu zit met de invoering van de nieuwe vakken Techniek en Verzorging, hoe sommige leraren worstelen met vernieuwde leerstofgebieden (wiskunde kreeg nieuwe stof als *kijklijnen*, Nederlands kreeg nieuwe stof als *uiterlijke verzorging van teksten*), of leerlingen door de vernieuwde leerstof meer gemoti-

veerd worden, hoe de harmonisering tussen vakken verloopt, .. het blijven onbeantwoorde vragen in dit rapport. Wellicht dat de VOCL-data zich er niet voor leenden, maar ondertussen suggereerde de titel meer dan het rapport waarmaakt.

Een tweede punt van kritiek is dat dit rapport zes jaar na invoering van de basisvorming uitkomt, terwijl het slechts de eerste drie jaar ervan beschrijft. Voor beleidsaanbevelingen is een *informatiegap* van drie jaar te veel. Het zou bijvoorbeeld zo kunnen zijn, dat de ingezet- te veranderingen inmiddels alweer een stap verder zijn.

Maar hoewel de auteurs alleen de eerste invoeringsjaren in beeld brengen, komt toch een beeld naar voren van scholen in verandering. Voorbeelden hiervan zijn de verbreding van scholen, vernieuwing van schoolmethoden, duidelijkere regelgevingen, en wellicht ook minder instructie tijdens de lessen. Er zijn ook zaken onveranderd gebleven, zoals de autonomie van leraren en de manier waarop cijfers tot stand komen. En er zijn verslechtingen zoals de vervroeging van het keuzemoment van het schooltype.

De conclusies van het rapport zijn genuanceerder dan de krantekoppen na de presentatie suggereerden, zeker als in aanmerking wordt genomen dat de geconstateerde veranderingen zich voordeden in een relatief kort tijdsbestek. Gezien de logheid van een groot apparaat als het onderwijs zijn deze verschuivingen een prestatie van formaat.

Drs. F.P. Vos

*Universiteit Twente, afd. Curriculumtechnologie
Postbus 217, 7500 AE Enschede*

Vittorio Busato

Leerstijlen nader geanalyseerd

Dissertatie Universiteit van Amsterdam, 1998, ISBN 90-5470-079-3

Na het verschijnen van het proefschrift van Jan Vermunt in 1992 is er in de afgelopen jaren een stroom van Nederlands onderzoek gepubliceerd over leerstijlen. De belangstelling voor leren en leerprocessen past bij de eis van deze tijd dat mensen voortdurend nieuwe kennis en informatie kunnen verwerven. De wetenschappelijke belangstelling voor het werk van Vermunt berust op zijn originele aanpak, in het bijzonder de wijze waarop hij het begrip leerstijl een nieuwe plaats verschafte in de onderwijsleerpsychologie.

Theoretisch vatte Vermunt leerstijl op als een combinatie van cognitieve strategieën, meta-cognitieve modellen en motivationele kenmerken van de lerende. Om leerstijlen vast te kunnen stellen ontwikkelde Vermunt een leerstijldiagnose-instrument, de Inventaris Leerstijlen (ILS) die inmiddels in verschillende onderwijssettings op zijn merites is beoordeeld. Met de ILS worden vier leerstijlen gemeten die in een puntenscore worden uitgedrukt. Het gaat om een betekenisgerichte, een toepassingsgerichte, een reproductiegerichte en een ongerichte leerstijl. Aangenomen wordt dat één van de leerstijlen overheerst en dat daar dan ook een verklaring ligt voor verschillen in studiesucces.

Het proefschrift van Busato is een kritische reflectie op de psychometrische kwaliteiten van de ILS. Het doel van de studie van Busato was een empirische bijdrage en wetenschappelijke onderbouwing te leveren aan de waarde en de toepassingsmogelijkheden van de ILS als (diagnostisch) meetinstrument. Het tweede doel was de relevantie van de leerstijlentheorie van Vermunt kritisch te beschouwen.

Onderzoeksvragen en opzet van het onderzoek

Het proefschrift omvat acht hoofdstukken. De inleiding (*hoofdstuk 1*) bestaat grotendeels uit een samenvatting van het proefschrift van Vermunt. De eerste studie (*hoofdstuk 2*) is een replicatie-onderzoek naar de leerstijlentheorie uitgevoerd bij eerstejaars studenten psychologie in Amsterdam (UVA). De vraagstelling is of bij deze groep studenten dezelfde vier leerstijlen

worden gevonden als bij de studenten aan de Open Universiteit (Heerlen) en aan de Katholieke Universiteit Brabant (Tilburg) uit het onderzoek van Vermunt. Dezelfde analyses werden uitgevoerd, de interne consistentie en stabiliteit van de leerstijlen zijn nagegaan en de samenhang met studiesucces. Een verschil met Vermunt is dat Busato in het replicatie-onderzoek de variabele intelligentie meeneemt om na te gaan of daar een verklaring ligt voor de variantie in leerstijlen.

De tweede studie is een toegepast onderzoek naar het effect van een cursus studievaardigheden op betekenis- en ongerichte leerstijlen. De verwachting was dat studenten met een ongerichte leerstijl behoefte hebben aan begeleiding en zich vaker inschrijven voor een dergelijke cursus dan studenten met andere leerstijlen. Ook werd verondersteld dat de cursus een ongerichte leerstijl doet verminderen en dat een ongerichte leerstijl zich op korte termijn laat veranderen in een betekenisgerichte leerstijl.

Deelstudie drie (*hoofdstuk 4*) beschrijft twee validatie-onderzoeken. Onderzocht werd of de studenten inderdaad zo studeren als zij volgens eigen zeggen (ILS) gewoonlijk doen. Met behulp van hardop denkprotocollen werd empirische steun gezocht voor de verschillende leerstijlen zoals die met de ILS worden gemeten.

In *hoofdstuk vijf* is met behulp van een cross-sectionele en longitudinale studie onderzocht of de leerstijl van de studenten in de loop van de tijd verandering ondergaat. Ook hier werd de relatie van leerstijlen met studiesucces onderzocht. De verwachting was dat studenten naarmate hun studie vordert hoger scoren op productieve, c.q. toepassingsgerichte en betekenisgerichte, leerstijlen dan in de lagere studie jaren.

Centrale vraag in de vijfde studie (*hoofdstuk 6*) was de relatie tussen leerstijl, persoonlijkheid, intelligentie, prestatie-motivatie en studiesucces. Onderzocht werd wat de eigen bijdrage van leerstijlen is als voorspeller van studiesucces bovenop persoonlijkheid en motivatie. Nagegaan is of de vier leerstijlen (ILS) gerelateerd zijn aan eigenschappen zoals Extraversie, Vriendelijkheid, Gewetensvolheid, Neuroticisme en Ontwikkeling (The Big Five van Tupes en Christal, 1961/1992), gemeten met de Vijf persoonlijkheidsfactoren Test (5PFT) van Elshout en Akkerman, 1975.

Het doel van *hoofdstuk 7* was om met behulp van een 'third party-review' een zo goed mogelijk beeld te krijgen van de interne validiteit van de ILS. Onderzoek naar de interne structuur dient ter ondersteuning van de begripsvaliditeit van de test. De onderzoeksvraag was of een ander model met meer dan vier leerstijlen de voorkeur verdient.

Het proefschrift besluit met een algemene discussie (*hoofdstuk 8*) waarin wordt teruggekomen op de waarde van de ILS als (diagnostisch) meetinstrument en de leerstijlentheorie van Vermunt nog eens tegen het licht wordt gehouden.

Proefpersonen voor de eerste drie studies vond Busato onder de Amsterdamse eerstejaars studenten psychologie in 1993, die bij aanvang van hun studie verplicht waren aan een testweek deel te nemen waar een grote verscheidenheid aan psychologische tests werden afgenomen. Daar waar herhaling of aanvulling van tests nodig waren, werden (geselecteerde) groepen studenten apart aangeschreven. Voor de tweede reeks onderzoeken werden ook de gegevens gebruikt van de cohorten psychologie studenten die in 1994 en 1995 instroomden. De data werden geanalyseerd met factoranalyse, betrouwbaarheidsanalyse, (m)anova, regressie-analyse, clusteranalyse en factoranalyse met Lisrel.

Belangrijkste resultaten

We geven eerst de belangrijkste resultaten in het licht van de doelen die Busato zich had gesteld. Busato wilde nagaan hoe betrouwbaar de ILS is, of de ILS een geschikt instrument is om leerstijlen te meten; wat de mogelijkheden van de ILS in de praktijk zijn; wat de relaties zijn met andere begrippen; en of de ILS met een indeling in vier leerstijlen wel klopt.

Hoe betrouwbaar de ILS is, komt in drie onderzoeken aan de orde. Het replicatie-onderzoek liet zien dat in Amsterdam dezelfde vier leerstijlen werden gemeten als in Tilburg en Heerlen. Uit het cross-sectionele en longitudinale onderzoek bleek dat de ILS ook in de loop

van de tijd dezelfde leerstijlen blijft meten. Hoewel uit nadere toetsing naar voren kwam dat er meer dan vier factoren te onderscheiden zijn, werd op praktische gronden geconcludeerd dat een indeling in vier leerstijlen de voorkeur verdient. Op basis van de resultaten van hoofdstuk 2, 5 en 7 was de conclusie dan ook dat wat de ILS meet intern consistent is en door de tijd heen op hetzelfde uitkomt.

Over de *geschiktheid* van de ILS om *studiegedrag* te meten zijn meer twijfels. Hoe beter een test bepaald gedrag buiten de testsituatie voorspelt, des te hoger is de validiteit. Wat het (meta)cognitieve deel van de ILS betreft werd in de validatiestudies vrijwel geen samenhang gevonden tussen de vier leerstijlen en regulier, waarneembaar studeergedrag. De conclusie van hoofdstuk 4 is dat de ILS waarschijnlijk niet de 'echte' leerstijl van studenten meet, maar hun opvattingen over hun manier van studeren. Deze vaststelling is niet zonder gevolgen voor de mogelijkheden van de ILS als diagnostisch instrument, temeer daar de ILS inmiddels in de praktijk een populair instrument is geworden.

Wat zijn de mogelijkheden van de ILS in de *praktijk*? Onderzoek naar leerstijlen is ook bedoeld om de kwaliteit van onderwijs en leerprocessen te verbeteren. Een van de achterliggende gedachten daarbij is dat de kwaliteit van leerprocessen, de manier waarop studenten hun studie aanpakken, de kwaliteit bepaalt van de studieresultaten. In de eerste studie, het replicatie-onderzoek, werd alleen een negatieve correlatie aangetroffen tussen een ongerichte leerstijl en studiesucces. In dit verband moet opgemerkt worden dat in het replicatie-onderzoek intelligentie uitsluitend besproken wordt in verband met leerstijlen, intelligentie als voorspeller van studiesucces komt niet aan de orde.

In het eerste validatie-onderzoek werden bij studenten die hoog of laag op een bepaalde leerstijl scoorden ($N=27$), in tegenstelling tot resultaten van Vermunt (1992, hoofdstuk 7), geen significante verbanden gevonden tussen leerstijlen en tentamencijfers. In de tweede validatie-studie ($N=96$) werd wel een bescheiden empirische ondersteuning gevonden die erop wijst dat de ongerichte leerstijl een negatieve voorspeller is van studiesucces. Ook in de vergelijking met andere kenmerken van de student bleek de ILS het af te laten weten als een geschikt instrument om studiesucces te voorspellen (hoofdstuk 6).

De resultaten van het onderzoek naar de *relatie tussen leerstijlen en persoonlijkheid, intelligentie en prestatiemotivatie* lieten zien dat er weliswaar een geringe, maar systematische overlap is tussen leerstijlen en andere kenmerken van de student. Daarom heeft het volgens Busato wel zin om naast leerstijlen deze persoonlijkheidskenmerken mee te nemen. Als we echter kijken naar de voorspellende waarde van het eerste tentamen voor later studiesucces, 48 procent verklaarde variantie, ligt het voor studiebegeleiders meer voor de hand om dat als aanknopingspunt te kiezen voor hun werk. Het is zuiniger en minder belastend voor studenten dan leerstijlen- en persoonlijkheidsonderzoek.

Dit geldt temeer daar de resultaten van het onderzoek naar de effecten van een cursus begeleid studeren uitwezen dat weliswaar een kortstondige daling van de ongerichte leerstijl bewerkstelligd kon worden, maar dat dit effect na drie maanden weer was verdwenen.

In het laatste onderzoek ging Busato na of het vier-factor model het beste model is. Aanleiding daartoe was dat de vierde factor, de toepassingsgerichte leerstijl niet geheel stabiel is, en dat er bovendien theoretisch noch empirisch gronden waren om een model met meer factoren uit te sluiten. Het resultaat van een aantal exploratieve analyses was de conclusie dat een zeven-factor of een acht-factor model ook mogelijk is. Zoals gezegd, op praktische gronden gaf Busato toch maar de voorkeur aan een vier-factor model.

Commentaar

Het risico van een bespreking van het proefschrift van Busato is dat het uitloopt op een bespreking van het werk van Vermunt. Het is net een Droste-plaatje, als je de een ziet, dan zie je ook de ander. Om recht te doen aan het werk van Busato gaan we in op de vraag wat nu typisch is aan het onderzoek van Busato en wat de verdienste ervan is.

In de eerste plaats moet gezegd worden dat enige moed is vereist om een kritisch onder-

zoek te doen naar een zo bejubelde aanpak als die van Vermunt. Het proefschrift in zijn geheel in ogenschouw nemend, dan lijkt het, gezien de volgorde van de deelstudies, dat Busato in de loop van zijn onderzoek steeds meer afstand heeft genomen van de ILS. De eerste hoofdstukken staan het dichtst bij het werk van Vermunt.

Na de vaststelling op grond van het replicatie-onderzoek dat er inderdaad sprake is van vier leerstijlen, volgt direct daarop al het onderzoek naar de toepassing van de ILS in de cursus begeleid studeren. Gezien het ontbreken van een overtuigende samenhang tussen leerstijlen gemeten met de ILS en studiesucces, getuigt dit nog van vertrouwen in het instrument. Zo'n cursus wordt namelijk opgezet met het oog op prestatieverbetering. Belangrijke aannames bij de opzet van de cursus waren dat de betekenisgerichte en de toepassingsgerichte leerstijl kwalitatief de betere manieren van studeren zijn; en dat een ongerichte leerstijl zich door gerichte training laat omzetten in een betekenisgerichte leerstijl. Dat bleek niet zo te zijn, de eerste twijfels rond de geschiktheid van de ILS komen nu naar boven: heeft de student misschien een verkeerd idee van zijn studiegedrag of ligt het aan de opzet en inhoud van de cursus?

In het daarop volgende onderzoek staat de validatie centraal. Waar eerst nog de diagnostische waarde van de ILS werd genoemd als selectie-instrument voor studenten met een ongerichte leerstijl (p. 23), wordt de toon kritischer. De ILS kan maar beter worden gezien als een aanknopingspunt om sterke en zwakke punten in de manier van studeren op te sporen, dan als een instrument om studenten in te delen (p. 44).

Naarmate het onderzoek vordert, worden de onderzoeksvragen interessanter. Doel van het longitudinale onderzoek is de ontwikkeling van leerstijlen over een langere periode na te gaan. De uitkomsten zijn niet geheel eenduidig, dit komt omdat er veel details worden gepresenteerd die nu eens wel en dan weer niet overeenstemmen met de verwachtingen dat er in de tijd een toename zal optreden van meer betekenis- en toepassingsgericht leren en dat reproductief en ongericht leren minder zal voorkomen. Er worden echter nog steeds geen kernvragen gesteld: hoe stabiel zijn leerstijlen, zijn ze wel veranderbaar, en zo ja, op korte of lange termijn? Ook in de studie naar de relaties met persoonlijkheidskenmerken waarin Busato de leerstijlen toch in een aanmerkelijk bredere context plaatst, komen dergelijke vragen niet aan de oppervlakte. Lag hier niet de kans om de leerstijlen in een theoretisch model te zetten? Iets wat al na het replicatie-onderzoek had gekund. Ook deze deelstudie blijft in correlaties steken. Tot slot het onderzoek naar de begripsvaliditeit dat bij gebrek aan een theoretisch model als leidraad er maar een beetje bij bungelt.

Afgezien van deze kritische kanttekeningen blijft het de verdienste van Busato dat hij, na aanvankelijk in het kielzog van Vermunt te hebben gewerkt, op basis van zijn eigen bevindingen zich langzamerhand loszingt van het idee dat met de ILS een instrument is verschenen dat de onderwijspraktijk met een gerust hart kan gebruiken.

Hoewel het begrip 'leerstijl' niet meer weg te denken is uit het onderzoek naar leren, is het duidelijk dat leren ingewikkelder in elkaar zit dan in dit proefschrift is beschreven. Om ecologisch valide uitspraken over leerprocessen en leerresultaten te doen is er meer nodig. Elk leerproces speelt zich af in een context van vakinhoud, leerdoelen, instructie. In essentie is studiesucces het resultaat van de confrontatie tussen deze contextkenmerken en bepaalde persoonlijkheidskenmerken en eerdere leerervaringen die een student met zich meebrengt. Uit alle omschrijvingen is duidelijk dat 'leerstijl' wordt opgevat als een relatief stabiele leeraanpak en manier van denken. Wat, gezien de definitie van het begrip leerstijl en ook de resultaten van Busato, in (vervolg)onderzoek nodig is, is dat aan éénzelfde groep studenten meer en verschillende typen taken gegeven worden. Dit om na te gaan of studenten, afhankelijk van de context – vakinhoud, instructie en leerdoel – verschillende 'leerstijlen' gebruiken. Dan wordt met behulp van verschillende reële studietaken inzicht verkregen of er één leerstijl is die domineert. Er zijn aanwijzingen uit onderzoek dat goede studenten, studenten met goede studieresultaten, geen vaste leerstijl hebben, maar dat ze hun verwerkings- en regulatiestrategieën afstemmen op de taak. Dat ze een helder taakbewustzijn hebben en daar hun aan-

pak op afstemmen. Met een vragenlijst die niet is afgestemd op verschillende taken is dat dus niet te meten.

Een andere kwestie is de steekproef. We gaan er slechts kort op in, omdat Busato het zelf ook niet onder stoelen of banken steekt dat de resultaten van zijn onderzoeken bij een beperkte steekproef van studenten psychologie verkregen. Wie de pretentie heeft om uitspraken te doen over validiteit, weet dat zo' beperkte steekproef geen sterk punt is (Van den Brink en Mellenbergh, 1998).

Het onderzoek naar leerstijlen in een bredere context

De resultaten en conclusies van Busato zijn een aanmoediging voor onderzoekers om zich nog meer te verdiepen in de vele vragen over de wijze waarop leerprocessen verlopen en studiesucces tot stand komt. Voor gebruikers van de ILS in de onderwijspraktijk is er alle aanleiding tot terughoudendheid.

De ILS wat kun je ermee? Diagnostisch gezien is de ILS bruikbaar als instrument om opvattingen over studeren te meten, als basis voor gerichte studiebegeleiding komt de ILS al minder in aanmerking; als selectie-instrument is de ILS af te raden; en het voorspellen van studiesucces met de ILS is niet aan de orde. Naar aanleiding van het proefschrift van Busato is onze conclusie dat uitspraken op basis van de ILS het predikaat 'evident based practice' niet verdienen.

Tot slot mag niet onvermeld blijven dat het proefschrift van Busato bijzonder plezierig leest.

Literatuur

- Brink, W.P. van den & Mellenbergh, G.J. (1998). *Testleer en testconstructie*. Amsterdam: Boom.
- Elshout, J.J. & Akkerman, A.E. (1975). *Vijf Persoonlijheids-factoren Test SPFT*. Nijmegen: Berkhout Nijmegen B.V.
- Tupes E.C. & Christal, R.E. (1992). Recurrent personality factors based on trait rating. *Journal of personality*, 60, 225-251 (original work published in 1961).
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs*. Amsterdam/Lisse: Swets & Zeitlinger.

Anneke van der Hoeven-van Doornum
 ITS, Katholieke Universiteit Nijmegen
 Postbus 9048, 6500 KJ Nijmegen
 e-mail: A.vd.Hoeven@its.kun.nl

Redactionele medewerkers

Aan jaargang 23 en 24 werd redactionele medewerking verleend door:

Aarnoutse, C.A.J.	Katholieke Universiteit Nijmegen
Admiraal, W.	IVLOS Universiteit Utrecht
Akker, J.J.H. van den	Universiteit Twente Enschede
Ax, J.	Universiteit van Amsterdam
Batenburg, Th.A. van	GION Groningen
Beijaard, D.	ICLON Universiteit Leiden
Bergh, H. van den	Universiteit Utrecht
Boer, P. den	GION Groningen
Bosker, R.J.	Universiteit Twente Enschede
Brandsma, T.F.	OCTO Universiteit Twente Enschede
Clason, C.E.	Rijksuniversiteit Groningen
Cohen-Schotanus, J.	Rijksuniversiteit Groningen
Dam, G. ten	Universiteit van Amsterdam
Dekkers, H.P.J.M.	ITS Nijmegen
Dijkstra, S.	Universiteit Twente Enschede
Driel, J. van	ICLON Universiteit Leiden
Dronkers, J.	SCO-Kohnstamm Instituut Amsterdam
Elshout-Mohr, E.	SCO-Kohnstamm Instituut Amsterdam
Erkens, G.	Universiteit Utrecht
Esch, W. van	ITS Nijmegen
Frietman, J.	ITS Nijmegen
Glebbeeck, A.C.	Rijksuniversiteit Groningen
Glopper, K. de	SCO-Kohnstamm Instituut Amsterdam
Grift, W. v.d.	Inspectie van het Onderwijs Utrecht
Guldemond, H.	GION Groningen
Haan, D. de	OPTIC Open Universiteit
Harms, T.	GION Groningen
Hartog, J. den	Universiteit van Amsterdam
Hoeben, W.Th.J.G.	GION Groningen
Hofman, R.	GION Groningen
Hofman, W.H.A.	RISBO Erasmus Universiteit Rotterdam
Hofstee, W.K.B.	Rijksuniversiteit Groningen
Hox, J.J.	SCO-Kohnstamm Instituut Amsterdam
Imants, J.	Universiteit Leiden
Jansen, E.	GION/COWOG Groningen
Jansen, G.G.H.	Rijksuniversiteit Groningen
Jong R. de	GION Groningen
Jong, U. de	SCO-Kohnstamm Instituut Amsterdam
Jörg, T.	IVLOS Universiteit Utrecht
Jungbluth, P.	ITS Nijmegen
Kanselaar, G.	Universiteit Utrecht
Kiers, H.A.L.	Rijksuniversiteit Groningen
Kuhlemeijer, T.	CITO Arnhem
Leij, D.A.V. van der	Vrije Universiteit Amsterdam
Leseman, P.	Universiteit van Amsterdam
Leune, J.M.G.	Erasmus Universiteit Rotterdam

- | | |
|------------------------|-----------------------------------|
| Lodewijks, H.G. | Katholieke Universiteit Brabant |
| Maas, C.J.M. | Universiteit Utrecht |
| Maeyer, S. De | Universiteit Antwerpen |
| Mandemaker, K. | Vrije Universiteit Amsterdam |
| Meijer, J. | SCO-Kohnstamm Instituut Amsterdam |
| Meijer, R.R. | Universiteit Twente Enschede |
| Meijer, W. | Rijksuniversiteit Groningen |
| Mellenberg, G.J. | Universiteit van Amsterdam |
| Mulder, L. | ITS Nijmegen |
| Oosterloo, S. | Universiteit Twente Enschede |
| Peer, D. van | Rijksuniversiteit Groningen |
| Peetsma, T.T. | SCO-Kohnstamm Instituut Amsterdam |
| Peschar, J.L. | Rijksuniversiteit Groningen |
| Pijl, Y.J. | GION Groningen |
| Pilot, A. | Universiteit Utrecht |
| Prins, J. | IOWO Nijmegen |
| Roeleveld, J. | SCO-Kohnstamm Instituut Amsterdam |
| Ruijsenaers, A.J.J.M. | Universiteit Leiden |
| Sanden, J.M.M. van der | Katholieke Universiteit Brabant |
| Schoot, E. van | SCO-Kohnstamm Instituut Amsterdam |
| Severiens, S. | Universiteit van Amsterdam |
| Snijders, T.A.B. | Rijksuniversiteit Groningen |
| Stokking, K.M. | Universiteit Utrecht |
| Streumer, J.N. | Universiteit Twente Enschede |
| Sühre, C. | GION Groningen |
| Tesser, P. | SCP Rijswijk |
| Tomich, W. | Open Universiteit Heerlen |
| Vandenbergh, R. | Katholieke Universiteit Leuven |
| Veenman, S.A.M. | Katholieke Universiteit Nijmegen |
| Verhelst, N.D. | CITO/Universiteit Twente Enschede |
| Vermeer, H. | Universiteit Leiden |
| Vermunt, J.D.H.M. | ICLON Universiteit Leiden |
| Vos, P. | Universiteit Twente |
| Vrieze, G. | ITS Nijmegen |
| Waslander, S. | Twijnstra & Gudde Amersfoort |
| Willemsen, T.M. | Katholieke Universiteit Brabant |
| Witziers, B. | Universiteit Twente Enschede |