



Pedagogische studiën 1999 : Tijdschrift voor onderwijskunde en opvoedkunde

<https://hdl.handle.net/1874/205431>

P E D A

G O G I

S C H E

studiën

**Tijdschrift voor
onderwijskunde en
opvoedkunde**

Jaargang 76

Nummer 1/1999

Van de redactie

Vaksecties en leerprestaties

**Vragenlijst 'Zorg voor
leerlingen met onderwijs-
leerproblemen'**

**Het concept activerende
instructie gemeten**

Discussie

**'Kwaliteitsbeoordeling van
scholen'**

- Naar een ander systeem
Validiteit en nuttigheid
openbare prestatie-
gegevens
- Het betere blijft vijand
van het goede

Boekbespreking

Mededelingen

WOLTERS-NOORDHOFF

Pedagogische Studiën

Jaargang 76

Nummer 1/1999

Van de redactie 1

J.B. Kuhlemeier en H. van den Bergh
Vaksecties en leerprestaties bij
aardrijkskunde in het voortgezet
onderwijs 2

E. Maes, R. Vandenberghen en
P. Ghesquière
Ontwikkeling en validering van de
vragenlijst 'Zorg voor leerlingen met
onderwijsleerproblemen' 21

R. Lamberigts, P.J. den Brok, K. Derksen
en T. Bergen
Het concept activerende instructie
gemeten via de perceptie van
leerlingen 36

Discussie

M.P.C. van der Werf en
B.P.M. Creemers
Naar een vollediger systeem van
kwaliteitsbeoordeling van scholen 51

R.J. Bosker en J. Scheerens
Openbare prestatiegegevens van scho-
len; nuttigheid en validiteit 61

J. Dronkers
Het betere blijft nog steeds de vijand
van het goede. Een antwoord op Van
der Werf en Creemers, en Bosker en
Scheerens 74

Boekbespreking 83

Mededelingen 86

Pedagogische Studiën

Jaargang 76

Nummer 2/1999

Van de redactie 89

J.J. Hox

De standaardmethode voorbij.

Introductie op het thema 90

E.D. de Leeuw

Vragen naar gevoelige informatie:
een overzicht 92

M.J.C. van Hattum en S. Kef

Data verzamelen bij basisschool-
leerlingen en adolescenten met een
visuele beperking met behulp van de
computer 104

F. Wester en V. Peters

Kwalitatieve analyse van interviews.
Fasen, stappen en computergebruik in
intensief onderzoek 117

J.J. Hox

Computerintensieve data-analyse:
Bootstrap- en randomisatiemethoden
in de statistiek 132

Discussie

Aad Goddijn

Realiteit van wiskundeonderwijs VWO
ter discussie 142

Frans Keune

Kwaliteit wiskundeonderwijs VWO ter
discussie 151

Aad Goddijn

Reactie 155

Pedagogische Studiën

Jaargang 76

Nummer 3/1999

S. Banis, D.R. van Peer, &

T.P.B.M. Suurmeijer

Pedagogische onzekerheid en over-
bescherming bij ouders van kinderen
met hemofilie: verklarende factoren en
gevolgen 157

D. Uerz, H. Dekkers, & J. Dronkers

Reken- en taalvaardigheid als voor-
speller van B-keuzen in het voortgezet
onderwijs 170

L. Verschaffel, E. de Corte, & S. Lasure

Realistisch modelleren en oplossen van
Problematische wiskundige
toepassingsopgaven op de basisschool
Een review van het recent onderzoek
183

A.J.M. Vosse

Effecten en implementatie van een
tutorprogramma voor risico-
leerlingen 201

Boekbesprekingen 211

Mededelingen 217

Pedagogische Studiën

Jaargang 76

Nummer 4/1999

A.A.M. Houtveen, E.C. Roelofs, &
J.N. Streumer

Onderwijs en leeropbrengsten in de
Basisvorming: een tussenbalans.

Inleiding op het themanummer 221

H. van den Bergh, M. Peters-Sips, &
M. Zwarts

Deelstudies in het kader van de
evaluatie van de Basisvorming 224

E.C. Roelofs, & A.A.M. Houtveen

Didactiek van authentiek leren in de
Basisvorming. Stand van zaken bij
docenten Nederlands en wiskunde 237

E.C. Roelofs, H.A.M. Franssen, A.A.M.
Houtveen, & N.A.J. Lagerweij

Een dieptestudie naar didactisch
handelen in de Basisvorming: docent-
gedrag, methodengebruik en leerling-
percepties 258

G. van der Werf, M. Lubbers, &
H. Kuiper

Onderwijsopbrengsten en onderwijs-
kansen voor en na invoering van de
basisvorming 273

Kroniek

H. Dekkers (red.)

The annual meeting of the American
Educational Research Association, april,
29-23, 1999, Montreal. On the thresh-
hold of the 21st century 289

Mededelingen 306

Pedagogische Studiën

Jaargang 76

Nummer 5/1999

A. Kassenberg en R. Bosman
Schoolproblemen van basisschool-
leerlingen en de uiteenlopende
perceptie van ouders en
leerkrachten 307

J.H. van Huizen en J.W. Veerman
Leerkrachtbeoordelingen van
emotionele en gedragsproblemen bij
kinderen van een Pedagogisch
Instituutsschool 320

G.W.J.M. Driessen en J.J. Bezemer
Achtergronden, kenmerken en
opbrengsten van islamitisch basis-
onderwijs 332

M.M. Pool, T.H.A. van der Voort, J.W.J.
Beentjes en C.M. Koolstra
De invloed van achtergrondtelevisie op
de uitvoering van gemakkelijke en
moeilijke huiswerktaken 350

Kroniek

C.C. Wolhuter
Sociaal-wetenschappelijke literatuur
over onderwijs in Zuid-Afrika van ver-
zuiling tot eensgezindheid vanuit ver-
scheidenheid 36*

ORD '99 37*

Mededelingen 395

Afscheid mevrouw Hamburg

Met ingang van 1999 is mevrouw C. Hamburg teruggetreden als bureauredactrice van Pedagogische Studiën. In 1971 trad zij als bureauredactrice in dienst, dus zij heeft deze sleutelrol bij ons tijdschrift niet minder dan 28 jaar vervuld.

Mevrouw Hamburg was verantwoordelijk voor de totale afhandeling van zowel de kopij als het administratief functioneren van de redactie. Dus ook de afhandeling van correspondentie met de auteurs (uiteraard in samenwerking met de redactiesecretaris) en de 'nazorg' voor de in drukproeven verschenen kopij behoorde tot haar takenpakket.

Zij heeft haar werkzaamheden al deze jaren met de grootst mogelijke zorgvuldigheid uitgevoerd. Hoewel wellicht weinig auteurs haar 'in levenden lijve' ontmoet zullen hebben, was haar naam inmiddels bij een groot deel van onderwijskundig Nederland bekend. Hoewel zij maar een beperkte aanstellingsomvang had, voelde zij zich kennelijk dermate verbonden met het functioneren van het tijdschrift dat zij vrijwel iedere werkdag op haar werkplek te vinden was.

Vanaf het allereerste begin bevond deze werkplek zich bij de Onderwijsbegeleidingsdienst te Arnhem, de dienst waarvan het toenmalige redactielid, Dr. K. Doornbos, destijds directeur was en waaraan ook Drs. S. de Witt, de per 1 januari 1971 benoemde redactiesecretaris, als stafid verbonden was. De OBD Arnhem heeft het tijdschrift al die jaren tegen een slechts geringe vergoeding onderdak geboden.

Niet alleen kende een groot deel van onderwijskundig Nederland en Vlaanderen de naam van mevrouw Hamburg, omgekeerd vertoonde mevrouw Hamburg een enorme kennis van de 'onderwijskundige kaart van Nederland'. In mijn tijd als redactiesecretaris heb ik meermaalen ervaren dat zij goed op de hoogte was van de onderzoeksgebieden waarop men zich bij de diverse instituten richtte, een type kennis dat men niet snel elders bij een bureauredactrice zal aantreffen. Wanneer door een bepaalde auteur een manuscript werd aangeboden, wist zij niet zelden aan te geven of en wanneer door deze auteur op hetzelfde terrein eerder een artikel was aangeboden, welke de beoordelaars waren geweest, en dergelijke. Kortom, zij beschikt over een fenomenaal geheugen, hetgeen, gecombineerd met haar uiterst consciëntieuze aanpak, voor iedere redactiesecretaris het werken met mevrouw Hamburg tot een genoegen maakte.

Hoewel wij er als redactie uiteraard begrip voor hebben dat zij op haar 78 jarige leeftijd het vaandel wil overdragen, realiseren wij ons dat met mevrouw Hamburg een bureauredactrice bij Pedagogische Studiën verdwijnt die niet eenvoudig te vervangen zal zijn. De redactie is haar diep dankbaar voor het vele voortreffelijke werk dat zij al deze jaren voor het tijdschrift gedaan heeft.

Namens de redactie,
Prof. dr. N. Verloop

Vaksecties en leerprestaties bij aardrijkskunde in het voorgezet onderwijs

J.B. Kuhlemeier en H. van den Bergh

Samenvatting

Recente schoolorganisatorische theorieën wijzen op het belang van vaksecties voor het verhogen van de kwaliteit van het onderwijs. Tot de middelen die secties daartoe kunnen aanwenden behoren coördinatie via overleg, via standaardisatie van procedures, output, kennis en vaardigheden, via gemeenschappelijke opvattingen, waarden en normen, en via onderwijskundig leiderschap. In dit artikel is nagegaan welke bijdrage vaksecties aardrijkskunde leveren aan de aardrijkskundeprestaties (rekening houdend met verschillen tussen secties in het type leerlingen waarmee zij te maken hebben). De bijdrage van de vaksectie aardrijkskunde aan de aardrijkskundeprestaties van derdeklassers blijkt klein. In het VWO, HAVO, MAVO en LBO verklaren de sectiekenmerken respectievelijk 1%, 1%, 0% en 2% van de totale variantie (gegeven de achtergrond- en instroomkenmerken en de voorkennis aan het begin van het derde leerjaar). Daarnaast is nagegaan in hoeverre de gevonden kenmerken van effectieve secties consistent zijn over opleidingstype (VWO, HAVO, MAVO en LBO) en schoolstructuur (categoriale scholen, smalle scholengemeenschappen en brede scholengemeenschappen). Vrijwel alle gevonden sectie-effecten blijken of beperkt tot één opleidingstype of zijn inconsistent in de zin dat de richting ervan van opleidingstype tot opleidingstype verschilt. Slechts één kenmerk heeft in twee opleidingstypen een vergelijkbaar effect. Het betreft de mate waarin de directie (volgens de sectieleider) invloed uitoefent op het handhaven van de regels en de orde. Op scholen waar de directie actieve bemoeienis heeft met de leefregels in de klas, met ordeproblemen in specifieke klassen en met de registratie en aanpak van spijbelaars presteren de leerlingen relatief hoog. Daarnaast blijkt geen enkel sectiekenmerk zowel effectief in categoriale scholen als in smalle en brede scholengemeenschappen.

1 Inleiding

Onderzoek naar schooleffectiviteit heeft aanemelijk gemaakt dat het schoolbeleid de leerprestaties van de leerlingen slechts in beperkte mate beïnvloedt (o.a. Creemers, 1991; Scheerens, 1989; Wang, Haertel & Walberg, 1993). Een van de mogelijke verklaringen is de lange causale keten tussen het schoolbeleid en het leren in de klas. Dit vestigt de aandacht op de mogelijke rol van het middenmanagement: de hiërarchische laag tussen directie en vakdocenten. In het schoolorganisatieonderzoek is het middenmanagement nog grotendeels over het hoofd gezien (Smets, 1986a; Turner, 1996; Worner & Brown, 1993). In het Nederlandse voorgezet onderwijs kan een onderscheid worden gemaakt in de meer algemeen-organisatorische en de meer vakgericht-inhoudelijke middenmanagers. Tot de eerste categorie behoren de leerjaar- en afdelingscoördinatoren en tot de tweede categorie de sectieleiders voor de verschillende vakgebieden. Tot de taken van middenmanagers behoren onder meer de coördinatie van secties, afdelingen en leerjaren, leerlingbegeleiding, stages, onderwijskundige vernieuwing en begeleiding van collega's (Gooren, 1989; Smets, 1986b).

Hylkema (1990) en Marx (1992) onderscheiden drie typen vaksecties van toenemende organisatorische zwaarte. Allereerst heeft de sectie een sociale of samenlevingsfunctie. De sectie is een informeel contactorgaan waar vakspecialisten met vergelijkbare belangstelling, expertise en ervaring informatie uitwisselen en elkaar ondersteunen (Siskin, 1991; Van Wesum, 1997a). Daarenboven kan de sectie ook een (adviserend) staforgaan zijn en een intermediërende rol vervullen bij het implementeren van het schoolbeleid (HMI, 1984; Turner, 1996). De sectieleider is dan een schakel tussen enerzijds de sectieleden en anderzijds de directie en andere vaksecties. Ten slotte kan de sectie ook een onderwijskundig sturingsorgaan

zijn. De sectieleider vervult dan een sturende en bewakende rol met betrekking tot de vakinhoudelijke, vakdidactische en organisatorische aspecten van het lesgeven. Tot het aandachtsgebied van de vaksectie behoren onder meer de onderwijsdoelen, leergangen en lesmateriaal, te behandelen leerstof, leer- en hulpmiddelen, onderwijsvernieuwing, proefwerken en normering, huiswerk, inwerken van nieuwe collega's, regelen van onderlinge vervanging en beheer van het sectiebudget (voor een inventarisatie: zie Worner & Brown, 1993).

De vaksectie wordt geacht van vitaal belang te zijn voor het welbevinden van de docenten (Hylkema, 1990), de kwaliteit van het onderwijs en de leerresultaten van de leerlingen (Bosker & Witziers, 1995; Luyten, 1994; Van Marwijk-Kooy, 1984; Scheerens & Bosker, 1997; Turner, 1996; Van Wessum, 1997a). Het belang van de sectie voor de leerprestaties van de leerlingen is nog nauwelijks onderzocht.

Wel heeft onderzoek aangetoond dat een aanzienlijk deel van de effectiviteitsverschillen tussen scholen is toe te schrijven aan verschillen tussen vakken binnen scholen. Zo analyseerde Luyten (1994) de gemiddelde examenresultaten van scholen voor voortgezet onderwijs in zeventien vakken over een reeks van jaren. Er bleken grote effectiviteitsverschillen te bestaan tussen de verschillende vakken die binnen een school onderwezen worden. Van de totale variatie in de schoolgemiddelden was maar liefst 40% toe te schrijven aan verschillen tussen vakken binnen scholen; slechts 25% kwam voor rekening van de algemene schooleffecten (d.w.z.: schooleffecten die consistent waren over vakken en stabiel in de tijd). Het gegeven dat de verschillen tussen vakken binnen scholen groter zijn dan de verschillen tussen scholen wijst volgens Luyten (1994) in de richting van de vaksectie als een belangrijke determinant van leerprestaties (zie ook Scheerens & Bosker, 1997).

De hamvraag is nu welke kenmerken van secties voor de veronderstelde positieve bijdrage aan de leerprestaties van de leerlingen verantwoordelijk zouden zijn. In dit verband wordt regelmatig gewezen op het belang van coördinatie (o.a. Creemers, 1991; Levine & Lezotte, 1990; Scheerens, 1989) waarbij zes mechanismen worden onderscheiden (Mintzberg, 1979, 1989; Hofman & Lugthart, 1991; Smets, 1986b).

Een eerste mechanisme dat vaksecties kunnen toepassen is coördinatie via overleg en overreding (o.a. Hamlin, 1990; Turner, 1996). Indicatoren zijn onder meer de frequentie van het overleg binnen de sectie, het overleg met andere secties en het overleg van de sectieleider met de directie.

Een tweede mechanisme is coördinatie via standaardisatie van procedures. Daarvan is onder meer sprake als secties afspraken maken en regels opstellen over de te behandelen leerstof, het behandeltempo, de inhoud van gemeenschappelijke proefwerken, de normering en de aard en omvang van het huiswerk.

Een derde mechanisme is standaardisatie van output. Vaksecties treffen dan maatregelen teneinde de leerlingprestaties te verhogen, bijvoorbeeld door middel van een leerlingvolgsysteem of een integraal systeem van leerlingbegeleiding. Van standaardisatie van output is ook sprake als een vaksectie bindende afspraken maakt die tot doel hebben de prestatieverschillen tussen parallelklassen te minimaliseren.

Een vierde mechanisme is coördinatie via onderwijskundig leiderschap (direct leiding geven). Sterk onderwijskundig leiderschap behoort tot de kenmerken waarop effectieve scholen zich zouden onderscheiden van minder effectieve (o.a. Creemers, 1991; Hamlin, 1990; Levine & Lezotte, 1990; Scheerens, 1989). School- en sectieleiders geven bijvoorbeeld blijk van onderwijskundig leiderschap als zij directe bemoeienis hebben met het lesgeven.

Een vijfde mechanisme is coördinatie via standaardisatie van kennis en vaardigheden. Voorbeelden zijn het volgen van na- of bijscholing, het houden van functioneringsgesprekken of het bijwonen van elkaars lessen (Earley & Fletcher-Campbell, 1989).

Een zesde mechanisme waarvan secties gebruik kunnen maken is coördinatie via gemeenschappelijke opvattingen, waarden en normen (Turner, 1996). Het verkrijgen van consensus binnen het schoolteam wordt beschouwd als een belangrijke voorwaarde voor een effectief schoolbeleid en een effectief instructieproces in de klas (o.a. Creemers, 1991; Levine & Lezotte, 1990). Van teamhomogeniteit is bijvoorbeeld sprake als de sectieleden het in hoge mate met elkaar eens zijn over doelstellingen, inhoud, werkvormen, evaluatie en discipline.

Uit het onderzoek van Witziers (1992; 1993a,b) onder vaksecties wiskunde, Engels en geschiedenis komt naar voren dat secties zich vooral bezig houden met de leerdoelen, de leerstof en de evaluatie. Andere terreinen, zoals de orde in de klas en de pedagogisch-didactische aspecten van het lesgeven, zijn nog vrijwel geheel het domein van de individuele docent. Tevens bleken docenten binnen secties vrij homogeen in hun opvattingen over het (vak-)onderwijs. Al met al concludeert Witziers (1993a, p. 16) dat 'secties vrij hechte organen zijn en dat de coördinatie binnen secties meer voorstelt dan op grond van de literatuur aangenomen mocht worden'.

Hoe leerlingen presteren, zou direct gerelateerd zijn aan de wijze waarop vaksecties hun onderwijs plannen, uitvoeren en evalueren (HMI, 1984). Witziers (1992; 1993b) onderzocht de effectiviteitsvraag bij 39 opleidingen voor MAVO en HAVO. Secties wiskunde bleken voor de wiskundeprestaties van belang te zijn. Directies en wiskundesecties met een sterk ontwikkeld onderwijsbeleid bewerkstelligden bij hun leerlingen relatief hoge wiskundeprestaties. De mate waarin de sectieleider als teamleider optrad en de mate van samenwerking en overleg tussen de leden van de sectie bleken daarentegen negatief met de wiskundeprestaties samen te hangen. Andere sectiekenmerken, waaronder de consensus binnen het team, waren niet op de prestaties van invloed. In Groot-Brittannië onderzochten Harris, Jamieson en Russ (1995) de kenmerken van vaksecties waar de leerlingen hogere prestaties behaalden dan op grond van hun achtergrond verwacht mocht worden. Uit de bij sectieleiders, docenten en leerlingen afgenomen interviews bleek onder meer dat effectieve secties regelmatig en doelgericht vergaderen. Bovendien bleken zij veel tijd en energie te steken in de keuze en behandeling van de leerstof, het registreren en evalueren van de prestaties en het geven en controleren van huiswerk.

In dit artikel staat de vaksectie aardrijkskunde centraal. De gegevens zijn afkomstig uit een landelijk peilingsonderzoek naar de inhoud en het niveau van het vak aardrijkskunde in het derde leerjaar van het voortgezet onderwijs (Kuhlemeier, Van den Bergh, Notté, Wagenaar, Verstralen & Cappers, 1994). In deze peiling is de sectieleiders aardrijkskunde gevraagd

naar een aantal kenmerken van hun sectie als sturingsorgaan (Kuhlemeier & Van den Bergh, 1994c). Gevraagd is naar de coördinatie via intern en extern overleg, via standaardisatie van procedures, via standaardisatie van output, via onderwijskundig leiderschap en via gemeenschappelijke opvattingen (vgl. Mintzberg, 1979; 1989). Daarnaast is informatie verzameld over onder meer de sociaal-etnische achtergrond en de schoolloopbanen van de leerlingen (Kuhlemeier & Van den Bergh, 1994b). In dit artikel worden de kenmerken van de sectie aardrijkskunde in verband gebracht met de aardrijkskundeprestaties, rekening houdend met verschillen tussen secties in het type leerlingen waarmee zij te maken hebben.

Lange tijd is men ervan uitgegaan dat de kenmerken van effectieve scholen dezelfde zijn voor verschillende typen scholen. Echter, volgens het zogeheten contingentieperspectief (vgl. Scheerens & Stoel, 1987) is de effectiviteit van het schoolbeleid en de instructie mede afhankelijk van omgevingsfactoren. Contextuele factoren – zoals schoolgrootte, schoolstructuur, opleidingstype, urbanisatiegraad, denominatie en de 'kwaliteit' van de instroom – kunnen daarbij een faciliterende dan wel belemmerende rol spelen. Wat 'werkt' in de ene context hoeft niet te 'werken' in de andere context (vgl. Brandsma & Knuver, 1991). In dit artikel wordt tevens nagegaan in hoeverre de kenmerken van hoog en laag effectieve secties consistent zijn over opleidingstype (VWO, HAVO, MAVO en LBO) en schoolstructuur (categoriale scholen, smalle scholengemeenschappen en brede scholengemeenschappen)¹.

2 Methode van onderzoek

2.1 Steekproef en proefpersonen

De gegevens over de sectie, de leerprestaties, en de achtergrondkenmerken van de leerlingen zijn verzameld in het kader van een grootschalig landelijk peilingsonderzoek naar de inhoud en het niveau van het vak aardrijkskunde (Kuhlemeier et al., 1994). Het onderzoek is uitgevoerd in het derde leerjaar van opleidingen voor VWO, HAVO, MAVO en LBO. De aardrijkskundeprestaties van de leerlingen zijn zowel aan het begin als aan het einde van het schooljaar 1991-1992 gemeten.

Tabel 1

Verdeling van scholen, klassen, en derdeklassers in de totale responsgroep, per opleidingstype en per schoolstructuur

	Scholen	Klassen	Leerlingen
Opleidingstype			
VWO	35	107	2513
HAVO	37	99	2302
MAVO	45	153	3423
LBO	24	88	1167
Schoolstructuur			
categoriaal	30	121	2396
smalle sg	55	191	3622
brede sg	56	154	3387
Totaal	141	447	9405

Aan het einde van het schooljaar 1991-1992 zijn in totaal 233 sectievragenlijsten verzonden. Deze zijn geadresseerd aan de 233 scholen die tijdens de werving in het voorjaar van 1991 te kennen gaven aan de begin- en eindmeting te willen deelnemen. Van de 233 potentiële respondenten heeft 91 procent ($N = 211$) de vragenlijst in verwerkbaar vorm geretourneerd. Van de 211 scholen beschikten er 156 over een sectie aardrijkskunde (eventueel geïntegreerd met een of meer andere vakken). Tabel 1 toont de aantallen scholen, klassen en leerlingen die uiteindelijk beschikbaar waren voor de analyse van de relaties tussen de sectiekenmerken en de aardrijkskundeprestaties aan het begin en einde van het derde leerjaar².

2.2 De sectie aardrijkskunde

De kenmerken van de sectie aardrijkskunde zijn geïnventariseerd met een vragenlijst. Als er een sectie aardrijkskunde op de school aanwezig was, is de vragenlijst ingevuld door de sectieleider aardrijkskunde. Gevraagd is naar diverse in de literatuur onderscheiden coördinatiemechanismen (Mintzberg, 1979, 1989; Smets, 1986b).

Het eerste gedeelte van de vragenlijst bevat tien vragen die een algemene typering van de sectie aardrijkskunde mogelijk maken. Deze vragen zijn vooral indicatief voor coördinatie via overleg. Allereerst is gevraagd naar de aanwezigheid van een vaksectie aardrijkskunde op school (eventueel geïntegreerd met één of meer andere vakken). De overige vragen gaan over de omvang van de sectie, het aantal leden dat doorgaans op de sectievergadering aanwezig

is, de frequentie en tijdsduur van het sectieoverleg, de frequentie en aard van de contacten tussen enerzijds de sectie(leider) aardrijkskunde en anderzijds de directie en andere vaksecties, het vergaderen op een vast tijdstip en op een vaste plaats en de mate van structurering (al of geen voorzitter, agenda, verslag, afsprakenlijst, vaste werkverdeling).

Het tweede deel van de vragenlijst gaat vooral over de coördinatie via standaardisatie van procedures, regels en output. Gevraagd is naar de aandacht voor vakinhoudelijke en vakdidactische onderwerpen, de leerlingbegeleiding, de discipline in de klas, de coördinatie tussen docenten, klassen en leerjaren, en de regels waaraan ieder sectielid geacht wordt zich te houden.

In het derde deel van de vragenlijst is gevraagd naar de mate van overeenstemming tussen de sectieleden over de vakinhoud, de vakdidactiek, de evaluatie, de discipline, de invoering van de basisvorming en de vak- en leerstofinhoudelijke afstemming van het aardrijkskundeonderwijs. Deze vragen zijn indicatief voor de coördinatie via gemeenschappelijke opvattingen, waarden en normen.

Het vierde deel gaat over coördinatie via onderwijskundig leiderschap. In dit laatste deel is de sectieleiders gevraagd naar hun mening over de invloed van de directie op het aardrijkskundeonderwijs in het derde leerjaar van het opleidingstype waarvoor de school aan het onderzoek deelnam.

De antwoorden op de 83 vragen uit het tweede, derde en vierde gedeelte van de vragenlijst zijn middels principale componenten-

analyse ingedikt tot negen schalen. Deze kunnen als volgt worden beschreven:

- 1) De aandacht in het sectieoverleg voor leerlingbegeleiding: de vorderingen en aanpak van individuele, zwakke en 'snelle' leerlingen.
- 2) De aandacht in het sectieoverleg voor de vak- en leerstofinhoudelijke afstemming van de te behandelen leerstof, het gebruik van eigen lesmateriaal, het behandeltempo, de inhoud van proefwerken en de taakverdeling tussen de sectieleden onderling.
- 3) De mate waarin de sectie eigen regels heeft opgesteld op het terrein van proefwerken en huiswerk: het aankondigen van een proefwerk, het tijdstip van aankondiging, de aard en omvang van het huiswerk, het bespreken van schriftelijk werk en de correctietermijn.
- 4) De mate waarin de sectieleden overeenstemmen over de leerlingbegeleiding: de vorderingen en aanpak van individuele, zwakke en snelle leerlingen.
- 5) De mate waarin de sectieleden overeenstemmen over de vak- en leerstofinhoudelijke afstemming van het aardrijkskunde-onderwijs: onderwijsdoelen, leergang, te

behandelen leerstof/onderwerpen, behandeltempo, proefwerken, afstemming tussen leerjaren.

- 6) De mate waarin de sectieleden overeenstemmen over de orde en regels: de leefregels in de klas, ordeproblemen in specifieke klassen en de registratie en aanpak van spijbelaars.
- 7) De (gepercipieerde) invloed van de directie op de leerlingbegeleiding.
- 8) De (gepercipieerde) invloed van de directie op de vak- en leerstofinhoudelijke aspecten van het aardrijkskunde-onderwijs.
- 9) De (gepercipieerde) invloed van de directie op de handhaving van orde en regels.

De eerste drie schalen zijn indicatief voor coördinatie via overleg en via standaardisatie van procedures en output. De schalen 4 t/m 6 betreffen de teamhomogeniteit en de schalen 7 t/m 9 het onderwijskundig leiderschap van de directie ten aanzien van het vak aardrijkskunde.

De tien vragen over het sectieoverleg uit het eerste gedeelte van de vragenlijst zijn als afzonderlijke variabelen gehandhaafd. Vanwege de hoge correlatie tussen de omvang van de

Tabel 2

Gemiddelde ($\bar{X}$), standaarddeviatie (sd), laagste score (min), hoogste score (max), aantal uitspraken (N_{uit}) en Cronbachs alpha (α) per sectiekenmerk

	$\bar{X}$	SD	min	max	N_{uit}	α
Algemene sectiekenmerken						
Geïntegreerde sectie	.08	.28	0	1	1	nvt
Vergaderfrequentie	5.42	6.87	0	40	1	nvt
Vergaderduur	84.61	52.33	0	240	1	nvt
Vaste tijd en plaats	.24	.43	0	1	1	nvt
Omvang sectie	3.90	1.50	2	8	1	nvt
Aantal leden op sectievergadering	3.42	1.26	0	7	1	nvt
Taakverdeling	1.90	.76	1	3	1	nvt
Frequentie overleg met directie over algemene zaken	1.68	2.16	0	20	1	nvt
Frequentie overleg met directie over het aardrijkskundeonderwijs	.71	1.62	0	15	1	nvt
Frequentie overleg met andere secties	1.95	1.05	1	5	1	nvt
Sectieoverleg over						
leerlingbegeleiding	10.95	3.31	7	21	7	.88
afstemming	21.41	4.69	9	33	9	.90
Eigen sectieregels	3.43	1.20	1	5	5	.80
Teamhomogeniteit t.a.v.						
leerlingbegeleiding	14.50	4.05	8	24	8	.96
afstemming	20.37	3.62	9	27	9	.90
orde en regels	6.09	1.72	3	9	3	.91
Invloed directie op de						
leerlingbegeleiding	9.91	2.53	8	18	8	.93
afstemming	11.36	2.54	10	28	10	.86
orde en regels	5.86	1.68	3	9	3	.86

Tabel 3

Gemiddelde (X) en standaarddeviatie (SD) op de begin- en eindmeting en de gemiddelde leerwinst per opleidingstype

	Beginmeting		Eindmeting		Leerwinst
	X_1	SD_1	X_2	SD_2	$X_2 - X_1$
VWO	109.90	13.51	117.68	14.89	7.78
HAVO	103.43	12.51	109.87	13.48	6.44
MAVO	95.83	11.46	102.27	12.94	6.44
LBO	84.14	13.80	87.14	13.99	3.00
Totaal	100.00	15.00	106.37	16.68	6.37

sectie en het aantal leden dat de sectievergadering doorgaans bijwoont, is de eerstgenoemde variabele niet in de analyse betrokken. Dit brengt het totale aantal verklarende variabelen met betrekking tot de sectie aardrijkskunde op achttien. Tabel 2 geeft onder meer informatie over het aantal uitspraken per schaal en de betrouwbaarheid in de totale responsgroep. De weergegeven betrouwbaarheidsmaat is Cronbachs alpha, een indicatie voor de interne consistentie of homogeniteit van een groep uitspraken. De betrouwbaarheid van de negen schalen is voldoende hoog om verdere analyse te rechtvaardigen, zeker wanneer we het kleine aantal uitspraken in aanmerking nemen.

2.3 De aardrijkskundeprestaties

Zes zogeheten kerntoetsen zijn zowel aan het begin als aan het einde van het derde leerjaar aan de leerlingen voorgelegd. De in totaal 67 verschillende opgaven bestrijken de kennis van negen belangrijke onderdelen van het vak aardrijkskunde. Vanwege de grote vaardigheidsverschillen tussen leerlingen van verschillende opleidingstypen zijn aparte kerntoetsen voor VWO/HAVO, MAVO en LBO geconstrueerd. Voor ieder opleidingstype waren er twee toetsversies met elk 22 à 25 opgaven. De toetsversies voor VWO/HAVO bevatten deels moeilijker opgaven dan die voor MAVO, die op hun beurt weer uit moeilijker opgaven bestaan dan die voor LBO. De helft van de leerlingen maakte bij de eindmeting dezelfde toetsversie als bij de beginmeting; de andere helft kreeg bij de eindmeting de andere toetsversie voorgelegd.

Omdat de zes kerntoetsen via ankeritems met elkaar verbonden zijn, kunnen alle opgaven in beginsel op één schaal gebracht worden. De schaling is uitgevoerd met het programma

OPLM (One Parameter Logistic Model) waarbij van zogeheten WARM-schatters gebruik is gemaakt (Verhelst, Glas & Verstralen, 1994). Van de in totaal 67 opgaven bleken er zes om psychometrische redenen verwijderd te moeten worden.

Een overzicht van de prestaties van de 141 scholen met een sectie aardrijkskunde op de begin- en eindmeting is opgenomen in Tabel 3. De geschaalde scores van de derdeklassers op de begin- en eindmeting zijn zo getransformeerd dat de scores op de beginmeting een gemiddelde van honderd en een standaarddeviatie van vijftien hebben. Voor de nameting resulteert deze lineaire transformatie in een gemiddelde van 106.37 en een standaarddeviatie van 16.68.

Tabel 3 laat onder meer zien dat het verschil tussen de gemiddelde score op de begin- en eindmeting niet voor elk opleidingstype gelijk is. In het AVO bedraagt de leerwinst zo'n zes à acht punten en in het LBO slechts drie.

2.4 Achtergrondkenmerken

De achtergrondkenmerken van de leerlingen zijn met een vragenlijst geïnventariseerd. Voor dit artikel zijn die leerlingkenmerken gekozen die in een eerdere analyse een substantiële en over opleidingstypen redelijk consistente bijdrage aan de aardrijkskundeprestaties leverden (Kuhlemeier & Van den Bergh, 1994b). Het betreft sekse, doubleergeschiedenis, advies van het hoofd van de basisschool, opleidingsniveau van de vader en de moeder, thuistaal van de ouders en van eventuele broers en zussen, geboorteland ouders, vakantiedrag en gezinssituatie (geen-moedergezin).

2.5 Statistische analyse

Als de ene sectie hogere prestaties behaalt dan de andere, hoeft dat niet per se aan de kwaliteit van het onderwijs of van de sectie te liggen. Ook de samenstelling van de leerlingbevolking op de school in kwestie speelt een rol. Op sommige scholen is de instroom 'gunstiger' dan op andere, waardoor het behalen van hoge prestaties voor de ene sectie gemakkelijker is dan voor de andere. Omwille van een eerlijke vergelijking zijn de volgende achtergrondkenmerken als covariaten in de analyse betrokken: sekse, doubleergeschiedenis, advies van het hoofd van de basisschool, opleidingsniveau van de vader en de moeder, thuistaal van de ouders en van eventuele broers en zussen, geboorteland ouders, vakantiegedrag en gezinssituatie (geen-moedergezin). Ter verdere typering van de instroom (Willms, 1992) zijn enkele naar het niveau van de klas of school geaggregeerde leerlingkenmerken gebruikt: het percentage jongens op school, het gemiddeld advies van het hoofd van de basisschool, het percentage leerlingen op school uit zogeheten 'geen-moedergezinnen' en het percentage leerlingen op school dat thuis met beide ouders een andere taal dan Nederlands spreekt. In de analyses per opleidingstype is daarnaast gecorrigeerd voor de structuur van de school, terwijl in de analyses per schoolstructuur de opleidingstypen als covariaten zijn opgenomen.

Willen we vaksecties vergelijken in termen van effectiviteit, dan zullen we ook rekening moeten houden met verschillen in de voorkennis van de leerlingen waarmee secties te maken hebben. Door de voormeting op te nemen als covariaat, worden verschillen in het bereikte eindniveau gecorrigeerd voor verschillen in de voorkennis van de leerlingen aan het begin van het derde leerjaar. Daarnaast zijn de gemiddelde score van de klas op de beginmeting en de gemiddelde score van de school op de beginmeting in de analyse betrokken.

Al met al onderscheiden we vier analysemodellen die verschillen in de mate waarin rekening wordt gehouden met de meer of minder gunstige omstandigheden waaronder secties hun werk moeten doen (vgl. Bosker, 1991; Bryk & Raudenbush, 1992; Willms, 1992). In *model A* wordt uitsluitend het effect van de sectiekenmerken op de prestaties aan het einde van het derde leerjaar (de eindmeting) geschat.

In dit model is noch rekening gehouden met de achtergrondkenmerken van de leerlingen noch met de prestaties aan het begin van het schooljaar. In *model B* wordt – behalve de sectiekenmerken – ook het effect van de achtergrond- en instroomkenmerken geschat. In *model C* wordt *model A* uitgebreid met de prestaties aan het begin van het derde leerjaar. De prestaties aan het einde van het derde leerjaar worden nu gecorrigeerd voor initiële prestatieverschillen tussen secties ten gevolge van onder meer het aardrijkskundeonderwijs in de eerste twee leerjaren. In *model D* breiden we *model C* uit met de achtergrondkenmerken van de leerlingen. De verschillen tussen secties in de prestaties aan het einde van het derde leerjaar worden nu gecorrigeerd voor zowel de beginkennis als de achtergrond- en instroomkenmerken (het zogeheten 'netto-effect' van de sectie).

De vier modellen analyseren we in elk van de vier opleidingstypen en in elk van de drie schoolstructuren¹ afzonderlijk. De relatie tussen prestaties en leerlingkenmerken hoeft immers niet voor elk opleidingstype en voor elke schoolstructuur gelijk te zijn. Bovendien kunnen relaties gecontamineerd zijn met selectie-effecten ten gevolge van de ongelijke verdeling van leerlingen over scholen voor voortgezet onderwijs (vgl. Kuhlemeier & Van den Bergh, 1994b).

De vier modellen zijn geanalyseerd met behulp van een multiple regressie-analyse (Goldstein, 1987). De analyse is uitgevoerd op de geschaalde aardrijkskundeprestaties. Gebruik is gemaakt van het programma ML3 (Prosser, Rasbash & Goldstein, 1991). In de multilevel analyse zijn drie analyseniveaus onderscheiden: a) tussen-scholen (schoolniveau), b) tussen-klassen-binnen-scholen (klasseniveau) en c) tussen-leerlingen-binnen-klassen (leerlingniveau).

Naarmate de inhoudelijke verwantschap tussen de predictoren groter is, zal men meer rekening moeten houden met multicollineariteit en eventuele over- en onderschatting van regressiegewichten als een gevolg daarvan³. Bij elk analysemodel is daarom een stapsgewijze procedure toegepast. Volgens de zogeheten backward selectieprocedure zijn de verklarende variabelen eerst alle tegelijkertijd in de analyse ingebracht. Daarna zijn niet-significante ($p > .05$) predictoren stapsgewijs uit de analyse

verwijderd (waarbij de predictoren met de naar verhouding grootste standaardfout het eerst geëlimineerd worden). De selectieprocedure is beëindigd op het moment dat er alleen significante predictoren resteren. De uitkomst van de analyse is het meest spaarzame model dat de data met zo weinig mogelijk predictoren zo goed mogelijk beschrijft.

Een belangrijke vraag is die naar het relatieve belang van de sectiekenmerken gegeven de achtergrond-, instroom- en beginkenmerken. Dit additieve of 'netto' effect verwijst naar de speelruimte die de sectie heeft om de leerprestaties via een actief beleid te beïnvloeden, gegeven de randvoorwaarden die niet of nauwelijks te veranderen zijn. Ter beantwoording van de vraag naar het relatieve belang van de sectie onderscheiden we drie analysemodellen. In het zogeheten onconditionele model (*model 1*) wordt de totale variantie op de eindmeting ontleed in drie componenten: school, klas en leerling. In *model 2* voegen we aan model 1 de achtergrond-, instroom- en beginkenmerken toe die in het eerder beschreven model D significantie te zien gaven. In *model 3* breiden we model 2 uit met de sectiekenmerken. Alleen die sectiekenmerken zijn toegevoegd die in model D significantie te zien gaven. Model 3 is derhalve identiek aan model D.

De grootte van het additieve of 'netto' effect van de sectiekenmerken evalueren we aan de hand van de toename van de toetsingsgrootheid $-2 \text{ Log Likelihood}$. Dit verschil is Chi-kwadrat verdeeld met het bijbehorende verschil in het aantal vrijheidsgraden c.q. in het aantal te schatten parameters (Bentler & Bonet, 1980). Een tweede criterium is de toename van de proportie verklaarde variantie (R^2_{totaal}) van model 2 naar model 3. R^2_{totaal} wordt berekend als het kwadraat van de correlatie tussen de geobserveerde scores en de op grond van het model voorspelde scores. Dit tweede criterium geeft inzicht in het additieve effect van de sectiekenmerken (model 3) gegeven het effect van de achtergrond-, instroom- en beginkenmerken (model 2).

3 Resultaten

3.1 De vaksectie aardrijkskunde

Op ruim één derde van de MAVO-opleidingen en bijna de helft van de LBO-opleidingen ontbreekt een sectie aardrijkskunde (al dan niet geïntegreerd met andere secties). De hieronder gepresenteerde gegevens zijn gebaseerd op de scholen die wel over een sectie aardrijkskunde beschikken.

Secties aardrijkskunde vergaderen relatief weinig. Tegen het einde van het schooljaar 1991-1992 was de doorsnee sectie ongeveer vijf keer bij elkaar gekomen. Ter vergelijking, wiskundesecties vergaderen per jaar gemiddeld veertien keer (Witziers, 1992) en secties bewegingsonderwijs gemiddeld negen keer (Kamphorst & Tuinenga, 1995). De vergaderingen van de aardrijkskundesecties duren gemiddeld bijna anderhalf uur. Tussen scholen bestaan grote verschillen in het aantal keren dat er vergaderd wordt en de tijdsduur van de vergaderingen. Op jaarbasis vergaderde ongeveer een derde van de secties minder dan drie uur en een kwart kwam ten minste acht uur bijeen.

Zo'n 70% van de secties aardrijkskunde kent geen vaste vergadertijd en -plaats. Bovendien wordt er doorgaans op informele en weinig gestructureerde wijze vergaderd (d.w.z.: zonder voorzitter, agenda, verslag, afsprakenlijst, vaste werkverdeling).

De gemiddelde sectie(leider) aardrijkskunde heeft in het schooljaar 1991-1992 nog geen twee keer overleg gevoerd met de directie. In dit overleg blijkt het aardrijkskundeonderwijs in het opleidingstype waarvoor de school aan de peiling deelneemt maar weinig ter sprake te komen. Ook voeren sectieleiders maar weinig overleg met andere secties.

De sectieleider is gevraagd naar de onderwerpen die in het sectieoverleg aan bod komen. Geen van de bevroegde onderwerpen op vakinhoudelijk, vakdidactisch, evaluatief en disciplinair gebied komt in de sectievergadering vaak aan bod. Nog het meest wordt gesproken over de invoering van de basisvorming en de vak- en leerstofinhoudelijke afstemming van het aardrijkskundeonderwijs (de keuze van de leerstof, het tempo waarin de leerstof behandeld wordt, het gebruik van eigen lesmateriaal, de inhoud en de normering van proefwerken en het huiswerk).

Tabel 4

Regressiegewichten (β) voor het effect van de sectie op de aardrijkskundeprestaties in het VWO, HAVO, MAVO en LBO in model A, B, C en D (Model A: alleen sectiekenmerken; Model B: sectie- en achtergrondkenmerken; Model C: sectie- en beginkenmerken; Model D: sectie-, achtergrond- en beginkenmerken)

	Model A		Model B		Model C		Model D	
	β	se	β	se	β	se	β	se
Voorbereidend Wetenschappelijk Onderwijs (VWO)								
Vergaderfrequentie	-.18	.09						
Sectieoverleg over afstemming			.41	.10	.21	.10	.33	.08
Eigen regels	1.15	.54						
Invloed directie op de orde en regels			1.32	.28	.60	.27	1.04	.23
Hoger Algemeen Voortgezet Onderwijs (HAVO)								
Vergaderduur	.01	.00	.01	.00	.01	.00	.01	.00
Aantal leden aanwezig					-1.48	.44		
Algemeen overleg directie			.40	.14			.36	.11
Overleg directie ak	2.67	.63	1.28	.49	1.32	.41		
Overleg met andere secties			-.74	.32				
Sectieoverleg over de leerlingbegeleiding	-.48	.21			-.42	.13	-.57	.14
Teamhomogeniteit t.a.v. leerlingbegeleiding	-.33	.13	-.35	.09				
afstemming	.63	.17	.44	.13			.26	.10
Invloed directie op de leerlingbegeleiding							.26	.11
Middelbaar Algemeen Voortgezet Onderwijs (MAVO)								
Aantal leden aanwezig	1.14	.54	1.16	.55				
Taakverdeling	-2.34	.98	-2.41	.98				
Lager Beroepsonderwijs (LBO)								
Geïntegreerde sectie					2.43	1.02	2.53	.96
Taakverdeling	3.84	.88	3.09	.74				
Overleg directie ak	-2.43	.72	-1.74	.63				
Eigen regels	-3.97	.62	-3.06	.54	-1.68	.45	-1.65	.42
Teamhomogeniteit t.a.v. leerlingbegeleiding	.67	.18	.61	.15	.36	.13	.33	.12
Invloed directie op de leerlingbegeleiding	-1.91	.57	-1.80	.48				
afstemming	2.37	.55	2.08	.47				
orde en regels	1.80	.50	1.65	.42	.54	.27	.58	.25
Categoriale scholen								
Aantal leden aanwezig	2.28	.70	1.58	.62				
Eigen regels			-1.82	.72				
Smalle scholengemeenschappen								
Vaste taakverdeling	-1.36	.60						
Overleg met andere secties					-.56	.27		
Teamhomogeniteit t.a.v. afstemming			.60	.22	.47	.18	.56	.18
Invloed directie op orde en regels	.53	.26						
Brede scholengemeenschappen								
Vergaderduur	.04	.01	.03	.01	.03	.01	.03	.01

Ten minste vier van de vijf secties hebben regels ontwikkeld voor het al dan niet aankondigen van proefwerken en het tijdstip van aan-

kondiging (zoveel dagen of weken van te voren). Meer dan de helft van de secties kent regels voor de aard en omvang van huiswerk,

het bespreken van schriftelijk werk of de correctietermijn van schriftelijk werk. Volgens de sectieleiders houden de vakdocenten zich redelijk tot goed aan deze regels.

Binnen secties bestaat weinig verschil van mening over de vakinhoud, de vakdidactiek, de evaluatie, de discipline, de invoering van de basisvorming en de vak- en leerstofinhoudelijke afstemming van het aardrijkskunde-onderwijs. De teamhomogeniteit is het grootst voor de interne afstemming van de leerstofkeuze, het behandeltempo en de proefwerken.

Volgens de sectieleiders oefent de directie vrijwel geen invloed uit op het aardrijkskunde-onderwijs. Dit geldt zowel ten aanzien van vakinhoudelijke, vakdidactische, evaluatieve als afstemmingskwesties. De enige uitzondering betreft de invoering van de basisvorming en de registratie en aanpak van spijbelaars.

3.2 Vaksectie en prestaties

In een preliminaire analyse is nagegaan of de aardrijkskundeprestaties van scholen met een vaksectie aardrijkskunde inderdaad hoger zijn dan die van scholen zonder zo'n sectie (vgl. Kuhlemeier & Van den Bergh, 1994c). Zowel binnen het MAVO als het LBO blijkt er geen significant verschil te bestaan tussen de gemiddelde prestaties van scholen met en scholen zonder een sectie aardrijkskunde. Deze bevinding houdt ook stand als rekening wordt gehouden met verschillen tussen beide groepen scholen qua achtergrond-, instroom- en beginkenmerken. De scholen zonder een sectie aardrijkskunde zijn in de verdere analyses dan ook buiten beschouwing gelaten.

Het effect van de sectie is onderzocht in de vier opleidingstypen (VWO, HAVO, MAVO, LBO) en de drie schoolstructuren (categoriale scholen, smalle scholengemeenschappen en brede scholengemeenschappen) afzonderlijk¹. De uitkomsten van de analyse van de vier onderscheiden modellen (zie paragraaf 2.5) zijn weergegeven in Tabel 4. Omdat in dit artikel de sectie aardrijkskunde centraal staat en het effect van de achtergrond-, instroom- en beginkenmerken grotendeels overeenkomt met eerder gepubliceerde gegevens (vgl. Van den Bergh & Kuhlemeier, 1997; Kuhlemeier & Van den Bergh, 1994b) worden de effecten ervan op de prestaties hier niet gerapporteerd. Omwille van de overzichtelijkheid worden

alleen de regressiegewichten voor het effect van de sectiekenmerken in de tabellen vermeld. Voor de volledige uitkomsten zij verwezen naar Kuhlemeier en Van den Bergh (1994c).

Effectieve secties in het VWO

In het VWO blijkt de vergaderfrequentie een negatieve bijdrage te leveren aan de prestaties aan het einde van het derde leerjaar aardrijkskunde (model A). Hoe vaker er wordt vergaderd, hoe lager de prestaties. Tevens blijken secties die eigen regels hebben ontwikkeld op het terrein van de proefwerken en het huiswerk hogere prestaties te realiseren (model A). Houden we echter rekening met verschillen in het type leerlingen waarmee de secties te maken hebben, dan is dit effect niet meer significant (model B, C en D).

Onder constanthouding van de achtergrond-, instroom- en beginkenmerken zijn twee sectiekenmerken in het VWO van belang. Secties die vaak vergaderen over de vak- en leerstofinhoudelijke afstemming van het aardrijkskunde-onderwijs behalen hogere prestaties. Dit geldt ook voor directies die (volgens de sectieleiders) invloed uitoefenen op disciplinair gebied.

Effectieve secties in het HAVO

In het HAVO heeft de duur van de sectievergadering een positief effect: langer vergaderen gaat gepaard met hogere aardrijkskundeprestaties (zie Tabel 4). Het aantal leden dat op de sectievergadering aanwezig is, levert een negatieve bijdrage. Hoe groter de sectievergadering, hoe lager de prestaties. Dit effect treffen we overigens alleen aan in model B. Algemeen overleg met de directie en overleg met de directie over het aardrijkskunde-onderwijs leveren een positieve bijdrage⁴. De frequentie waarin de sectieleider met andere secties overlegt, is van negatieve invloed, maar ook dit effect treedt slechts één keer op (model B).

Naarmate in het sectieoverleg meer gesproken wordt over de vorderingen en de aanpak van zwakke en snelle leerlingen zijn de leerlingprestaties in het HAVO lager. Oefent de directie (volgens de sectieleider) invloed uit op de leerlingbegeleiding, dan komt dit de prestaties van de leerlingen ten goede (model D).

De effecten van de teamhomogeniteit in het HAVO zijn zo op het eerste gezicht wat incon-

sistent. Overeenstemming over de begeleiding van individuele leerlingen levert een negatieve bijdrage aan de aardrijkskundeprestaties, maar teamhomogeniteit inzake vak- en leerstofinhoudelijke afstemmingskwesties is van positief belang. Het negatieve effect van overeenstemming over de leerlingbegeleiding verdwijnt evenwel als we rekening houden met verschillen tussen secties in de aardrijkskundige kennis van de leerlingen waarmee zij aan het begin van het schooljaar geconfronteerd worden.

Effectieve secties in het MAVO

Over de relatie tussen de sectiekenmerken en de prestaties van MAVO-leerlingen kunnen we kort zijn (zie Tabel 4). Het aantal sectieleden dat op een sectievergadering aanwezig is, hangt in positieve zin samen met de aardrijkskundeprestaties.

De secties met een relatief duidelijke taakverdeling tussen de sectieleden (voorzitter, agenda, verslag, afsprakenlijst, vaste werkverdeling) bereiken een lager prestatieniveau. Overigens verdwijnt dit effect zodra rekening gehouden wordt met verschillen tussen secties in de aardrijkskundige kennis van de leerlingen aan het begin van het derde leerjaar.

Effectieve secties in het LBO

In het LBO hebben drie sectiekenmerken in elk van de vier modellen een vergelijkbaar effect⁵. Naarmate secties meer bindende regels ten aanzien van proefwerken en huiswerk kennen, behalen de leerlingen wat lagere prestaties. Overeenstemming binnen de sectie over de leerlingbegeleiding levert daarentegen een positieve bijdrage aan de prestaties. Op scholen waar de directie (volgens de sectieleider) invloed uitoefent op de orde en regels zijn de leerprestaties van de leerlingen iets hoger.

In het LBO verschilt het effect van de invloed van de directie al naar gelang het specifieke onderwerp. Invloed op de orde en regels hangt zoals we zagen positief samen met de leerprestaties. Bemoeienis met het vak- en leerstofinhoudelijke sectiebeleid levert eveneens een positieve bijdrage. Heeft de directie aandacht voor de begeleiding van met name zwakke leerlingen en klassen, dan zien we lagere prestaties. Deze laatste twee effecten sneuvelen echter zodra rekening gehouden wordt met de verschillen tussen secties in de (gemiddelde)

aardrijkskundige kennis van de leerlingen aan het begin van het derde leerjaar.

Geïntegreerde secties (meestal met geschiedenis) realiseren hogere prestaties dan secties waarin alleen aardrijkskunde vertegenwoordigd is. Ten slotte, als er vergaderd wordt met een duidelijke taakverdeling behalen de leerlingen hogere prestaties (model A en B).

Effectieve secties in categoriale scholen

In categoriale scholen blijkt de sectie er niet of nauwelijks toe te doen (zie Tabel 4). Slechts twee sectiekenmerken zijn (soms) significant. Naarmate de sectie(vergadering) meer leden telt, zijn de prestaties aan het einde van het derde leerjaar hoger. Tevens constateren we relatief lage aardrijkskundeprestaties bij secties die naar verhouding veel eigen regels hebben opgesteld. Corrigeren we voor verschillen tussen secties in het cognitieve aanvangsniveau van de leerlingen, dan verdwijnen beide effecten.

Effectieve secties in kleine scholengemeenschappen

In kleine scholengemeenschappen hangen vier sectiekenmerken samen met de aardrijkskundeprestaties (zie Tabel 4): de taakverdeling (negatief), de contacten met andere secties (negatief), de overeenstemming over de vak- en leerstofinhoudelijke afstemming van het aardrijkskundeonderwijs (positief), en de invloed van de directie op disciplinair gebied (positief). Slechts één kenmerk geeft in meer dan één model significantie te zien. Hoe meer overeenstemming over vak- en leerstofinhoudelijke kwesties, hoe hoger de aardrijkskundeprestaties.

Effectieve secties in brede scholengemeenschappen

Binnen brede scholengemeenschappen levert één sectiekenmerk een consistente bijdrage aan de aardrijkskundeprestaties (zie Tabel 4). Een naar verhouding lange tijdsduur van de sectievergadering gaat samen met relatief hoge prestaties.

3.3 Consistentie over opleidingstype en schoolstructuur

Hiervoor brachten we verslag uit van de invloed van de sectiekenmerken binnen elk

Tabel 5

Consistentie van het effect van de sectie over opleidingstype (Model A: alleen sectiekenmerken; Model B: sectie- en achtergrondkenmerken; Model C: sectie- en beginkenmerken; Model D: sectie-, achtergrond- en beginkenmerken)

	Model A				Model B				Model C				Model D			
	VW	HA	MA	LB	VW	HA	MA	LB	VW	HA	MA	LB	VW	HA	MA	LB
Geïntegreerde sectie												+				+
Vergaderfrequentie	-															
Vergaderduur		+				+			+					+		
Aantal leden aanwezig			+			-	+									
Formalisering																
vaste tijd/plaats																
taakverdeling			-	+		-	+									
Overleg met																
directie						+								+		
directie over ak		+	-			+	-		+							
andere secties						-										
Sectieoverleg over																
leerlingbegeleiding	-								-					-		
afstemming						+			+					+		
Eigen regels		+				-			-					-		-
Teamhomogeniteit t.a.v.																
leerlingbegeleiding		-	+			-	+			+						+
afstemming		+					+								+	
orde en regels																
Invloed directie op																
leerlingbegeleiding			-			-								+		
afstemming			+			+										
orde en regels			+	+		+	+	+	+	+	+	+	+	+	+	+

opleidingstype en binnen elke schoolstructuur afzonderlijk. Een belangrijke vraag is die naar de generaliseerbaarheid van de bevindingen over opleidingstype en schoolstructuur. Tabel 5 geeft een overzicht van de consistentie van de kenmerken van hoog en laag effectieve secties over opleidingstypen. Met het plus- en minteken is de richting (d.w.z.: het algebraïsche teken) van het verband aangeduid. We spreken van een consistent effect als aan twee voorwaarden is voldaan. Ten eerste moet het effect in ten minste twee opleidingstypen optreden. Ten tweede moet de richting ervan gelijk zijn, ook als het effect zich in meer dan twee opleidingstypen zou voordoen.

Slechts één kenmerk heeft in twee opleidingstypen een soortgelijk effect (zie Tabel 5). Zowel in het VWO als het LBO heeft de mate waarin de directie (volgens de sectieleider) invloed uitoefent op het handhaven van de regels en de orde een positief effect op de aardrijkskundeprestaties. De overige effecten zijn of beperkt tot één opleidingstype of inconsistent over opleidingstypen in de zin dat de richting van het verband verschillend is.

Vergelijking van de uitkomsten in categoriale scholen, smalle en brede scholengemeenschappen brengt geen gemeenschappelijke kenmerken van laag of hoog effectieve secties aan het licht.

3.4 Sectie versus achtergrond, instroom en beginkennis

De uitkomsten van de analyse van het (additieve) effect van de sectiekenmerken gegeven de achtergrond van de leerlingen, de instroom van de school en de aardrijkskundige kennis aan het begin van het derde leerjaar zijn opgenomen in Tabel 6. Per opleidingstype is eerst weergegeven hoe de totale variantie verdeeld is over de drie niveaus: school, klas en leerling. Tevens worden de drie modellen met elkaar vergeleken aan de hand van de toetsingsgrootheid -2 Log Likelihood en de toename van het percentage verklaarde variantie (R^2_{totaal}).

De achtergrond-, instroom- en beginkenmerken blijken van groot belang voor de aardrijkskundeprestaties. Op grond van de in Tabel 6 gerapporteerde passingsgegevens⁶ kunnen we

Tabel 6

Variantie op school-, klas- en leerlingniveau, modelpassing en proportie verklaarde variantie in het VWO, HAVO en LBO (Model 1: onconditioneel; Model 2: plus achtergrond-, instroom- en beginkenmerken; Model 3: plus sectiekenmerken; df: verschil in aantal vrijheidsgraden vergeleken met het voorgaande model; onderstreept: niet significant op 5%-significantieniveau)

	Model 1	Model 2	Model 3
Voorbereidend Wetenschappelijk Onderwijs (VWO)			
Varianties			
school	10.58	<u>3.86</u>	.00
klas	16.33	5.30	5.25
leerling	243.70	187.30	187.30
-2 Log Likelihood ¹	210.90	203.57	203.35
Vershil df		7	2
Passingsevaluatie		< .01	< .01
R^2_{totaal}		.29	.30
Hoger Algemeen Voortgezet Onderwijs (HAVO)			
Varianties			
school	13.43	<u>1.25</u>	.00
klas	<u>1.72</u>	<u>1.40</u>	.07
leerling	247.00	184.20	184.20
-2 Log Likelihood ¹	192.80	185.68	185.41
Vershil df		5	5
Passingsevaluatie		< .01	< .01
R^2_{totaal}		.29	.30
Lager Beroepsonderwijs (LBO)			
Varianties			
school	27.88	5.29	.60
klas	16.75	<u>3.37</u>	<u>3.54</u>
leerling	182.80	136.80	136.50
-2 Log Likelihood	9487.74	9094.93	9077.54
Vershil df		5	8
Passingsevaluatie		< .01	< .01
R^2_{totaal}		.37	.39

1 * 10^{-2}

concluderen dat de achtergrond-, instroom- en beginkenmerken 29% à 37% van de totale variantie verklaren (R^2_{totaal}). Zowel in het VWO, HAVO als LBO zien we een aanzienlijke reductie van de variantie tussen scholen, klassen en leerlingen (met uitzondering van de tussen-klassenvariantie in het HAVO, maar deze was al verwaarloosbaar klein in model 1). In het VWO en HAVO wijkt de tussen-scholenvariantie na het toevoegen van de achtergrond-, instroom- en beginkenmerken niet meer significant af van nul en dat geldt ook voor de tussen-klassenvariantie in het LBO. In het VWO resteert na uitschakeling van deze

'covariaten' nog een aanzienlijke hoeveelheid tussen-klassenvariantie; hetzelfde geldt voor de tussen-scholenvariantie in het LBO.

De sectiekenmerken blijken in het VWO, HAVO en LBO een significante bijdrage te leveren aan de voorspelling van de prestaties, boven hetgeen voorspeld kan worden op grond van de achtergrond-, instroom- en beginkenmerken. Tevens is te zien dat de additieve bijdrage van de sectiekenmerken in het VWO en HAVO erg klein is. Ten gevolge van de sectiekenmerken stijgt R^2_{totaal} zowel in het VWO als het HAVO met 1%. In het LBO leveren de sectiekenmerken een iets grotere bijdrage aan het

voorspellen van de prestaties: R^2_{totaal} stijgt van 37% naar 39%. Dit is ook te zien aan de tussen-scholenvariantie: in het LBO deze wordt deze tot vrijwel nul teruggebracht.

Binnen de groep categoriale scholen levert geen van de sectiekenmerken een additieve bijdrage aan de leerprestaties. Binnen de smalle en brede scholengemeenschappen verklaren de achtergrond-, instroom- en beginkenmerken 60% respectievelijk 42% van de totale variantie in leerlingprestaties; de additieve bijdrage van de sectiekenmerken is in beide gevallen minder dan 1% (vgl. Kuhlemeier & Van den Bergh, 1994c).

4 Discussie

In recente publicaties krijgt de vaksectie als intermediair tussen directie en leerresultaten een cruciale rol toebedeeld. Een en ander veronderstelt dat secties regelmatig vergaderen en regelmatig overleg voeren met de directie en met andere secties. Op grond van onze studie moeten we concluderen dat secties aardrijkskunde weinig vergaderen, en als ze dat al doen, gebeurt dat op een informele en weinig gestructureerde wijze (vgl. Hannay & Denby, 1994). Gezien het nagenoeg ontbreken van contacten met de directie en met andere vaksecties (zie ook: Van Wessum, 1997a,b) lijken aardrijkskundesecties nauwelijks verankerd in de schoolorganisatie. Veel secties aardrijkskunde laten een belangrijke middel om invloed uit te oefenen onbenut (vgl. HMI, 1984; Worner & Brown, 1993). Deze bevinding vormt een ondersteuning voor de 'los-zand-theorie' van schoolorganisaties (Weick, 1976), zeker als het om het vak aardrijkskunde gaat.

Zoals in de inleiding is uiteengezet kan een sectie drie, in organisatorische zwaarte toenemende functies vervullen (Marx, 1992): contactorgaan, staforgaan en sturingsorgaan. Het onderhavige onderzoek doet vermoeden dat de sturingsfunctie vooral betrekking heeft op de vak- en leerstofinhoudelijke afstemming van het aardrijkskundeonderwijs. Voor zover secties de vrijheid van individuele docenten inperken, heeft dat vooral betrekking op kwesties als de leerstofkeuze, het behandeltempo, proefwerken en huiswerk. Een groot deel van de secties heeft bindende regels ontwikkeld

voor het omgaan met proefwerken en huiswerk. Bovendien worden deze regels volgens de sectieleider door de vakdocenten goed nageleefd. Deze gegevens komen in grote lijnen overeen met die van Van Wessum (1997a) en Witziers (1992). Laatstgenoemde concludeert zelfs dat vaksecties (wiskunde) ten aanzien van het curriculaire domein sterk ontwikkelde organen zijn die het handelen van de leden in sterke mate reguleren.

Sterk onderwijskundig leiderschap is één van de kenmerken waarin effectieve scholen zich zouden onderscheiden van minder effectieve scholen (o.a. Creemers, 1991; Lugthart, Roeders, Bosker & Bos, 1989; Scheerens, 1989). Effectieve schoolleiders passen niet alleen op de winkel, maar zijn ook bereid en in staat om daadwerkelijke bemoeienis met het gegeven onderwijs te hebben. Voor zover er ten onzent al onderzoeksgegevens beschikbaar zijn, wijzen deze in de richting van een zwakke en dikwijls negatieve relatie tussen onderwijskundig schoolleiderschap en effectiviteitscriteria (o.a. Brandsma & Stoel, 1987; Van der Grint & Akkermans, 1991; Vermeulen, 1987). Uit ons onderzoek blijkt dat de directie volgens de sectieleiders nauwelijks enige bemoeienis heeft met zaken die traditioneel tot het werkterrein en de verantwoordelijkheid van vakdocenten behoren. Onderwijskundig schoolleiderschap met betrekking tot het aardrijkskunde-onderwijs in het derde leerjaar is vrijwel non-existent. Wel heeft de directie enige invloed op vak- en leerjaaroverstijgende kwesties, zoals het spijbelen en de invoering van de basisvorming. Dat onderwijskundig leiderschap ten opzichte van het vak aardrijkskunde zwak ontwikkeld is, blijkt ook uit de vragenlijst die door schoolleiders is ingevuld (Kuhlemeier & Van den Bergh, 1994a). Daaruit komt naar voren dat schoolleiders zelden overleggen met de sectie(leider) aardrijkskunde. Ook in de sfeer van de begeleiding van aardrijkskundeleraars speelt de schoolleiding een weinig actieve rol. Zaken als functioneringsgesprekken, beoordelingsgesprekken en het bijwonen van aardrijkskundelessen komen nauwelijks voor.

In het onderzoek is een onderscheid gemaakt tussen enerzijds sectiekenmerken en anderzijds achtergrond-, instroom- en beginkenmerken. Deze laatste groep kenmerken ver-

tegenwoordigt de randvoorwaarden die vanuit het perspectief van de vaksectie niet of nauwelijks te beïnvloeden zijn (vgl. Bosker, 1991; Bryk & Raudenbush, 1992; Willms, 1992). De achtergrond-, instroom- en beginkenmerken blijken van veel groter belang voor de leerlingprestaties dan de wel veranderbare sectiekenmerken (zie ook Ellemers, 1976). De onderzochte kenmerken van de vaksectie aardrijkskunde leveren slechts een kleine bijdrage aan de voorspelling van de aardrijkskundeprestaties in het derde leerjaar. Nadat gecorrigeerd is voor verschillen tussen secties in de 'kwaliteit' van de leerlingen waarmee zij te maken hebben, verklaren de sectiekenmerken in het VWO, HAVO, MAVO en LBO respectievelijk 1%, 1%, 0% en 2% van de totale variantie in leerlingprestaties.

Echter, omdat de sectiekenmerken gecorrigeerd zijn met de achtergrond-, instroom- en beginkenmerken dreigt het gevaar dat met het elimineren van het effect ervan ook een deel van het effect van de sectiekenmerken verwijderd is. De sectiekenmerken zouden dan geen faïre kans hebben gekregen om effectief te blijken en hun effect zou onderschat zijn. Dat dit gevaar niet denkbeeldig is, blijkt uit het feit dat er na correctie voor achtergrond-, instroom- en beginkenmerken maar weinig variantie resteerde tussen scholen en tussen klassen binnen scholen (zie ook Tabel 6). Om deze reden is het effect van de sectiekenmerken ook geschat vóór correctie voor achtergrond, instroom en beginkennis. Dan blijken de sectiekenmerken in VWO, HAVO, MAVO en LBO respectievelijk 2%, 3%, 1% en 14% van de totale variantie tussen leerlingen in leerprestaties te verklaren. Vóór correctie leveren de sectiekenmerken in het LBO de grootste bijdrage aan de prestaties. Een verklaring voor dit laatste is moeilijk te geven. Mogelijk speelt de wat grotere spreiding van de sectiekenmerken in het LBO daarbij een rol.

Slechts één kenmerk heeft in ten minste twee opleidingstypen een vergelijkbaar effect. Het betreft de invloed van de directie op het handhaven van de regels en de orde. Opmerkelijk is dat de invloed van de directie op disciplinair gebied de minst zwak ontwikkelde indicator is voor het onderwijskundig leiderschap van de directie. De invloed van de directie op de leerlingbegeleiding en de vak- en leerstof-

inhoudelijke afstemming van het aardrijkskunde-onderwijs is nog kleiner. Een mogelijke verklaring voor het ontbreken van de veronderstelde relatie tussen sectie en leerprestaties is dat veel secties nog te zwak ontwikkeld zijn om een effect op de leerprestaties te kunnen hebben. Deze verklaring is in overeenstemming met Hofman en Lugthart (1991, p. 170) die concluderen dat scholen voor voortgezet onderwijs nog nauwelijks beschikken over interne capaciteiten op het gebied van coördinatie en management.

Volgens het contingentieperspectief (vgl. Scheerens & Stoel, 1987) is de effectiviteit van schoolinterne processen ook afhankelijk van omgevingsfactoren. Schoolexterne factoren – zoals schoolgrootte, schoolstructuur, opleidingstype, urbanisatiegraad, denominatie en de 'kwaliteit' van de instroom – kunnen daarbij een faciliterende dan wel belemmerende rol spelen. Ons onderzoek werpt nieuw licht op de interacties tussen sectiekenmerken, leerprestaties en de contingentiefactoren opleidingstype en schoolstructuur. De kenmerken van hoog en laag effectieve secties blijken sterk van opleidingstype tot opleidingstype te variëren. Zoals gezegd heeft slechts één kenmerk in twee van de vier opleidingstypen een vergelijkbaar effect. Daarnaast bleek geen enkel sectiekenmerk een uniform effect te hebben over schoolstructuur. Voor de effectiviteit van de aardrijkskundesecties binnen categoriale scholen zijn andere kenmerken relevant dan voor de effectiviteit van secties in smalle en in brede scholengemeenschappen. Hiermee lijkt ondersteuning gevonden te zijn voor het contingentieperspectief. De relevantie van deze bevinding kan echter worden betwijfeld. Het belang van de aardrijkskundesectie voor de leerprestaties van de leerlingen lijkt immers nog gering.

Uit onderzoek naar effectieve scholen blijkt steevast dat het schoolbeleid de leerprestaties slechts in beperkte mate beïnvloedt (o.a. Creemers, 1991; Lugthart et al., 1989; Scheerens, 1989). Een van de mogelijke verklaringen is de lange causale keten tussen schoolbeleid en leerprestaties en de verwaarlozing van het middenmanagement in onderzoek naar instructie- en schooleffectiviteit (Turner, 1996; Worner & Brown, 1993). Op grond van ons onderzoek lijkt de vaksectie als alternatieve verklaring voor de zwakte van de gevonden schooleffecten vooralsnog weinig plausibel.

Een belangrijke onderwijsvernieuwing in de eerste fase van het voortgezet onderwijs omvat het onderwijzen van de algemene, vakoverstijgende vaardigheden van de basisvorming, zoals leren communiceren, leren samenwerken, leren onderzoeken en leren reflecteren op het leren en de toekomst (Ministerie van OC&W, 1997). De overladenheid van het onderwijsprogramma laat het niet toe dat elk vak aan al deze algemene vaardigheden evenveel aandacht kan besteden. Dit betekent dat afstemming binnen de school en taakverdeling tussen secties noodzakelijk is (Oosterhuis & Kapteijn, 1992). De zwakte van de sectie aardrijkskunde en het nagenoeg ontbreken van dwarsverbanden met andere secties kan de invoering van onderwijsvernieuwingen ernstig belemmeren (vgl. Van Wessum, 1997a).

Ten slotte zij opgemerkt dat het onderzoek is uitgevoerd in het schooljaar 1991-1992, aan de vooravond van de invoering van de basisvorming in 1993. Met de basisvorming is een groot aantal vakinhoudelijke en vakdidactische vernieuwingen nagestreefd. Tegelijkertijd heeft de schaalvergrotingsoperatie geleid tot een kleiner aantal scholen(gemeenschappen) van grotere omvang en complexiteit. Hiermee is de noodzaak tot coördinatie door onder meer vaksecties alleen maar groter geworden (vgl. Smets, 1986a). Dit roept de vraag op naar de generaliseerbaarheid van onze bevindingen naar de situatie anno 1999. Herhaling van het onderhavige onderzoek verdient derhalve aanbeveling.

Noten

1 De variabele 'schoolstructuur' is geconstrueerd op basis van het aantal verschillende opleidingstypen van de school in het vierde leerjaar. Categoriele scholen beschikken slechts over één opleiding voor VWO (Atheneum en/of Gymnasium), HAVO, MAVO of LBO (waarbij we scholen met slechts één opleidingstype voor LTO, LHNO of LEAO/LMO/LAO tot de categoriele LBO-scholen rekenen). Tot de smalle scholengemeenschappen worden de scholen gerekend die twee van deze vier opleidingen aanbieden: VWO+HAVO, HAVO+MAVO, MAVO+LBO of LBO+LBO. De LBO+LBO-scholengemeenschappen kennen ten minste twee van de opleidingen

voor LTO, LHNO of LEAO/LMO/LAO. Tot de brede scholengemeenschappen behoren de scholen die ten minste drie van de opleidingen voor VWO, HAVO, MAVO en LBO aanbieden: de scholengemeenschappen voor VWO+HAVO+MAVO, HAVO+MAVO+LBO, VWO+HAVO+MAVO+LBO en AVO+LBO.

- 2 Bij de vertegenwoordiging van het LBO in de responsgroep van 141 scholen met een aardrijkskundesectie is een kanttekening te maken. Van de 1167 derdeklassers op de 24 LBO-scholen met een sectie aardrijkskunde kregen 644 geen les in aardrijkskunde. Van deze 644 leerlingen zaten er 167 op een LBO-opleiding waar geen van de onderzochte derdeklassers aardrijkskunde volgde (maar waar er wel een sectie aardrijkskunde aanwezig was).
- 3 Een opmerking over de interpretatie van de regressiegewichten is hier op zijn plaats. In ons geval geven die aan hoe de aardrijkskundeprestaties veranderen bij een verandering van één eenheid op de onafhankelijke variabele, rekening houdend met het effect van alle andere verklarende variabelen in het model. De grootte van een regressiegewicht geeft inzicht in het effect op de leeruitkomsten onder constanthouding van alle andere verklarende variabelen in de analyse. Een niet-significant regressiegewicht hoeft niet altijd op het ontbreken van een univariante samenhang (b.v. correlatie) te duiden. Als bijvoorbeeld etniciteit voor de leerresultaten van groter belang is dan sociaal milieu, zou de oorspronkelijke correlatie tussen milieu en prestaties kunnen verdwijnen op het moment dat deze gecorrigeerd wordt voor de invloed van etniciteit. Ook het omgekeerde is mogelijk: een effect verschijnt pas na correctie voor andere variabelen in het model. Zo zou het effect van de sociale achtergrond pas op kunnen treden nadat gecorrigeerd is voor een mogelijk ongelijke uitgangssituatie aan het begin van het schooljaar. Met andere woorden: partiële regressiegewichten kunnen niet zonder meer univariaat geïnterpreteerd worden.
- 4 Opmerkelijk is dat het positieve effect van het algemene overleg met de directie pas verschijnt na uitpartialisering van het effect van de achtergrondkenmerken (model B en D). Wellicht is de aanleiding tot algemeen overleg met de directie niet zozeer gelegen in lage aardrijkskundeprestaties als wel in problemen die te maken hebben met de achtergrondkenmerken van de leerlingen.

- 5 Voor de analyse zijn slechts 24 LBO-scholen met een sectie aardrijkskunde beschikbaar. Om overparametrisering te voorkomen is een deel van de school- en instroomkenmerken niet in de analyse betrokken (voor meer informatie: zie Kuhlemeier & Van den Bergh, 1994c).
- 6 In Tabel 6 zijn de gegevens van het MAVO niet opgenomen omdat in model B geen van de sectiekenmerken significantie te zien gaf.

Literatuur

- Bentler, P.M., & Bonet, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bergh, H. van den, & Kuhlemeier, H. (1997). Multiniveau modellen voor de analyse van leerwinst vergeleken. *Tijdschrift voor Onderwijsresearch*, 22, 54-75.
- Bosker, R.J. (1991). De consistentie van schooleffecten in het basisonderwijs. *Tijdschrift voor Onderwijsresearch*, 16, 206-218.
- Bosker, R.J., & Witziers, B. (1995). Een multilevel benadering voor het meten van organisatiecultuur variabelen. In B. Creemers (Red.), *Onderwijsonderzoek in Nederland en Vlaanderen 1995* (pp. 245-246). Groningen: GION.
- Brandsma, H.P., & Knuver, A. (1991). De contextgebondenheid van effectiviteitsbevorderende schoolkenmerken. *Tijdschrift voor Onderwijsresearch*, 16, 218-230.
- Brandsma, H.P., & Stoel, W.G.R. (1987). Schoolleiders in het voortgezet onderwijs. Een exploratieve studie naar de determinanten en effecten van schoolleiders-kenmerken. In J. Scheerens & W.G.R. Stoel (Red.), *Effectiviteit van onderwijsorganisaties* (pp. 99-114). Lisse: Swets & Zeitlinger.
- Bryk, A.S., & Raudenbush, S.W. (1992). *Hierarchical linear models: application and data analysis methods*. Newbury Park: Sage Publications.
- Creemers, B.P.M. (1991). *Effectieve instructie: een empirische bijdrage aan de verbetering van het onderwijs in de klas*. 's-Gravenhage: Instituut voor Onderzoek van het Onderwijs.
- Earley, P., & Fletcher-Campbell, F. (1989). *The time to manage? Department and faculty heads at work*. London: Routledge.
- Ellemers, J.E. (1976). Veel kunnen verklaren of iets kunnen veranderen: krachtige versus manipuleerbare variabelen. *Beleid & Maatschappij*, 11, 281-290.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Griffin.
- Gooren, W.A.J. (1989). *De sergeant op school? Samenvattingen en conclusies naar middenmanagement in het voortgezet onderwijs*. Tilburg: IVA.
- Grift, W. van der, & Akkermans, W. (1991). Onderwijskundig leiderschap en leerlingprestaties in het basisonderwijs. *Tijdschrift voor Onderwijsresearch*, 16, 230-243.
- Hamlin, B. (1990). The competent manager in secondary schools. *Educational Management and Administration*, 18 (3), 3-10.
- Hannay, L.M., & Denby, M. (1994). *Secondary school change: the role of department heads*. Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 4-8, 1994) [ED 380 877].
- Harris, A., Jamieson, I., & Russ, J. (1995). A study of effective departments in secondary schools. *School Organisation*, 15, 283-299.
- HMI Her Majesty's Inspectorate Scotland (1984). *Departmental organisation in secondary schools*. Cardiff: HMSO.
- Hofman, R.H., & Lugthart, E. (1991). *Interne capaciteit in het voortgezet onderwijs*. Groningen: RION.
- Hylkema, W.F.S. (1990). *Docenten en hun vaksecties. Een onderzoek naar de perceptie van docenten Nederlands, wiskunde en geschiedenis van het functioneren van de vaksectie als vakinhoudelijk en vakdidactisch besluitvormingsorgaan*. Nijmegen: Katholieke Universiteit Nijmegen.
- Kamphorst, J.C., & Tuinenga, M.G. (1995). Succes- en faalfactoren voor het functioneren van vaksecties. In B. Creemers (Red.), *Onderwijsonderzoek in Nederland en Vlaanderen 1995* (pp. 256-258). Groningen: GION.
- Kuhlemeier, J.B., & Bergh, H. van den (1994a). *Scholen in het voortgezet onderwijs: leerlingbegeleiding, regels en verzuim*. PPON-VO Werkdocument Nr 3. Arnhem: Instituut voor Toetsontwikkeling.
- Kuhlemeier, J.B., & Bergh, H. van den (1994b). *De invloed van de sociaal-etnische achtergrond, schoolloopbaan, vrije tijdsbesteding en inzet voor school op de aardrijkskundeprestaties*. PPON-VO Werkdocument Nr 4. Arnhem: Instituut voor Toetsontwikkeling.
- Kuhlemeier, J.B., & Bergh, H. van den (1994c). *De invloed van de vaksectie aardrijkskunde op de aardrijkskundeprestaties van derdeklassers in*

- het voortgezet onderwijs. PPO-VO Werkdocument Nr 6. Arnhem: Instituut voor Toetsontwikkeling.
- Kuhlemeier, J.B., Bergh, H. van den, Notté, H., Wagenaar, H., Verstralen, H., & Cappers, R. (1994). *Balans van het aardrijkskunde-onderwijs in het derde leerjaar van het voortgezet onderwijs*. Arnhem: Instituut voor Toetsontwikkeling.
- Levine, D.L., & Lezotte, L.W. (1990). *Unusually effective schools*. Madison: National center for effective schools research & development.
- Lugthart, E., Roeders, P.J.B., Bosker, R.J., & Bos, K.T. (1989). *Effectieve schoolkenmerken in het voortgezet onderwijs*. Deel 1: Een literatuuroverzicht. Groningen: RION.
- Luyten, H. (1994). Stability of school effects in Dutch secondary education: The impact of variance across subjects and years. *International Journal of Educational Research*, 21, 197-215.
- Marwijk-Kooy, M. van (1984). *Scholen verschillen: een verkennend vergelijkend onderzoek naar het intern functioneren van vijftientwintig scholengemeenschappen VWO-HAVO-MAVO*. Groningen: Wolters-Noordhoff.
- Marx, E. (1992). Sectieleider, organisatiekundig bezien. *NGL-blad*, 24 (23), 6-7.
- Ministerie van OC&W (1997). Besluit kerndoelen en adviesrentabel basisvorming 1998-2003. *Staatsblad van het Koninkrijk der Nederlanden*, 484, 1-64.
- Mintzberg, H. (1979). *The structuring of organizations: a synthesis of the research*. Englewood Cliffs, NJ: Prentice-Hall.
- Mintzberg, H. (1989). *Mintzberg on management: inside our strange world of organizations*. Englewood Cliffs, NJ: Prentice-Hall.
- Oosterhuis, P., & Kapteijn, M. (1992). *Algemene vaardigheden in de tweede fase voortgezet onderwijs*. Instituut voor Didactiek en Onderwijspraktijk. VU Amsterdam. Amsterdam.
- Prosser, R., Rasbash, J., & Goldstein, H. (1991). *Software for three-level analysis. Users' guide for V.2*. London: Institute of Education.
- Scheerens, J. (1989). *Wat maakt scholen effectief? Samenvattingen en analyses van onderzoeksresultaten*. 's-Gravenhage: SVO.
- Scheerens, J., & Bosker, R.J. (1997). *The foundations of educational effectiveness*. Amsterdam: Pergamon.
- Scheerens, J., & Stoel, W.G.R. (1987). Conceptuele en methodologische problemen bij onderzoek naar de effectiviteit van schoolorganisaties. In J. Scheerens & W.G.R. Stoel (Red.), *Effectiviteit van onderwijsorganisaties* (pp. 1-16). Lisse: Swets & Zeitlinger.
- Siskin, L.S. (1991). Departments as different worlds: subject subcultures in secondary schools. *Educational Administration Quarterly*, 27 (2), 134-169.
- Smets, P. (1986a). Het Nederlandse schoolorganisatie-onderzoek: een overzicht. *Meso*, 1, 9-21.
- Smets, P. (1986b). Middenmanagement in het onderwijs. In A.M.L. van Wieringen (Red.), *Management van onderwijsinstellingen* (pp. 176-190). Groningen: Wolters-Noordhoff.
- Turner, D.K. (1996). The roles and tasks of a subject head of department in secondary schools in England and Wales. A neglected area of research? *School Organisation*, 16 (2), 203-217.
- Verhelst, N.D., Glas, C.A.W., & Verstralen, H.H.F.M. (1994). *OPLM: one parameter logistic model. A computer program and manual*. Arnhem: Instituut voor Toetsontwikkeling.
- Vermeulen, C.J. (1987). De effectiviteit van onderwijs bij 17 Rotterdamse stimuleringscholen. *Pedagogische studiën*, 64, 49-58.
- Wang, M.C., Haertel, G.D., & Walberg, H.J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63, 249-295.
- Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1-19.
- Wessum, L. van (1997a). *De sectie als eenheid. Samenwerking en professionaliseringsopvattingen van docenten in het voortgezet onderwijs*. Utrecht: Brouwer Uithof.
- Wessum, L. van (1997b). Vaksecties als los zand. *Het onderwijsblad*, 19, 18-19.
- Willms, J.F. (1992). *Monitoring school performance: a guide for educators*. Washington, DC: The Falmer Press.
- Witziers, B. (1992). *Coördinatie binnen scholen voor voortgezet onderwijs*. Enschede: TU Twente.
- Witziers, B. (1993a). Vaksecties en schoolleiders als indicatoren voor de kwaliteit van het onderwijs. *Mesomazine*, 66, 13-17.
- Witziers, B. (1993b). Coördinatie en de effectiviteit van het onderwijs op het vakgebied wiskunde. *Tijdschrift voor Onderwijsresearch*, 18, 279-294.
- Worner, W., & Brown, G. (1993). The instructional leadership team: a new role for the department head. *NASSP Bulletin*, 77, 37-45.

Auteurs

J.B. Kuhlemeier (1952) is als onderwijskundig onderzoeker werkzaam bij het Instituut voor Toetsontwikkeling te Arnhem.

Adres: CITO, bedrijfsgroep basisvorming, Postbus 1034, 6801 MG Arnhem.

H. van den Bergh (1957) is verbonden aan de Universiteit Utrecht waar hij onderzoek doet binnen het Utrechts Instituut voor Linguïstiek (UIL-OTS) en onderwijs verzorgt bij het Onderwijsinstituut Nederlands.

Abstract

Departments and learner results in the third year of Dutch secondary education

J.B. Kuhlemeier & H. van den Bergh. Pedagogische Studiën, 1999, 76, 2-20.

In recent theories on school organization, the importance of departments for instruction and learning is stressed. This article examines the effect of geography departments on geography achievements, taking into account differences between departments in the type of students they are faced with. Geography departments appeared to be weakly developed. The contribution of geography departments on the geography achievements of third-year students was rather small. When background and entry characteristics were taken into account, only 0% to 2% of the total variance could be explained by the department. Moreover, the characteristics of effective departments that were found appeared to be rather inconsistent across curricular tracks and school structures.

Ontwikkeling en validering van de vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen'*

F. Maes, R. Vandenberghe en P. Ghesquière

Samenvatting

Naar aanleiding van alarmerende berichten over het vastlopen van leerlingen in het basisonderwijs en in het secundair onderwijs, blijft de overtuiging groeien dat het gewoon onderwijs meer aandacht moet hebben voor leerlingen bij wie het onderwijsleerproces niet vlot verloopt. Men stelt vragen over de signalerende, diagnosticerende en remediërende capaciteit van het gewoon onderwijs. Deze bekommernis is gerelateerd aan het streven naar zorgverbreding. In dit artikel beschrijven we de ontwikkeling en validering van een instrument dat het diagnosticerend en remediërend vermogen van lagere scholen meet. Hierbij vertrekken we van een heuristisch kader waarin de aspecten worden beschreven die aanwezig zouden moeten zijn in een school om hulp te bieden aan leerlingen. Op basis hiervan werd met behulp van factoranalyse een vragenlijst met 10 schalen ontwikkeld, 6 op mesoniveau en 4 op microniveau. Op basis van enkele exploratieve analyses blijkt dat er voldoende aanwijzingen zijn dat deze 10 schalen betrouwbaar en valide zijn.

1 Achtergrond en probleemstelling

Gedurende de laatste vijftien jaar groeit de overtuiging dat het gewoon onderwijs meer aandacht moet hebben voor leerlingen bij wie het onderwijsleerproces niet vlot verloopt. Er zijn immers alarmerende berichten over het vastlopen van leerlingen in het basisonderwijs

en in het secundair onderwijs. Deze berichten hebben betrekking op drie vaststellingen. Zo wijst men ten eerste op het grote aantal zittenblijvers en leerlingen met schoolachterstand in het basisonderwijs en in het secundair onderwijs (Knuver & Reezigt, 1991; OESO, 1991; Van Damme, 1992-1993, 1993; Vlaamse Onderwijsraad, Vl.O.R., 1993). Daarnaast is er de groei van het buitengewoon onderwijs (speciaal onderwijs) (Ghesquière, De Fever & Van Hove, 1996; OESO, 1991; Pijl, 1997a, 1997b; Van Dun, 1990; Van Dun & Ruijsenaars, 1991) en ten slotte valt de toename op van allerlei vormen van buitenschoolse hulpverlening.

Deze vaststellingen roepen vragen op over de signalerende, diagnosticerende en remediërende capaciteit van het gewoon onderwijs. Deze capaciteit is nauw gerelateerd aan de begrippen 'zorgbreedte' en 'zorgverbreding'. Hieronder verstaan we het wijzigen, aanpassen en verdiepen van de aandacht, de maatregelen en de activiteiten op een school opdat elk kind maximale kansen krijgt om zich optimaal te ontwikkelen. Zorgverbreding is een belangrijke opdracht van het onderwijs en staat of valt met de capaciteit van het schoolteam om bij leerlingen problemen vast te stellen, te analyseren en indien nodig te remediëren.

Dit artikel heeft als doel het beschrijven van de ontwikkeling en validering van een instrument dat de signalerende, diagnosticerende en remediërende capaciteit van een lagere school kan vaststellen.

2 Aandacht voor zorgverbreding

Zoals vermeld past de ontwikkeling van zo'n instrument in de aandacht die men besteedt aan zorgverbreding. De gedachte dat *elke* leerling, ondanks zijn/haar verschil in achtergrond en aanleg, optimale kansen moet krijgen om zich te ontwikkelen blijft actueel (Stokking, 1992; Vl.O.R.-Forum Basisonderwijs, 1994). In dit

* Dit artikel kwam tot stand op basis van een onderzoek dat uitgevoerd werd in het kader van het Onderwijskundig Beleids- en Praktijkgericht Wetenschappelijk Onderzoek (OBPWO 94.01) in opdracht van het Ministerie van de Vlaamse Gemeenschap, Departement Onderwijs.

artikel sluiten we sterk aan bij deze gedachte, die men plaatst onder de noemer 'zorgverbreding'. De Nederlandse Adviesraad voor het Basisonderwijs (ARBO, 1984, p. 21) definieert zorgverbreding als *"de uitbreiding en versterking van maatregelen en activiteiten, vooral op school- en groepsniveau, ten behoeve van een zo intensief mogelijke zorg voor leerlingen, in het bijzonder voor hen die specifieke pedagogisch-didactische behoeften tonen, opdat zij in een zekere hoeveelheid tijd door de school te bepalen doelen en tussendoelen bereiken"*.

Zo veel mogelijk leerlingen tot een zo hoog mogelijk niveau van leren (cognitief, sociaal, affectief, fysiek-motorisch) brengen, is vanuit dit opzicht een belangrijke doelstelling van het onderwijs. Deze doelstelling omvat twee essentiële elementen. Enerzijds kan men zorgverbreding opvatten als het verbreden van de zorg wat betreft het aantal leerlingen. Hier wordt de bekommernis uitgedrukt voor alle kinderen en niet enkel voor de gemiddelde leerling. Concreet betekent dit dat men meer aandacht moet hebben voor die leerlingen die in het gewoon onderwijs uitvallen of dreigen uit te vallen. Anderzijds kan zorgverbreding ook betekenen dat men de zorg verbreedt wat betreft het aantal ontwikkelingsdomeinen. Hier wordt beklemtoond dat er op school mogelijkheden moeten worden voorzien om aan alle ontwikkelingsdomeinen te werken. De school moet met andere woorden, naast cognitieve vaardigheden, ook andere (bv. sociaal-emotionele of psychomotorische) capaciteiten en vaardigheden aanspreken en stimuleren (Laevers, 1994).

We hebben ervoor gekozen om dit onderzoek te richten op een duidelijk omschreven doelgroep, namelijk 'leerlingen met onderwijsleerproblemen'. In navolging van onder andere Van der Leij (1996) en Van Huynegem en Verammen (1991) gebruiken we de term 'onderwijsleerproblemen' omdat deze term de oorzaak van de problemen niet a priori beperkt. Concepten zoals leerproblemen, gedragsproblemen en leerstoornissen suggereren in sterke mate dat de oorzaak bij leerlingen zelf ligt. Datzelfde geldt voor een begrip als pedagogisch-didactische problemen dat uitsluitend naar de school of leerkracht verwijst (Van der Leij, 1996). Met de term 'onderwijsleerproble-

men' komen we tegemoet aan de complexe interactie van het lerende kind en de onderwijs-situatie. We willen hiermee de wisselwerking tussen verschillende factoren in de ontstaans-geschiedenis van het probleem benadrukken zonder een uitspraak te doen over de oorzaak van het probleem (Van der Leij, 1996).

Het signalerend, diagnosticerend en remediërend vermogen van een school wordt bijgevolg ingevuld als de zorg die een bepaalde school besteedt aan leerlingen met onderwijsleerproblemen.

Het realiseren van de zorg voor leerlingen met onderwijsleerproblemen is zowel een taak voor de individuele leerkracht als voor de school als organisatie. Dit betekent vooreerst dat het concrete handelen van de individuele leerkracht in de klas of groep ook moet zijn gericht op leerlingen met onderwijsleerproblemen. Het gaat hier onder meer om het signaleren, diagnosticeren en remediëren van onderwijsleerproblemen. De opeenvolging (cyclus) van signaleren, diagnosticeren en remediëren vormt tevens de basis van het model voor planmatig handelen (Kool & Van der Leij, 1985; Van der Leij & Kappers, 1994).

Zorg dragen voor leerlingen met onderwijsleerproblemen vraagt niet enkel inspanningen van individuele leerkrachten, maar moet tevens gedragen worden door de (ondersteunende) organisatorische structuren, procedures en activiteiten op het schoolniveau. Als op schoolniveau immers geen maatregelen worden genomen, wordt de opvang en aanpak van leerlingen met onderwijsleerproblemen uitsluitend bepaald door de inzet en kennis van de individuele leerkrachten. De maatregelen op schoolniveau creëren met andere woorden de voorwaarden waarbinnen signalerende, diagnosticerende en remediërende activiteiten in een klas tot stand kunnen komen. Het is derhalve een opdracht voor de school om werkcondities en een schoolcultuur te creëren waardoor leerkrachten gestimuleerd worden om aandacht te geven aan leerlingen met onderwijsleerproblemen (Vandenberghe, 1990). Dit impliceert nieuwe vaardigheden en een professionele samenwerking tussen leerkrachten, schoolhoofd, ouders en schoolbegeleidingsdiensten. Daarom is de school niet enkel een plaats waar leerlingen moeten leren, maar het is

evenzeer een plaats waar leerkrachten kansen moeten krijgen om te leren (Smylie, 1994, 1995, 1996). Het is daarom noodzakelijk dat, aangezien nieuwe vaardigheden verondersteld worden van leerkrachten, leerkansen en leer-ruimte gecreëerd worden om de professionele ontwikkeling van leerkrachten te ondersteunen (Clement, 1995). Kortom, als men de professionele kwaliteit van een leerkracht wil verbeteren, moet men evenzeer de kwaliteit van de werkcondities op een school verbeteren (Vandenbergh, 1992).

Daarbij is het essentieel dat de nodige inspanningen gelijktijdig en complementair gebeuren op beide niveaus (Maes, Vandenbergh & Ghesquière, in press). Het is immers mogelijk dat er op schoolniveau een systeem van signaleren, diagnosticeren en remediëren bestaat, maar dat dit mank loopt bij de reële uitvoering in de klas. Omgekeerd is het ook mogelijk dat de hulpverlening in de klas op basis van persoonlijke initiatieven van een leerkracht verloren gaat omdat een begeleidings-systeem op schoolniveau ontbreekt. Bij de begeleiding van leerlingen met onderwijsleerproblemen wordt duidelijk dat het gebeuren op klasniveau en op schoolniveau sterk met elkaar verweven is (zie ARBO-definitie en Gielis & Gerits, 1989).

3 Opbouw van de vragenlijst

Er werd gezocht naar een heuristisch kader voor het genereren van concrete variabelen en items. Dit kader pretendeert zeker niet een afgebakend conceptueel model te zijn. Het heeft enkel en alleen een heuristische functie en dient dus als aanzet tot het genereren van variabelen en items die een indicatie kunnen geven van het diagnosticerend en remediërend vermogen van lagere scholen. Dit kader bevat twee triaden. De eerste triade omvat drie hoofdactoren die een belangrijke rol spelen bij het realiseren van de signalerende, diagnosticerende en remediërende taken op een school, namelijk de leerkracht, het schoolteam en het schoolhoofd¹. De tweede triade behandelt de inhoudelijke dimensies die aanwezig zijn bij elke hoofdactor. Hiermee bedoelen we de visie of opvattingen, de activiteiten die worden gesteld en de organisatorische kenmerken. In

een volgende stap integreren we beide triaden.

De triade 'leerkracht-schoolhoofd-schoolteam'

Zoals reeds aangeduid zal een goed uitgebouwd systeem van zorg voor leerlingen met onderwijsleerproblemen zich zowel op meso- als op microniveau manifesteren. Op deze twee niveaus zijn verscheidene actoren van belang. Op microniveau is de wijze waarop de leerkracht in zijn klas werkt belangrijk. Op meso-niveau onderscheiden we twee actoren: het schoolhoofd en het schoolteam. Het schoolhoofd neemt uiteraard een centrale plaats in. Hij/zij is een belangrijke figuur in het schoolgebeuren (Gielis & Gerits, 1989; Staessens, 1991a; Van de Grift, 1987). Verder is het evident dat ook het schoolteam als geheel belangrijk is. Men kan pas over de klassen heen zorg dragen voor een leerling als de inspanningen van de individuele leerkrachten en het schoolhoofd worden gedragen door het volledige schoolteam.

De triade 'visie-organisatie-activiteiten'

Bij elk van de drie actoren van de eerste triade kunnen we drie dimensies onderscheiden die voor de realisatie van de signalerende, diagnosticerende en remediërende taken op een school een belangrijke rol spelen. Men kan voorerst een onderscheid maken tussen organisatorische aspecten enerzijds en de concrete activiteiten of handelingen anderzijds. Onder het eerste verstaan we de *organisatorische kenmerken* die aanwezig zijn om de zorg voor leerlingen met onderwijsleerproblemen te realiseren. Zo zijn de structuren die in het leven worden geroepen (bv. teamvergaderingen, orthotheek, leerlingvolgsysteem, leerlingdossier), de afspraken die worden gemaakt tussen leerkrachten en de manier waarop alles binnen een klas wordt georganiseerd aspecten van de organisatorische dimensie die zich zowel op meso- als op microniveau manifesteert. De *activiteitsdimensie* betreft het gericht handelen van de individuele leerkracht, het schoolhoofd en het schoolteam om problemen bij de leerlingen te voorkomen (preventie) en te verhelpen. De samenwerking op school en de inspanningen die worden geleverd om continuïteit te realiseren zijn voorbeelden van dergelijke activiteiten van het schoolteam. De belangrijkste activiteiten van de leerkracht op klasniveau – binnen de

onderzochte problematiek – zijn het signaleren, diagnosticeren en remediëren van leerlingen met onderwijsleerproblemen. Voor het schoolhoofd is dit het initiëren, ondersteunen en opvolgen van de signalerende, diagnosticerende en remediërende taken op school. Tenslotte bepalen de *opvattingen (visie)* over de zinvolheid en de inhoud van de zorg voor leerlingen met onderwijsleerproblemen in hoge mate de vormgeving. Dit betreft zowel de opvattingen van de individuele leerkracht, van het schoolhoofd als van het schoolteam (opvattingen van het schoolteam over de belangrijkste taken van een school die leiden tot een schoolconcept). Opvattingen over bijvoorbeeld de haalbaarheid en zinvolheid om ook in het gewoon basisonderwijs remediërend te werken zijn een essentieel onderdeel van de signalerende, diagnosticerende en remediërende capaciteit van een school. Deze visies geven aanleiding tot verschillende activiteiten (handelingen) en organisatorische maatregelen in de school en in de klas.

Integratie van de twee triaden

Zowel de organisatie, de activiteiten als de visie zijn aanwezig bij de drie hoofdfactoren van de eerste triade. De combinatie van beide triaden leidt zo tot een negenveldentabel. Die verscheidene velden werden vervolgens ingevuld met één of meerdere variabelen (zie Tabel 1). Bij de keuze van de variabelen baseerden we ons op de literatuur over zorgverbreding en op bestaande zorgverbredingsinstrumenten (Dienst voor Onderwijsontwikkeling, 1994; Gielis & Gerits, 1989; Huistra, 1980; Span, 1990). Zodoende kwamen we tot variabelen waarvan we veronderstellen dat ze een belangrijke rol spelen in de signalerende,

diagnosticerende en remediërende capaciteit van een lagere school.

Er werd gekozen voor een schriftelijke, gesloten vragenlijst. Dit biedt enkele belangrijke voordelen. Zo is een gestandaardiseerde vragenlijst vrij gemakkelijk af te nemen van een grote groep en geschikt als het aantal te meten variabelen groot is (Swanborn, 1987). Om in functie van een periodiek peilingsonderzoek bruikbaar te zijn, moet een herafname van het instrument mogelijk en betrouwbaar zijn. Een gestandaardiseerde vragenlijst leent zich daarvoor (Janssens, 1988).

Er werden voor elke variabele items geformuleerd. Elk item is een uitspraak waarbij de respondent op een zespunten-Likertschaal dient aan te geven in welke mate deze uitspraak van toepassing is, gaande van 'helemaal niet van toepassing' tot 'helemaal van toepassing' (Swanborn, 1982).

De vragenlijst zelf werd opgesplitst in twee delen: 'onze school' en 'mijn klas' in overeenstemming met het meso- en het microniveau. Aangezien de problematiek omtrent zorgverbreding erg actueel is, bestaat er een reële kans dat op de items sociaal wenselijk wordt geantwoord. Er werden daarom maatregelen genomen om dit in de mate van het mogelijke tegen te gaan. Zo werden ten eerste de uitspraken binnen de twee delen ('onze school' en 'mijn klas') door elkaar geplaatst, ongeacht tot welke variabele ze behoren, maar de items op schoolniveau bleven altijd gescheiden van de items op klasniveau. Daarnaast werden zowel 'positieve' als 'negatieve' uitspraken voor de zorg voor leerlingen met onderwijsleerproblemen opgenomen, om een 'response-set' tegen te gaan (Bartelds, Jansen & Joostens, 1989; Bil-

Tabel 1

De verschillende variabelen voor de zorg voor leerlingen met onderwijsleerproblemen

	MICRONIVEAU	MESONIVEAU	
	LEERKRACHT	SCHOOLTEAM	SCHOOLHOOFD
VISIE	- visie leerkracht	- schoolconcept	- visie schoolhoofd
ACTIVITEITEN	- signaleren - diagnosticeren - remediëren	- continuïteit realiseren	- initiëren/ondersteunen
ORGANISATIE	- klasorganisatie	- structuren - procedures	

liet, 1990; Sudman & Bradburn, 1987). Verder werden in de voorbereidende onderzoeken en het definitief onderzoek die items weggelaten die sociaal wenselijk beantwoord worden. Hiervoor baseerden we ons op de standaarddeviatie en het gemiddelde van de items. Een kleine standaarddeviatie en een extreem gemiddelde kunnen immers duiden op een sociaal wenselijk item.

Als respondenten werd gekozen voor de klasleerkrachten, schoolhoofden en taakleerkrachten (remedial teachers). Voor die drie groepen van respondenten werd een afzonderlijke vragenlijst gemaakt. Die voor klas- en taakleerkracht zijn grotendeels gelijk met uitzondering van enkele items op klasniveau. De vragenlijst van het schoolhoofd bestaat enkel uit de items op schoolniveau waarbij dan nog de items die handelden over het schoolhoofd weggelaten waren.

4 Opzet en uitvoering van het onderzoek

De ontwikkeling van de vragenlijst is empirisch ondersteund door een pilotstudy bij enkele leerkrachten en deskundigen en door een vooronderzoek in elf scholen. Daarin stonden de duidelijkheid van de vragenlijst (wat lay-out, instructies en begrippen betreft) en de zinvolheid van de uitspraken centraal. Verder werd in het vooronderzoek ook gekeken naar het gemiddelde en de spreiding van de items. Op basis hiervan werd de vragenlijst aangepast. We kwamen uiteindelijk tot 71 items op schoolniveau en 77 items op klasniveau (voor een uitvoerigere beschrijving zie Maes, Luyckx, Vandenberghen & Ghesquière, 1996).

Op basis van de volledige populatie van Vlaamse lagere scholen werd in het definitief onderzoek een gestratificeerde steekproef getrokken van 150 scholen. Deze steekproef is representatief voor grootte (aantal leerlingen), onderwijsnet, onderwijsniveau (kleuter en lager / enkel lager) en provincie. Van de 150 aangeschreven scholen waren er 104 bereid mee te werken. Als belangrijkste reden om niet deel te nemen aan het onderzoek, vermeldden de schoolhoofden de te zware belasting voor het schoolteam. De 104 scholen die wel deelnamen, waren ook representatief voor de hierboven beschreven kenmerken.

De hier beschreven analyses gebeurden enkel op basis van de gegevens van klasleerkrachten. De motivering hiervan is enerzijds van praktische aard omdat we hierdoor aan het grote aantal respondenten konden komen dat nodig is voor de analyses en ook omdat er verschillen waren in de vragenlijsten voor deze verschillende respondenten. Anderzijds was de keuze ook door inhoudelijke argumenten ingegeven. De klasleerkracht wordt immers als eerste en belangrijkste persoon aangezien voor de zorgverbreding in de klas en in de school. Hierbij is het belangrijk om een onderscheid te maken tussen respondenten, dat zijn de personen die de vragenlijst invullen, en het onderzoekselement, namelijk de school (Janssens, 1988; Knapp, 1982; Segers, 1987). Billiet (1992) spreekt in dit verband over onderzoekseenheden (cf. onderzoekselementen) en waarnemingseenheden (cf. respondenten). Ofschoon de klasleerkrachten de vragenlijst invullen en bijgevolg de waarnemingseenheden zijn, ligt de focus van ons onderzoek niet enkel op 'hun' individuele zorg voor leerlingen met onderwijsleerproblemen, maar zijn we op de eerste plaats geïnteresseerd in de zorg van hun school als geheel. De belangrijkste onderzoekseenheid is de school.

Er werden 1110 vragenlijsten voor klasleerkrachten verzonden, waarvan er 688 (= 62%) werden teruggestuurd.

5 De vragenlijst en de schalen

De vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen' werd van het begin af opgesplitst in twee delen: een deel met items op schoolniveau ('onze school') en een deel met items op klasniveau ('onze klas'). Bij de verdere analyses werd deze onderverdeling consequent doorgetrokken.

5.1 De vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen': schoolniveau

Op de 71 items werd een principale factoranalyse met varimaxrotatie uitgevoerd (Fontaine, 1997; Tacq, 1992). De zesfactorenoplossing werd weerhouden. Omdat de beslissing over het aantal factoren in ons geval niet eenduidig gemaakt kon worden op basis van statistische criteria (Tacq, 1992) werd deze beslissing

Tabel 2

De zes weerhouden schalen, omschrijving, aantal items en Cronbachs α van de vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen' op schoolniveau

SCHALEN EN HUN OMSCHRIJVING	AANTAL ITEMS	CRONBACHS α
S1: Het schoolhoofd en de gezamenlijke zorg voor leerlingen met onderwijsleerproblemen	15	0,89
S2: Structureel overleg en registratie in functie van overleg	13	0,86
S3: Gebrek aan zorg voor leerlingen met onderwijsleerproblemen	11	0,77
S4: Samenwerking met P.M.S.-centrum en met ouders	8	0,72
S5: Klasoverschrijdende initiatieven en coördinatie	10	0,73
S6: Continuïteit	5	0,71

sing hoofdzakelijk genomen op basis van inhoudelijke of interpretatieve argumenten (Fontaine, 1997). Deze oplossing wordt in zekere mate ook ondersteund door de screen-test van Cattell (Tacq, 1992). Deze zesfactor-oplossing 'verklaart' 32% van de totale variantie. Verder werden alle items met een lading onder 1.201 weggelaten. Een item met een lage factorlading is immers geen goede indicator voor die factor. Uiteindelijk bleven 62 items over. Op basis hiervan werd van elke schaal de Cronbachs α berekend (zie Tabel 2).

We beschrijven hieronder kort elke schaal en geven telkens enkele items als voorbeeld.

S1) Het schoolhoofd en de gezamenlijke zorg voor leerlingen met onderwijsleerproblemen

Deze schaal duidt aan in welke mate de leerkrachten vinden dat het schoolhoofd de zorg voor leerlingen met onderwijsleerproblemen een prioriteit vindt voor de school en dat het schoolhoofd initiatieven daaromtrent onderneemt of ondersteunt. Daarnaast wordt ook nagegaan in welke mate er een gezamenlijke doelgerichtheid (missie) is om van de zorg voor leerlingen met onderwijsleerproblemen een prioriteit te maken.

- Het schoolhoofd laat duidelijk horen dat de opvang van leerlingen met onderwijsleerproblemen bijzondere aandacht moet krijgen.
- De zorg voor leerlingen met onderwijsleerproblemen is op onze school voor iedereen een prioritaire bekommernis. ♦

S2) Structureel overleg en registratie in functie van overleg

Deze schaal geeft aan of er op een school overleg is en in welke mate dit overleg gestructureerd is. Dit wil zeggen of het regelmatig

plaatsvindt, of alle betrokkenen aanwezig zijn, of het voorbereid wordt, of de conclusies neergeschreven worden in leerlingdossiers, enzovoort. Daarnaast wordt er ook gepeild of de evoluties van de leerlingen geregistreerd worden. Worden vorderingenregistraties/-screenings gebruikt? Is er een leerlingvolgsysteem?

- Er vindt in onze school per trimester minstens één multidisciplinair overleg plaats.
- Op onze school maken wij gebruik van een leerlingvolgsysteem.

S3) Gebrek aan zorg voor leerlingen met onderwijsleerproblemen

Deze schaal gaat na of er tegenindicaties zijn voor de zorg voor leerlingen met onderwijsleerproblemen. Het betreft een breed spectrum van aanwijzingen dat de zorg voor leerlingen met onderwijsleerproblemen geen prioritaire bekommernis is op de school.

- Op onze school laat men soms leerlingen met onderwijsleerproblemen overgaan naar een volgend leerjaar om geen leerlingen te verliezen.
- Op onze school vermijdt men om leerlingen door te verwijzen naar het buitengewoon onderwijs om geen leerlingen te verliezen.

S4) Samenwerking met Psycho-Medisch-Sociaal-centrum (P.M.S.-centrum) en met ouders

Het gaat hier om de samenwerking tussen school en P.M.S.-centrum enerzijds en tussen de school en de ouders anderzijds. Deze schaal omvat dus twee elementen. Ten eerste gaat de schaal na in welke mate het overleg met het P.M.S.-centrum gericht is op het uitvoeren van een diagnostisch onderzoek en/of op het ondersteunen van leerlingen en leerkrachten wat betreft de zorg voor leerlingen met onderwijs-

Tabel 3

De vier weerhouden schalen, omschrijving, aantal items en Cronbachs α van de vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen' op klasniveau

SCHALEN EN HUN OMSCHRIJVING	AANTAL ITEMS	CRONBACHS α
K1: Diagnostiserend en remediërend onderwijs	21	0,81
K2: Ondersteuning en overleg	13	0,77
K3: Opvattingen	17	0,76
K4: Binnenklasdifferentiatie	10	0,70

leerproblemen. Daarnaast wordt er ook gepeild naar het informatie inwinnen van en doorgeven aan de ouders.

- De samenwerking met het P.M.S.-centrum is gericht op ondersteuning van ouders van leerlingen met onderwijsleerproblemen.
- Het is op onze school gebruikelijk dat iemand van het schoolteam zelf contact opneemt met de ouders van leerlingen met onderwijsleerproblemen die afwezig waren op oudercontacten.

55) Klasoverschrijdende initiatieven en coördinatie

Deze schaal peilt naar de klasoverschrijdende initiatieven die op school genomen worden en naar de externe of interklassikale differentiatie van een school door middel van klasoverschrijdende niveaugroepen. Wordt er samengewerkt over de klassen heen? Wordt er met niveaugroepen gewerkt? Wordt er overlegd over die groepering van leerlingen? Naast de externe of interklassikale differentiatie komt in deze schaal ook de ondersteuning door de taakleerkracht (remedial teacher) of de ambulante leerkracht aan bod.

- Het is op onze school moeilijk om inzake de zorg voor leerlingen met onderwijsleerproblemen klasoverschrijdende initiatieven te nemen.
- Het schoolhoofd delegeert specifieke taken inzake de opvang van leerlingen met onderwijsleerproblemen naar één of meerdere leden van het schoolteam.

56) Continuïteit

Hier gaat het vooral om de aard van de informatie over leerlingen die wordt overgedragen van het ene leerjaar naar het andere. Daarnaast wordt ook gevraagd of er over de leerjaren heen afstemming is op elkaar wat betreft de leerinhouden en de gebruikte methoden.

- Wanneer leerlingen overgaan naar een volgend leerjaar wordt de volgende informatie overge-

dragen: * eventueel verleende hulp

* eventuele onderwijsleerproblemen

* rapportcijfers/schoolresultaten

- Op onze school stemmen de leerkrachten hun aanpak af op de gebruikte methoden van voorgaande en volgende leerjaren.

5.2 De vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen': klasniveau

Er werden 4 items (van de 77 van het deel 'klasniveau') weggelaten omwille van hun extreem gemiddelde en te lage spreiding. Dit waren de vier items met de kleinste standaarddeviatie. Alle andere items (73) werden behouden voor verdere analyse. Op basis van die 73 items werd ook hier een principale factoranalyse met varimaxrotatie uitgevoerd. De vierfactorenoplossing werd weerhouden. Ook hier baseerden we ons hoofdzakelijk op inhoudelijke criteria, waarmee we de interpreteerbaarheid van de factoren bedoelen (Fontaine, 1997). Deze vierfactorenoplossing verklaart 20% van de totale variantie. Ook hier wordt een gelijksoortige werkwijze gevolgd als met de factoren op schoolniveau en passen we enkele zaken aan (zie hoger). Uiteindelijk blijven 61 items over. Op basis van deze aangepaste indeling werden de α -coëfficiënten van Cronbach per schaal berekend (zie Tabel 3).

K1) Diagnostiserend en remediërend onderwijs

Deze schaal die we 'diagnostiserend en remediërend onderwijs' noemen, omvat items die hoofdzakelijk peilen naar het signaleren (opsporen), diagnosticeren (nadere exploratie) en remediëren van onderwijsleerproblemen.

- Ik ga na of mijn vermoedens kloppen omtrent de oorzaak van bepaalde fouten die een leerling met onderwijsleerproblemen maakt door hierop specifieke oefeningen te geven.
- Voor leerlingen met onderwijsleerproblemen gebruik ik aangepast oefenmateriaal.

K2) Ondersteuning en overleg

Hoewel deze schaal zich situeert op klasniveau sluit ze inhoudelijk nauw aan bij variabelen op schoolniveau. Het gezichtspunt is echter verschillend. Deze schaal peilt naar de mate waarin de individuele klasleerkracht zich ondersteund voelt door collega's, P.M.S.-centrum en anderen; in welke mate de leerkracht zich ook betrokken voelt bij overleg, informatie krijgt en geeft. Wordt de leerkracht op de hoogte gebracht van eventuele diagnostische onderzoeken bij een leerling? Onderneemt de leerkracht concrete stappen om informatie in te winnen bij P.M.S.-centrum, ouders, collega's,...? Kan de leerkracht terecht bij collega's voor eventuele problemen?

- Ik overleg soms met een andere leerkracht voor de aanpak van een onderwijsleerprobleem bij een leerling.
- Voor de opstelling van een handelingsplan kan ik binnen de school op weinig hulp rekenen.

K3) Opvattingen van de leerkracht

Hier wordt nagegaan hoe de leerkracht zijn/haar eigen mogelijkheden en deskundigheid inzake de zorg voor leerlingen met onderwijsleerproblemen percipieert. Daarnaast komen ook items aan bod over mogelijkheden en grenzen van de zorg voor leerlingen met onderwijsleerproblemen in de klas, over het vergroten van de zorg van het gewoon onderwijs, over het buitengewoon onderwijs, enzovoort.

- Ik mis de nodige vaardigheden om een handelingsplan in de praktijk goed uit te voeren.
- Ik vind dat het voor de ontwikkeling van leerlingen met onderwijsleerproblemen beter is als ze worden opgevangen in het buitengewoon onderwijs.

K4) Binnenklasdifferentiatie

In deze schaal komt het aspect van interne of binnenklasdifferentiatie aan bod. Het zijn items in verband met het differentiëren wat betreft het aantal taken, de inhoud van taken, wijze van evalueren en de mate waarin de leerkracht de klas organiseert met het oog op individueel en zelfstandig werken van leerlingen.

- Ik geef alle leerlingen uit mijn klas dezelfde taken.
- Om de leerlingen met onderwijsleerproblemen te helpen, verdeel ik de klas in deelgroepen.

6 Validiteit

In wat volgt bespreken we de belangrijkste bevindingen van de kwantitatieve analyses met het oog op een exploratie van de validiteit van de vragenlijst². Als eerste indicatie voor de validiteit van de vragenlijst staan we in Tabel 4 stil bij de correlaties tussen de 10 schalen onderling³. Ten eerste merken we op dat de correlaties tussen de schalen op schoolniveau (S1-S6) hoger liggen dan de correlaties tussen de schalen op schoolniveau enerzijds en de

Tabel 4

Correlatiematrix tussen alle schalen van de vragenlijsten 'Zorg voor leerlingen met onderwijsleerproblemen' op het niveau van de scholen (N=65 scholen)

	S1	S2	S3	S4	S5	S6	K1	K2	K3	K4
S1		.70*	-.73*	.47*	.60*	.46*	.23	.60*	.37*	.10
S2			-.74*	.66*	.59*	.31*	.30*	.74*	.24	-.09
S3				-.67*	-.57*	-.51*	-.14	-.72*	-.38*	.04
S4					.46*	.28*	.17	.67*	.12	-.28*
S5						.16	-.02	.54*	.39*	-.01
S6							.46*	.52*	.23	.25*
K1								.44*	.34*	.15
K2									.33*	-.08
K3										.21
K4										

* $p < 0,05$

schalen op klasniveau anderzijds. Evenwel vallen de hoge correlaties op tussen de schaal K2 ('Ondersteuning en overleg') en de schalen van dezelfde vragenlijst op schoolniveau (S1-S6). Dit bevestigt de validiteit van deze schaal (K2) aangezien ze, ofschoon ze gesitueerd is op klasniveau, ook inhoudelijk nauw aansluit bij het schoolniveau. Verder vinden we ook een bevestiging van de validiteit in het feit dat de schaal die een negatieve inhoud heeft (S3: 'Gebrek aan zorg voor leerlingen met onderwijsleerproblemen') negatief correleert met alle andere schalen van de vragenlijst (uitgezonderd K4, 'Binnenklasdifferentiatie').

Daarnaast werd per schaal onderzocht hoeveel procent van de variantie zich tussen en binnen scholen bevindt. Naarmate een schaal een aspect nagaat dat meer samenhangt met het niveau van de leerkracht, verwachten we hier relatief meer variantie binnen de school dan schalen die iets nagaan over de school. We verwachten met andere woorden dat de variantie tussen scholen bij schalen die zich op schoolniveau situeren verhoudingsgewijs groter zal zijn dan schalen die zich eerder op klasniveau situeren. In Tabel 5 geven we per schaal de procent variantie binnen en tussen scholen

Op basis van Tabel 5 merken we dat de resultaten grotendeels aan de verwachtingen voldoen. Schalen die op schoolniveau werden gesitueerd variëren meer tussen scholen dan schalen die op klasniveau werden gesitueerd. Enige uitzondering hierop is schaal S6 ('Continuïteit'). Wanneer we de inhoud van deze schaal nagaan, is het goed te begrijpen dat deze schaal inderdaad ook neigt naar het klasniveau,

meer dan naar het schoolniveau. Het doorgeven van informatie en afspraken omtrent leerstof kan zeer sterk afhankelijk zijn van leerkracht tot leerkracht.

Verder is het validiteitsonderzoek enerzijds gericht op de inhoudsvaliditeit en anderzijds op de begrips- of constructvaliditeit. Het op geregelde tijdstippen voorleggen van het instrument aan deskundigen en mensen uit de praktijk, alsook het verantwoorden van bepaalde keuzes vanuit de literatuur gebeurden met het oog op het garanderen van de inhoudsvaliditeit. Begripsvaliditeit behelst de vraag of een instrument in voldoende mate kan worden beschouwd als een maat van een theoretisch construct of begrip. Hiervoor gingen we de relatie na met bestaande vragenlijsten. In alle scholen van het definitief onderzoek werd bij de leerkracht naast de vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen' ook de vragenlijsten 'Professionele Cultuur van Basisscholen' en 'Onderwijsoriëntatie' afgenomen. Voor een uitvoerige beschrijving van het theoretisch kader en de constructie van deze vragenlijsten verwijzen we naar andere publicaties (voor de vragenlijst 'Professionele cultuur van basisscholen' zie Staessens, 1991b; voor de vragenlijst 'Onderwijsoriëntatie' zie Maes, Vandenberghe & Ghesquière, 1997). We geven hier slechts een summier beschrijving. De vragenlijst 'Professionele cultuur van basisscholen' situeert zich op schoolniveau en is opgebouwd rond drie domeinen waarin de professionele cultuur tot uiting komt. Na factoranalyse komt men tot vier schalen die aansluiten bij deze domeinen: 'Het schoolhoofd als cultuurbouwer en -drager' (C1), 'Overeen-

Tabel 5
Procent variantie tussen scholen per schaal van de vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen'

SCHALEN	VARIANTIE TUSSEN SCHOLEN	VARIANTIE BINNEN SCHOLEN
S1	39,0 %	61,0 %
S2	41,1 %	58,9 %
S3	23,5 %	76,5 %
S4	22,9 %	77,1 %
S5	35,4 %	64,6 %
S6	13,4 %	86,6 %
K1	9,0 %	91,0 %
K2	14,6 %	85,4 %
K3	13,7 %	86,3 %
K4	13,5 %	86,5 %

komst in doelgerichtheid' (C2), 'Professionele relaties tussen leerkrachten' (C3) en 'Gebrek aan een intern netwerk van professionele steun' (C4). De betrouwbaarheidscoëfficiënten liggen tussen .89 en .95. De vragenlijst 'Onderwijsoriëntatie van klasleerkrachten' situeert zich op leerkrachtniveau en omvat na factoranalyse twee schalen: 'Leerstofgerichte onderwijsoriëntatie' (O1) en 'Leerlinggerichte onderwijsoriëntatie' (O2). De betrouwbaarheidscoëfficiënten zijn respectievelijk .81 en .71.

Wat betreft de samenhang tussen de schalen van de ontwikkelde vragenlijst en de schalen van de reeds bestaande vragenlijsten (zie Tabel 6), merken we vooreerst de hoge correlaties op tussen S1 ('Het schoolhoofd en de gezamenlijke zorg voor leerlingen met onderwijsleerproblemen') enerzijds en C1 ('Schoolhoofd als cultuurbouwer en -drager'): $r=0,86$ en C2 ('Overeenkomst in doelgerichtheid'): $r=0,75$ anderzijds. Dit is volgens de verwachtingen. Verder correleren alle schalen op schoolniveau van beide vragenlijsten hoog met elkaar (S1-S6 en C1-C4). Daarbij zijn er telkens negatieve correlaties met de negatieve schaal van de vragenlijst 'Professionele cultuur van basisscholen': C4 ('Gebrek aan een intern netwerk van professionele steun'). Dit met uitzondering van S3 ('Gebrek aan zorg voor leerlingen met onderwijsleerproblemen') die zelf een negatieve schaal is.

K2 heeft tevens een grote, significante correlatie met alle schalen van de vragenlijst 'Professionele cultuur van basisscholen'. Dit sluit aan bij het feit dat deze schaal ook inhoudelijk meer aansluit bij het schoolniveau en meer specifiek bij het functioneren van het schoolteam. K4 ('Binnenklasdifferentiatie') correleert enkel significant – negatief – met O1 ('Leerstofgerichte onderwijsoriëntatie'). Ook hierin vinden we evidentie voor de begripsvaliditeit. Een school waarin de nadruk ligt op het doorgeven van leerstof, zal wellicht minder aan binnenklasdifferentiatie doen.

De schalen van de vragenlijst 'Professionele cultuur van basisscholen' correleren significant met nagenoeg alle schalen. De uitzonderingen zijn: S4 ('Samenwerking met het P.M.S.-centrum en met ouders') – misschien omdat daarin meer schoolexterne aspecten aan bod komen –, K1 (Diagnosticerend en remediërend onderwijzen) en K4 ('Binnenklasdifferentiatie'). Bij deze laatste twee betreft het duidelijk klasinterne aspecten.

Al bij al mogen we stellen dat de samenhang tussen de schalen van de ontwikkelde vragenlijst en de schalen van reeds bestaande vragenlijsten voldoende aanwijzingen geven voor de begripsvaliditeit van de schalen⁴. Schalen die een gelijkaardig concept (of 'begrip') nagaan, correleren immers hoog positief. Schalen met een negatieve of tegengestelde betekenis correleren eerder negatief.

Tabel 6

Correlatiematrix tussen de schalen van de vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen' enerzijds en de vragenlijsten 'Professionele cultuur van basisscholen' en 'Onderwijsoriëntatie' anderzijds, op het niveau van de scholen (N=65 scholen)

	C1	C2	C3	C4	O1	O2
S1	.86*	.75*	.68*	-.56*	-.02	.48*
S2	.48*	.37*	.45*	-.28*	-.17	.39*
S3	-.60*	-.53*	-.64*	.54*	.20	-.50*
S4	.38*	.24	.30*	-.16	-.16	.20
S5	.40*	.39*	.45*	-.28*	-.14	.31*
S6	.35*	.49*	.52*	-.52*	-.12	.38*
K1	.10	.18	.24	-.05	.15	.13
K2	.42*	.45*	.64*	-.39*	-.10	.50*
K3	.34*	.36*	.37*	-.25*	-.10	.25*
K4	.03	.16	.10	-.18	-.28*	.16

* $p < 0,05$

7 Normering

Er werd voor gekozen om de scores op de schalen van de vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen' op twee wijzen te normeren. Ten eerste werden de gegevens *relatief* genormeerd door van elke schaalscore de overeenkomstige percentielscore te berekenen. Zo kunnen we op basis van deze normen één bepaalde school situeren ten opzichte van haar vergelijkingsgroep.

Daarnaast werd op basis van een onderzoek bij de gemeenschapsinspectie van het gewoon basisonderwijs een criterium vastgesteld per schaal om aldus een *absoluut* vergelijkingspunt te verkrijgen. Door een relatieve normering kan men wel alle scores vergelijken met de referentiegroep, maar het is desalniettemin een meerwaarde om die scores ook te kunnen vergelijken met een absoluut criterium. Een vragenlijst werd voorgelegd aan alle gemeenschapsinspecteurs van het gewone basisonderwijs. Deze vragenlijst was identiek met deze die de klasleerkrachten ontvingen in het definitief onderzoek. Er werden wel enkele pagina's bijgevoegd met specifieke instructies en vragen over de respondent. Aan de inspecteurs werd gevraagd om per item een grens te trekken waarmee men kon aanduiden dat een score boven die grens (of onder die grens voor inhoudelijk negatief gestelde items) door hem/haar als positief wordt beoordeeld wat betreft de zorg voor leerlingen met onderwijsleerproblemen.

Op basis van de antwoorden van de inspectie ($N=42$) werden gemiddelden per schaal berekend op identieke wijze als van de gegevens van de scholen. Op die manier krijgen we een absoluut criterium waarmee de scores van scholen vergeleken kunnen worden. Zo kan men nagaan of een school of een groep van scholen voor de zorg voor leerlingen met onderwijsleerproblemen het criterium reeds behaalt of nog onder het criterium voorgesteld door de gemeenschapsinspectie blijft.

8 Enkele resultaten

Tot slot willen we enkele bevindingen beschrijven op basis van de ontwikkelde vragenlijst. Hierbij baseren we ons vooreerst op de

beschrijvende statistieken per schaal voor de steekproef van 65 scholen. Daarnaast vergelijken we deze statistieken met de criteria die we op basis van de antwoorden van de inspecteurs basisonderwijs vaststelden. In Tabel 7 vindt men in de tweede kolom het gemiddelde en de standaarddeviatie per schaal van de vragenlijst 'Zorg voor leerlingen met onderwijsleerproblemen' voor de 65 scholen. We vinden de laagste gemiddelden bij de schalen S5 ('Klasoverschrijdende initiatieven en coördinatie') en K4 ('Binnenklasdifferentiatie'), met uitzondering van de inhoudelijk negatieve schaal S3 ('Gebrek aan zorg voor leerlingen met onderwijsleerproblemen'). Blijkbaar vormen de interne differentiatie en externe differentiatie (daarop ligt immers het hoofddaccent in de schaal 'Klasoverschrijdende initiatieven en coördinatie') nog steeds een belangrijk knelpunt. Het hoogste gemiddelde vinden we bij S6 ('Continuïteit'). Wanneer we de standaarddeviaties bekijken dan zien we dat die kleiner zijn voor de schalen op klasniveau (K1-K4). Dit is te verklaren door het feit dat bij het aggregeren van resultaten van leerkrachtniveau naar schoolniveau de binnenschoolse variantie wordt 'weggemiddeld'. Dit komt vooral tot uiting bij schalen die veel binnenschoolse variantie vertonen, in casu de schalen op klasniveau, en dus minder bij schalen op schoolniveau waarbij variantie zich meer zal situeren tussen scholen.

In Tabel 7 geven we verder het gemiddelde en de standaarddeviatie per schaal voor de inspectie. Om na te gaan of deze gemiddelden significant verschillen van die van de scholen werd een t-toets uitgevoerd. De verschillen zijn niet significant voor drie van de tien schalen. We mogen met andere woorden stellen dat het gemiddelde van de scholen voor deze schalen gelijk is aan het gemiddelde (criterium) van de inspectie. Dit is zo voor de schalen 'Samenwerking met P.M.S.-centrum en met ouders' (S4), 'Continuïteit' (S6) en 'Diagnosticerend en remediërend onderwijzen' (K1). Voor alle andere schalen blijkt het gemiddelde van de scholen te verschillen van dat van de inspectie. Wat is nu de betekenis hiervan? Deze vaststelling zegt immers niets over de waarde van het criterium. We stellen enkel vast dat het gemiddelde van de scholen voor 7 van de 10 schalen

Tabel 7

Gemiddelden en standaarddeviaties van de schalen van de vragenlijst 'Zorg voor leerlingen met onderwijs-leerproblemen' voor scholen (N=65) en inspectie (N=42)

SCHAAL	SCHOLEN (65)		INSPECTIE (42)		t-TOETS	
	$\bar{x}$	SD	$\bar{x}$	SD		
S1	4,43	0,67	4,95	0,49	$t(105) = -4,334$	$p = 0,00^*$
S2	4,36	0,76	4,87	0,58	$t(105) = -3,705$	$p = 0,00^*$
S3	2,29	0,53	1,81	0,71	$t(105) = 3,996$	$p = 0,00^*$
S4	4,68	0,54	4,64	0,71	$t(105) = 0,330$	$p = 0,74$
S5	3,24	0,74	4,34	0,48	$t(105) = -8,536$	$p = 0,00^*$
S6	5,03	0,43	5,06	0,70	$t(105) = -0,275$	$p = 0,78$
K1	4,58	0,29	4,69	0,65	$t(105) = -1,195$	$p = 0,24$
K2	4,00	0,46	4,61	0,54	$t(105) = -6,253$	$p = 0,00^*$
K3	3,74	0,33	4,60	0,44	$t(105) = -11,529$	$p = 0,00^*$
K4	3,59	0,41	4,81	0,69	$t(105) = -11,476$	$p = 0,00^*$

* Deze gemiddelden verschillen significant van elkaar

significant onder dat van de inspectie ligt. We kunnen dit in twee richtingen interpreteren. Het kan betekenen dat volgens de inspectie de lagere scholen gemiddeld gezien wat betreft die zeven schalen nog een hele weg moeten afleggen. Men kan anderzijds stellen dat de normen (criteria) van de inspectie voor bepaalde variabelen misschien te hoog liggen.

We merken daarnaast op dat het gemiddelde van alle scholen voor drie schalen ver onder het criterium ligt. Dit is het geval voor 'Klasoverschrijdende initiatieven en coördinatie' (S5), 'Opvattingen van de leerkracht' (K3) en 'Interne differentiatie' (K4).

Tabel 8

Per schaal het percentage scholen (N=65) dat het criterium van de inspectie bereikt

SCHAAL	PERCENTAGE SCHOLEN DAT HET CRITERIUM BEREIKT
S1	18 %
S2	26 %
S3	16 %
S4	67 %
S5	4 %
S6	57 %
K1	34 %
K2	7 %
K3	0 %
K4	0 %

Er is evenwel spreiding in de schoolgemiddelden. Er zijn scholen die het criterium van de inspectie wel halen. Om na te gaan hoeveel scholen het gemiddelde criterium van de inspectie behaalden, bekeken we per schaal

welk percentiel overeenkomt met het criterium van de inspectie. Op basis daarvan geven we in Tabel 8 het percentage scholen dat het criterium van de inspectie haalt.

Op een andere manier merken we hier dat geen enkele of slechts een heel beperkt aantal scholen het criterium haalt voor de schalen 'Opvattingen van de leerkracht' (K3), 'Binnenklasdifferentiatie' (K4), 'Klasoverschrijdende initiatieven en coördinatie' (S5) en ook voor de schaal 'Ondersteuning en overleg' (K2). Daarnaast behaalt de meerderheid van de scholen het criterium voor de schalen 'Samenwerking met P.M.S. en ouders' (S4) en 'Continuïteit' (S6)⁵.

9 Besluit

Deze schriftelijk te beantwoorden vragenlijst laat volgens ons een screening toe van het diagnosticerend en remediërend vermogen van één of meerdere lagere scholen. Deze vragenlijst kan bijgevolg vooreerst op grote schaal worden afgenomen om op bepaalde tijdstippen snel en accuraat informatie over een brede groep lagere scholen en zelfs over alle lagere scholen te verkrijgen. Het kan met andere woorden dienen als instrument in een algemeen peilingsonderzoek dat het diagnosticerend en remediërend vermogen van het lager onderwijs wil nagaan.

Verder kunnen door middel van deze vragenlijst ook verschillen tussen scholen worden gedetecteerd of kan men één school situeren

ten opzichte van de normgroep en kan het instrument bijgevolg ook worden gebruikt voor meer schooldiagnostische doeleinden. De resultaten van elke school kunnen immers weergegeven worden door middel van een profiel op de onderscheiden schalen. Op die manier kunnen sterke en zwakke kanten van de zorg voor leerlingen met onderwijsleerproblemen in kaart gebracht worden. Hier moet men evenwel voor ogen houden dat dit instrument in deze context van schooldiagnose slechts kan dienen als aanzet tot verdere analyse. Deze vragenlijst kan dus van dienst zijn zowel voor de onderwijspraktijk (begeleiding, schoolhoofden, enz.), het onderwijsbeleid (ook inspectie) als het onderwijsonderzoek.

De vragenlijst heeft echter ook beperkingen. Ze laat immers niet toe om een gedetailleerd inzicht te geven in het specifieke van de zorg voor leerlingen met onderwijsleerproblemen van één bepaalde school. Aanvullende kwalitatieve gegevens (te verzamelen via andere onderzoekstechnieken zoals interviews, bezoeken, observaties) zijn daarom aan te raden indien men een rijker en gedetailleerder beeld wil krijgen van bepaalde scholen, zeker indien men concrete aanknopingspunten zoekt voor gerichte verbetering van de bestaande lokale onderwijspraktijk.

Noten

- 1 We willen hierbij ten eerste opmerken dat er naast deze hoofdactoren nog andere actoren zijn die instaan voor de leerlingenzorg. We denken hierbij in de eerste plaats aan het Psycho-Medisch-Sociaal Centrum (een schoolbegeleidingsdienst in Vlaanderen) en aan andere externe begeleiders. Ten tweede willen we benadrukken dat de leerkracht in dit model zowel een klasleerkracht als een taakleerkracht of ambulante leerkracht kan zijn.
- 2 De validiteit van de vragenlijst werd ook onderzocht door de scores verkregen met de vragenlijst te vergelijken met kwalitatieve data verzameld in vier scholen door middel van interviews (triangulering van de resultaten). Voor een uitgebreide bespreking hiervan verwijzen we naar Maes, Vandenberghé & Ghesquière (in press).
- 3 De hier vermelde analyses gebeurden met de schoolscores. Om de gegevens van de individuele

respondenten te aggregeren per school, werden enkel de gegevens van die scholen weerhouden waarvan we minstens 50% van de vragenlijsten van de klasleerkrachten en de vragenlijst van het schoolhoofd terugkregen. Er waren 65 scholen in dit geval. Deze groep van 65 scholen is representatief wat betreft de beschreven kenmerken (grootte, net, onderwijsniveau en provincie).

- 4 Voor de schalen op klasniveau (K1-K4) is de correlatiematrix van leerkrachten ($N=688$) meer aangewezen. Deze schalen betreffen immers kenmerken die meer samenhangen met leerkrachten zelf dan met de scholen. Die correlatiematrix bevestigt de validiteit van deze schalen nog beter. Zie hiervoor Maes, Vandenberghé en Ghesquière (1997).
- 5 De reden waarom 'slechts' 34% van de scholen het criterium van de inspectie haalt voor K1 ('Diagnostiserend en remediërend onderwijzen') ofschoon het de gemiddelden niet significant verschillen van elkaar, ligt in het feit dat de antwoordverdeling op deze schaal niet normaal verdeeld is (niet symmetrisch is) en bijgevolg de mediaan niet overeenkomt met het gemiddelde.

Literatuur

- ARBO: Adviesraad voor het basisonderwijs, speciaal onderwijs en de opleidingen (1984). *Het moet ons een zorg zijn. Advies over zorgverbreding in het basisonderwijs*. Zeist: Onderwijscentrum A.R.B.O.
- Bartelds, J.F., Jansen, E.P.W.A., & Joostens, T.H. (1989). *Enquête: het opstellen en gebruiken van vragenlijsten*. Groningen: Wolters-Noordhoff.
- Billiet, J. (1990). *Het ontwerpen van enquêtes: de gestandaardiseerde vragenlijst*. Leuven: Sociologisch Onderzoeksinstituut, Katholieke Universiteit Leuven.
- Billiet, J. (1992). *Methoden van sociaal-wetenschappelijk onderzoek: ontwerp en dataverzameling*. Leuven: Acco.
- Clement, M. (1995). *De professionele ontwikkeling van leerkrachten Basisonderwijs. De spanning tussen autonomie en collegialiteit*. Niet-gepubliceerd doctoraatsproefschrift, Centrum voor Onderwijsbeleid en -vernieuwing, Katholieke Universiteit Leuven.
- Damme, J. Van (1992-1993). *Schoolse vertraging en zittenblijven: frekwentie, betekenis en beleving*

- door de leerling. *Tijdschrift voor Onderwijsrecht en Onderwijsbeleid*, 2, 103-119.
- Damme, J. Van (1993). *Cahier 3. Over het aantal jongeren met vertraging in ons onderwijs*. Brussel: VI.O.R.
- Dienst voor Onderwijsontwikkeling (1994). *Periodiek peilingsonderzoek naar het diagnostiserend en remediërend vermogen van scholen gericht op minder uitval van leerlingen. Onderzoekstraject*. Niet-gepubliceerd document, Ministerie van de Vlaamse Gemeenschap, Departement Onderwijs.
- Dun, K. Van (1990). *De ontwikkeling van het Buitengewoon Onderwijs in België*. Niet-gepubliceerde licentiaatsverhandeling, Faculteit der Psychologie en Pedagogische Wetenschappen, Katholieke Universiteit Leuven.
- Dun, K. Van, & Ruijsenaars, A.J.J.M. (1991). Omvang en groei van het Buitengewoon/Special Onderwijs in Vlaanderen en Nederland. Een onderzoek naar de omvang en evolutie van het leerlingaantal. *Tijdschrift voor Orthopedagogiek, Kinderpsychiatrie en klinische Kinderpsychologie*, 16, 181-192.
- Fontaine, J. (1997). *Handleiding bij de practica bij Vraagstukken uit de psychologische statistiek. Deel III: Multivariate methoden met SAS*. Leuven: Acco.
- Ghesquière, P., Fever, F. De, & Hove, G. Van (1996). *Op weg naar inclusief onderwijs? Elementen voor de discussie rond de hervorming van het Buitengewoon Onderwijs in Vlaanderen in functie van de internationale tendens tot integratie*. Leuven/Brussel/Gent: Vlaamse Onderwijsraad, Vaste Themagroep: Relatie Gewoon-Buitengewoon Onderwijs.
- Gielis, H., & Gerits, R. (1989). Op weg naar een effectieve school. *School en Begeleiding*, 21, 49-57.
- Grift, W. van de (1987). *De rol van de schoolleider bij onderwijsvernieuwingen*. 's-Gravenhage: Lisbon/Vuga.
- Huistra, A. (1980). *Signaleren, diagnosticeren en remediëren als functie van de school*. Haren: RION.
- Huynegem, J. Van, & Vercammen, E. (1991). Onderwijsleerproblemen: een geïntegreerde benadering. In E. Vercammen & J. Van Huynegem (Red.), *Leerproblemen op school: niet inpakken, maar aanpakken* (pp. 15-37). Brussel: Centrale Raad van het Katholiek Lager- en Kleuteronderwijs (CRKLKO).
- Janssens, J.M.A.M. (1988). *Ogen doen onderzoek*. Lisse/Amsterdam: Swets & Zeitlinger.
- Knapp, T. R. (1982). The unit and the context of analysis for research in educational administration. *Educational Administration Quarterly*, 18, 1-13.
- Knuver, J. W. M., & Reezigt, G. J. (1991). Zittenblijvers in het basisonderwijs. *Didaktief*, 21, 4-9.
- Kool, E., & Leij, A. van der (1985). Planmatig handelen. In A. van der Leij (Red.), *Zorgverbreding. Bijdragen uit speciaal onderwijs aan basisonderwijs* (pp. 69-89). Nijkerk: Intro.
- Laevers, F. (1994). Zorgverbreding in het kleuteronderwijs. In VI.O.R.-Forum Basisonderwijs (Red.), *Zorgverbreding: onze zorg. Forum Basisonderwijs 30 april 1994* (pp. 127-146). Brussel: VI.O.R.
- Leij, A. van der (1996). De RVC. In R. de Groot, A.J.J. M. Ruijsenaars & H. Kapinga (Red.), *Inclusief onderwijs* (Tijdschrift voor Orthopedagogiek, jaargang 35, afl. 12) (pp. 83-100). Groningen: Wolters-Noordhoff.
- Leij, A., van der & Kappers, E.J. (Red.). (1994). *Zorgverbreding. Bijdragen uit speciaal onderwijs aan basisonderwijs* (6de herziene ed.). (Speciaal onderwijs en zorgverbreding). Nijkerk: Intro.
- Maes, F., Luyckx, E., Vandenbergh, R., & Ghesquière, P. (1996). *Het signalerend, diagnosticerend en remediërend vermogen van Vlaamse lagere scholen* (Tussentijds rapport OBPWO 94.01). Katholieke Universiteit Leuven, Centrum voor Onderwijsbeleid en -vernieuwing.
- Maes, F., Vandenbergh, R., & Ghesquière, P. (1997). *Het diagnosticerend en remediërend vermogen van lagere scholen* (Eindrapport OBPWO 94.01). Katholieke Universiteit Leuven, Centrum voor Onderwijsbeleid en -vernieuwing.
- Maes, F., Vandenbergh, R., & Ghesquière, P. (in press). The imperative of complementarity between the school level and the classroom level in educational innovation. *Journal of Curriculum Studies*.
- OESO (1991). *Het educatief bestel in België: van convergentie naar divergentie. (OESO: Doorlichting van het educatief overheidsbeleid)*. Brussel-Bruxelles-Eupen: Ministerie van de Vlaamse Gemeenschap, Departement Onderwijs - Ministère de l'Éducation, de la Recherche et de la Formation - Verwaltung der Deutschsprachigen Gemeinschaft, Abteilung Unterricht.
- Pijl, Y.J. (1997a). Basisonderwijs en de groei van LOM en MLK. *Tijdschrift voor Orthopedagogiek*, 36, 160-170.
- Pijl, Y. J. (1997b). Verblijfsduur als verklaring voor de groei van het speciaal onderwijs. *Tijdschrift voor Orthopedagogiek*, 36, 226-234.

Segers, J.H.G. (1987). *Methoden voor de sociale wetenschappen*. Assen: Van Gorcum.

Smylie, M.A. (1994). Redesigning teachers' work: connections to het classroom. *Review of Research in Education*, 20, 129-177.

Smylie, M.A. (1995). Teacher learning in the workplace. Implications for school reform. In T.R. Guskey & A.M. Huberman (Eds.), *Professional development in education. New paradigms and practices* (pp. 92-113). New York/London: Teachers College Press.

Smylie, M.A. (1996). From bureaucratic control to building human capital: the importance of teacher learning in education reform. *Educational Researcher*, 25(9), 9-11.

Span, B. (1990). *Meting van zorgverbreding. Instrumentontwikkeling* (Eindrapport SVO-project 8010). Groningen: RION.

Staessens, K. (1991a). *De professionele cultuur van basisscholen. Elke school heeft haar verhaal* (Studia Paedagogica). Leuven: Universitaire Pers.

Staessens, K. (1991b). Ontwikkeling en validatie van de Vragenlijst Professionele Cultuur van Basisscholen. *Pedagogische Studiën*, 68, 241-252.

Stokking, K.M. (1992). *Wijkende condities. Differentiatie en zorgverbreding in het basisonderwijs: onderzoek en beleid 1984-1990*. Utrecht: ISOR-afdeling Onderwijsonderzoek Universiteit Utrecht.

Sudman, S., & Bradburn, N.M. (1987). *Asking questions. A practical guide to questionnaire design*. San Francisco: Jossey-Bass.

Swanborn, P.G. (1982). *Schaaltechnieken. Theorie en praktijk van acht eenvoudige procedures*. Meppel: Boom.

Swanborn, P.G. (1987). *Methoden van sociaal-wetenschappelijk onderzoek*. Meppel: Boom.

Tacq, J. (1992). *Van probleem naar analyse. De keuze van een gepaste multivariate analysetechniek bij een sociaal-wetenschappelijk probleemstelling*. De Lier: Academisch Boeken Centrum.

Vandenberghe, R. (1990). Over schoolverbetering. Of hoe scholen leerkrachten kunnen steunen. *Impuls*, 20(4), 161-172.

Vandenberghe, R. (1992). *Onderwijsvernieuwing of schoolverbetering. Een discussienota voor de Commissie Onderwijsvernieuwing Basisonderwijs*. Centrum voor Onderwijsbeleid en -vernieuwing, Katholieke Universiteit Leuven.

VI.O.R.-Forum Basisonderwijs. (Red.). (1994). *Zorgverbreding: onze zorg. Forum Basisonderwijs 30 april 1994*. Brussel: VI.O.R.

Vlaamse Onderwijsraad. (1993). *Cahier 5. Zittenblijven in het basisonderwijs en in het secundair onderwijs. Advies van de Vlaamse Onderwijsraad*. Brussel: VI.O.R.

Manuscript aanvaard 5-10-1998

Auteurs

F. Maes is wetenschappelijk medewerker aan het Centrum voor Onderwijsbeleid en -vernieuwing, Katholieke Universiteit Leuven.

R. Vandenberghe is gewoon hoogleraar en hoofd van het Centrum voor Onderwijsbeleid en -vernieuwing, Katholieke Universiteit Leuven.

P. Ghesquière is als hoofddocent werkzaam aan de Afdeling Orthopedagogiek, Katholieke Universiteit Leuven.

Adres: Vesaliusstraat 2, 3000 Leuven, België.

E-mail: Frederik.Maes@ped.kuleuven.ac.be

Abstract

Construction and validation of the questionnaire 'The care for pupils with educational problems'

F. Maes, R. Vandenberghe & P. Ghesquière. *Pedagogische Studiën*, 1999, 76, 21-35.

Because of the alarming reports of school failure in primary and secondary schools, educationalists increasingly hold the conviction that regular education should pay more attention to children with learning problems. Central to this problem is the diagnostic and remedial capacity of the schools. This capacity is closely linked to the extension of the special needs provision. This article describes the construction and validation of an instrument that assesses the diagnostic and remedial capacity of a school. The instrument is based upon a framework that reflects the aspects that should be present in primary schools in order to help pupils with learning problems. Based upon this model, 10 scales were developed through a factor analysis, 6 scales on the school level and 4 scales on the classroom level. The reliability and validity of the questionnaire were studied and were satisfying.

Het concept activerende instructie gemeten via de perceptie van leerlingen

R. Lamberigts, P.J. den Brok, K. Derksen en T. Bergen

Samenvatting

Als leren van leerlingen omschreven kan worden als het ondernemen van leeractiviteiten, dan moet instructie omschreven worden als het nemen van maatregelen die bevorderen dat leerlingen die leeractiviteiten ook uitvoeren, zodat er daadwerkelijk sprake is van de beoogde veranderingen in kennis, vaardigheden en opvattingen. Het instructiegedrag van docenten dat dit beoogt te bewerkstelligen wordt door ons activerende instructie genoemd. Binnen het concept activerende instructie onderscheiden wij drie typen van activering; docent-geleide, medeleerling-geleide en leerling-geleide activering. In deze bijdrage wordt het concept activerende instructie gelegitimeerd vanuit verschillende constructivistische leer- en kennis-theorieën, en wordt de operationalisering van dit concept via de Vragenlijst InstructieGedrag (VIG) nader onderzocht. Voor de vooronderstelde theoretische relaties tussen de drie activeringscomponenten van de VIG wordt door middel van een structurele analyse (LISREL) empirische ondersteuning gevonden. Het blijkt mogelijk de voor activering veronderstelde a-priori structuur te repliceren. De betrouwbaarheid van de activeringsschalen in termen van interne consistentie blijkt bevredigend.

Inleiding

Het voortgezet onderwijs is onder invloed van het overheidsbeleid volop in beweging. Een belangrijke vernieuwing in het voortgezet onderwijs, naast de invoering van de Basisvorming, is de invoering in 1998 van het zogenaamde 'Studiehuis' in de tweede fase van het voortgezet onderwijs. Met de benaming 'Studiehuis' wordt een plaats of instelling bedoeld waar het accent ligt op het studeren door leerlingen in plaats van op het onderwijzen door docenten. Van docenten wordt verwacht dat zij

in het studiehuis het actief en zelfstandig leren van hun leerlingen bevorderen. Afgezien van de vraag hoe deze vernieuwingen kunnen worden ingevoerd (zie o.a. Bergen & Derksen, 1995; Bergen & De Jong, 1996; Bergen, 1996), moet in onze ogen eerst helderheid geschapen worden over welk gedrag van docenten het zelfstandig leren van de leerlingen kan bevorderen. Uit onderzoek blijkt dat het voor de effectiviteit van het leren en de cognitieve ontwikkeling van de leerling het best is wanneer informatieverwerking actief kan verlopen, met een nadruk op eigen initiatief, activiteit, autonomie en zelfregulatie (cf. Shuell, 1988; Resnick, 1987). Het gedrag van docenten richt zich dan op het stimuleren van de persoonlijke verantwoordelijkheid van leerlingen. Wij noemen dit gedrag 'activerende instructie'.

In deze bijdrage gaan wij nader in op de theoretische en empirische onderbouwing van het concept activerende instructie. Wij bespreken het concept activerende instructie eerst in algemene zin en plaatsen dit concept binnen een constructivistische opvatting over kennis en leren. Daarna presenteren wij een door ons ontwikkelde vragenlijst 'Vragenlijst InstructieGedrag' (VIG) die het concept activerende instructie via de perceptie van leerlingen beoogt te meten. Vervolgens zullen wij ingaan op de validiteit van de activeringsfactoren binnen de VIG, in het bijzonder op de inhoudsvaliditeit en op de constructvaliditeit. Ook besteden wij aandacht aan de betrouwbaarheid van de activeringsschalen, waarbij wij ons richten op de interne consistentie. Wij sluiten de bijdrage af met een discussie over de representativiteit en bruikbaarheid van het gevonden model en de rol die leerlingpercepties daarbij zouden kunnen spelen.

1 Het concept activerende instructie gelegitimeerd

Het concept activerende instructie is gebaseerd op een mix van opvattingen over leren, die allemaal stellen dat leren een actief, constructief, cumulatief en doelgericht proces van lerenden is (Shuell, 1988). Actief, omdat de lerende zich inspannt en zelf activiteiten onderneemt. Constructief, omdat nieuwe informatie wordt uitgediept en gerelateerd aan reeds aanwezige informatie om zowel eenvoudige leerstof te onthouden als complexe leerstof te begrijpen. Cumulatief, omdat alle nieuwe informatieverwerking stoelt op reeds aanwezige kennis en daarvan gebruik maakt. Doelgericht, omdat leren als succesvol wordt beleefd wanneer de leerling zich bewust is van een doel en weet waarom hij/zij leert. Psychologisch onderzoek naar leren (Brown, 1987; Shuell, 1988) laat zien dat het verwerven van kennis en vaardigheden niet zozeer het resultaat is van een directe overdracht van de docent op de leerlingen, maar eerder het resultaat is van denkactiviteiten van de leerlingen zelf. Leren wordt gezien als een in hoge mate interactief proces van construeren van nieuwe kennis en vaardigheden op basis van reeds aanwezige informatie bij een persoon (Glaser, 1991; Marzano, 1992). Deze recente opvattingen over leren hebben geleid tot een cognitief georiënteerde visie op instructie. Belangrijk in de nieuwe visie op instructie is het begrip constructivisme, dat pogt te verklaren hoe leerlingen zich een beeld construeren van de instructie die de docent geeft, waarbij zij zich een cognitieve representatie vormen van wat de gegeven instructie betekent. Deze centrale aanname – in wat nu een constructivistische kennis-, leer- en onderwijsfilosofie genoemd wordt – houdt in dat kennis en vaardigheden geen kant en klare producten zijn, die rechtstreeks kunnen worden overgedragen, maar het resultaat zijn van de

eigen leeractiviteiten van de lerenden (Glaser, 1991; Philips, 1995).

De constructivistische kennis- en leerfilosofie kent vele variëteiten die op hun beurt ieder eigen consequenties voor het onderwijzen hebben (Prawat & Floden, 1994). Hierbij zijn drie versies van het constructivisme dominant; een radicaal-individualistische; een sociaal-situatieve en een metacognitieve benadering. Vanuit deze kennis- en leertheorieën kunnen instructietheorieën worden afgeleid die hun uitwerking vinden in drie typen activering (zie Figuur 1). Deze typen geven tezamen vorm aan het concept activerende instructie.

De eerste kennis- en leertheorie die binnen het constructivisme kan worden onderscheiden – de radicaal-individualistische benadering – benadrukt de uniciteit van het individu en de singuliere wijze waarop deze de werkelijkheid ervaart en vormgeeft (Von Glasersfeld, 1987; 1995). Het radicaal-individualistische constructivisme legt de basis voor een extreme vorm van differentiatie in het onderwijs, waarbij ieder individu nagenoeg uniek is. Zulk een differentiatie is slechts mogelijk met behulp van moderne media in een krachtige leeromgeving in enge zin. In zulk een leeromgeving is autonoom leren door leerlinggeleide activering (zie § 1.1) mogelijk. De tweede benadering – de sociale, situatieve of sociaal-constructivistische – benadrukt de groep waarin kennis door samenwerking van individuen met elkaar wordt opgebouwd. Kennis zelf wordt beschouwd als een sociaal artefact, dat in stand gehouden wordt door een gemeenschap van ontwikkelingsgelijken. In de literatuur zijn binnen deze versie met het oog op onderwijskundige consequenties twee varianten geformuleerd. De ene (Vygotsky, 1987) stelt weliswaar dat al het denken sociaal van aard is, maar dat dit denken beïnvloed wordt door de expert (docent, tutor) binnen de zone van de naaste ontwikkeling. De andere (Dewey, 1960; Grec-

	Leerling-geleid	Medeleerling-geleid	Docent-geleid
Constructivistische kennis/leertheorie	Radicaal-individualistische benadering	Sociaal-situatieve benadering	Metacognitieve benadering
Instructietheorie	Krachtige leeromgevingen in 'enge' zin.	Coöperatief leren	Effectieve instructie op strategisch handelen

Figuur 1. Onderbouwing van de drie typen activering in schema

no, 1997) benadrukt de constructie van kennis door ontwikkelingsgelijken. In beide varianten wordt een legitimeringsbasis gelegd voor een leeromgeving waarin de sociale interactie van leerlingen centraal staat (Lamberigts, 1993). Het sociaal constructivisme legt de basis voor coöperatief leren in de brede zin van het woord als vorm van medeleerling-geleide activering (§1.2). Zowel de radicaal-individualistische als de sociaal-situatieve versie van het constructivisme op zichzelf leggen echter geen legitimeringsbasis voor het leren leren. De metacognitieve benadering (Flavell, Miller & Miller, 1993; Baker & Brown, 1984), die verwijst naar een stroming waarin kennis van en stuurvaardigheid in zelfstandig leren centraal staan, zou beschouwd kunnen worden als een variant van het constructivisme, dan wel als een aanvulling op de 'cognitieve wende', waarin de mens als informatieverwerker fungeert. In verband met het activerende karakter van het onderwijzen willen wij, binnen deze derde benadering, een basis leggen voor met name de vaardigheden en strategieën die direct verband houden met het leren leren. Daarbij wordt niet zozeer gedoeld op eenvoudige cognitieve leerstrategieën, maar meer op complexe vaardigheden die betrekking hebben op de vormgeving van het leren (regulatiestrategieën). In verband met onderwijzen is het ontwikkelen van deze centrale leervaardigheid van groot belang. Daarbij is deze centrale leervaardigheid gebonden aan vakspecifieke domeinen, waarin deze leervaardigheden aanpakstrategieën worden genoemd. De metacognitieve theorie of de theorie van het leren leren legt de basis voor een instructie van strategisch handelen. Het constructivisme met betrekking tot metacognitieve vaardigheden legitimeert effectief onderwijs in strategische leervaardigheden binnen een bepaald kennisdomein i.c. de aanpakstrategieën. Dit onderwijs heeft het karakter van docent-geleide activering (§1.3). Als nu het doel van een activerende instructie het bevorderen van het actief en zelfstandig kunnen leren van de lerende is, dan is het niet voldoende om aspecten van activerende instructie naast elkaar te beschrijven, maar moet ook aangegeven worden dat vanuit de praktijk van het onderwijs deze drie aspecten geïntegreerd kunnen en moeten worden (zie §1.4) in één integraal concept (Salomon, 1997).

1.1 Leerling-geleide activering

De literatuur met betrekking tot de krachtige leeromgeving en wel de 'enge' versie daarvan (cf. Lodewijks, 1993), sluit aan bij de radicaal-individualistische constructivistische visie. Bij de beschrijving van de krachtige leeromgeving in enge zin gaat het meestal om technologisch geavanceerde voorzieningen waarbij de geïntegreerde inzet van interactieve media kenmerkend is. Op deze wijze kunnen complete en rijke contexten uit de werkelijkheid nagebootst worden. Daardoor wordt het leren levensecht en interessant, maar – en dat is belangrijker – kan het leren op de individuele behoeften, leerstijlen en interesses van de lerenden worden toegesneden. Als de nabootsing van de werkelijkheid in levensechte casussen ook nog probleemgericht wordt aangeboden, worden de lerenden uitgenodigd om nieuwe kennis te verzamelen c.q. te construeren. Het betreft hier derhalve rijke en betekenisvolle probleemcontexten (zogenaamde 'anchored instruction' en 'cognitive apprenticeship' arrangementen), die tot volwaardig leren uitdagen.

Vanuit het concept activerende instructie is het van belang dat lerenden op deze manier de ruimte wordt gegeven om in eigen tempo te leren, te ontdekken dat alle kennis geconstrueerd is en opnieuw ge(re)construeerd kan worden en dat iedere kennis relatief is en vanuit verschillende perspectieven geïnterpreteerd kan worden.

Binnen het concept activerende instructie wordt zelfverantwoordelijk en autonoom leren beschouwd als een uiteindelijk doel, dat implementatietechnisch een laatste stap representeert en conceptueel aangevuld moet worden. In onze huidige onderwijscultuur kan een onmiddellijke focus op autonomie – hoe belangrijk ook – leiden tot het kapitaliseren op cognitieve, metacognitieve, motivationele en communicatieve vaardigheden die er nog niet zijn (Lamberigts, 1996). Zonder expliciete aandacht voor aanpakstrategieën in een docent-geleide activering (zie §1.3) bestaat het gevaar dat krachtige leeromgevingen verzanden in ongeleid project-onderwijs (Laurillard, 1993).

1.2 Medeleerling-geleide activering

De beweging van het coöperatief leren heeft vele onderwijsarrangementen opgeleverd (Lamberigts, 1988, 1993) die allen geënt zijn

op het sociaal-constructivisme, dat co-constructie van kennis beoogt. Lange tijd heeft de beweging van het coöperatief leren een geheel eigen en daardoor wat geïsoleerde positie binnen de onderwijsvernieuwing ingenomen, waardoor de bijdrage aan een concept van activerende instructie onderbelicht is gebleven.

Kenmerkend voor bijna alle arrangementen van coöperatief leren (Slavin, 1980; Johnson, Johnson, Holubee & Roy, 1984) is de noodzaak voor de lerenden om met elkaar samen te werken bij het uitvoeren van een leertaak om een resultaat voor alle betrokkenen te bereiken. Dat resultaat heeft betrekking op een gelijktijdige verbetering van de cognitieve, metacognitieve, sociale en motivationele competentie. Recente meta-analyses (Johnson & Johnson, 1992; Slavin, 1995) schetsen een zeer positief beeld van de effecten van onderwijs gericht op coöperatief leren. De belangrijkste mechanismen die verantwoordelijk zijn voor het succes en verklaren waarom arrangementen van coöperatief leren 'werken' zijn de volgende:

De positieve wederzijdse afhankelijkheid van de lerenden van elkaar

De lerenden worden zich ervan bewust dat ze hun persoonlijk doel slechts bereiken als alle andere lerenden in de groep ook hun doel bereiken. Dit stimuleert de eigen individuele verantwoordelijkheid maar ook de verantwoordelijkheid voor elkaar en voor de groep. Beide verantwoordelijkheden zijn intrinsiek met elkaar verbonden, zodat eigenbelang en het belang van anderen samengaan in een bepaalde groeps cultuur. Deze interdependentie maakt dat lerenden lotsverbonden en positief tot elkaar veroordeeld zijn. Dit in tegenstelling tot een wedijversituatie, waarin lerenden negatief tot elkaar veroordeeld zijn. Positieve interdependentie werkt uiteraard niet als deze wordt toegepast in een situatie die als een wedijversituatie wordt beleefd.

Het mechanisme van 'peer tutoring'

Bij peer-tutoring fungeert een meer ervaren lerende als docent/tutor voor een minder ervaren lerende (tutee of novice). In het algemeen zijn de effecten voor de tutor hoger dan voor de tutees. Daarom verdient het aanbeveling dat peer-tutoring rolwisselend plaatsvindt (Lambert, 1990). De situatie is leerzaam voor de

tutor, omdat deze in de rol van de docent status verkrijgt die hem/haar motiveert om de taak zo goed mogelijk uit te voeren. Het feit dat de tutor als docent zijn taak voorbereidt en uitvoert, leidt tot (her)structurering van kennis en activeert voorkennis (Webb, 1982). Het voordeel voor de tutees is dat zij sneller, frequenter en directer feedback krijgen van een docent naar (eigen) maat. Het heeft er alle schijn van, dat een docent-voor-medelerenden-zijn bevordert dat men ook een docent voor zichzelf kan zijn, inclusief alle docentvaardigheden die als zodanig strategische leervaardigheden zijn.

Het mechanisme van het 'echte' samenwerken of 'collaborative learning'

Bij collaborative learning (McCartney & McMahon, 1992) construeren ontwikkelingsgelijken kennis door met elkaar te discussiëren over bepaalde opgaven die meestal een probleemkarakter hebben. Het veronderstelde constructieproces wordt gefaciliteerd als controverses ontstaan of worden ingebouwd. De verschillen van kennis of mening zijn dan de noodzakelijke voorwaarden voor het ontstaan van een intern (socio-cognitief) conflict binnen iedere individuele persoon. Deze situatie wordt als onaangenaam beleefd en leidt onder bepaalde voorwaarden tot een gezamenlijke oplossing door het zoeken van aanvullende informatie.

Bovengenoemde mechanismen hebben ons geïnspireerd voor het medeleerling-geleide onderwijs, als onderdeel van ons concept activerende instructie. Aanpakstrategieën zijn wel ingebouwd in dit model maar vragen om een nadere en meer expliciete formulering in docent-geleide activering.

1.3 Docent-geleide activering

Binnen het terrein van de effectieve instructie (Rosenshine & Stevens, 1986) heeft zich een beweging voltrokken van een behavioristisch en taakanalytisch-georiënteerde opvatting over leren naar een meer cognitief georiënteerde opvatting. Men heeft nu meer aandacht voor de rol van cognities van leerlingen als vorm van mediëring tussen instructiegedragingen en leerresultaten (Veenman, 1992). Uiteraard heeft deze wijziging in opvatting consequenties voor het model van onderwijzen. Binnen het directe instructiemodel heeft zich derhalve ook een verschuiving voltrokken van een eng

naar een breed model. Dit impliceert dat de gerichtheid op overdracht van kennis en expliciete vaardigheden heeft plaatsgemaakt voor een gerichtheid op het modelleren van 'impliciete' vaardigheden.

Rosenshine en Edmonds (1990) bedoelen met deze impliciete vaardigheden leerstrategieën die van belang zijn bij het strategisch handelen en die in het model dat zich richtte op uitvoerend handelen niet expliciet geïmpliceerd werden. Het gaat hier derhalve om een wending van directe instructie van uitvoerend handelen naar strategisch handelen (cognitive strategy instruction). Wij vinden in de verbreding en verschuiving van dit model een legitimatie en voorbeeld voor ons begrip docent-geleide activering. Het onderwijzen van deze strategische vaardigheden geschiedt door middel van modelleren (hardop denken) door de docent van deze vaardigheden met betrekking tot procedures (stap-plannen) en handelingswijzers. Ook de instructie op zelfwerkzaamheid (het aangeven van het 'wat', 'hoe', 'met welk doel' etc. van de opdracht) vindt zijn legitimatie in dit verbrede concept. De bijdrage van het instructiemodel voor strategisch handelen is een relevante aanvulling op die van de krachtige leeromgeving in enge zin en van het coöperatief leren. De nadruk ligt op aanpakstrategieën die van belang zijn voor het uitvoeren van opdrachten binnen de les, maar ook voor het uitvoeren van huiswerk en proefwerken. Op deze wijze wordt werk gemaakt van onze opvatting dat zelfstandig leren geleerd wil worden. De aanpakstrategieën zijn naderhand (Rosenshine & Meister, 1992) 'hogere orde cognitieve strategieën' genoemd om te benadrukken dat deze vaardigheden eerder procesgericht dan productgericht zijn. Het gevaar bestaat namelijk dat binnen de traditie van de directe instructie deze vaardigheden als producten extern worden overgedragen in plaats van intern geïntegreerd in het bestaande arsenaal van vaardigheden.

1.4 Een integraal concept van activerende instructie

Tot nu toe zijn de drie typen van activerende instructie (leerling-geleid, medeleerling-geleid en docent-geleid) afzonderlijk beschreven en geïmpliceerd. Steeds meer wordt echter duidelijk dat de drie typen in combinatie met elkaar

een gunstig effect op elkaar en op het leren van de leerling hebben (Greeno, 1997; Salomon, 1997).

Medeleerling-geleide activering veronderstelt – wil deze zinvol en effectief zijn – dat eerst individuele – zij het voorlopige en beperkte – voorkennis is opgehaald. Medeleerling-geleide activering kan vervolgens als een middel voor leerling-geleide activering fungeren (Lamberigts, 1996). Het uiteindelijke doel van het onderwijs is weliswaar autonoom leren, maar bij het nastreven van dit doel wordt wel eens vergeten dat dit autonoom leren per vakdomein geleerd moet worden. Docent-geleide activering (Lamberigts, 1996) achten wij daartoe een voorwaarde.

Deze drie typen van activering zijn conceptueel met elkaar verbonden maar hebben ook een zekere sequentie als men denkt aan de implementering van activerende instructie. Zo sluit de docent-geleide activering goed aan bij het traditioneel goed, effectief en direct onderwijs. Met medeleerling-geleide activering richt de docent zich op het instrueren en begeleiden van de relatie tussen leerlingen, hetgeen in ons onderwijs aansluit bij de aandacht voor het groepswork. Bij leerling-geleide activering vindt een indirecte vorm van begeleiding plaats, die het verst afstaat van de gangbare onderwijspraktijk. Zowel voor medeleerling-geleide activering, waarin strategische aanpakstrategieën zichtbaar worden, als voor leerling-geleide activering, waarin gekapitaliseerd wordt op aanpakstrategieën, is docent-geleide activering een noodzakelijke voorwaarde. Het voordeel van deze sequentie voor de implementatie van activerende instructie is dat de docent geleidelijk aan wordt ingevoerd in de overgang van 'direct' naar 'indirect' onderwijzen, een overgang waarin weerstanden van zowel docenten als leerlingen overwonnen moeten worden.

2 Activerende instructie geoperationaliseerd: de ontwikkeling van de vragenlijst instructiegedrag (VIG)

In Nederland wordt al sinds enige tijd onderzoek gedaan naar het gedrag van docenten in

relatie tot het leren van de leerlingen door middel van vragenlijsten. Zo onderzochten Créton en Wubbels (1984) en Brekelmans (1989) door middel van de Vragenlijst Interpersoonlijk Leraarsgedrag (VIL) het omgangsgedrag van docenten met hun leerlingen. In hun meting trachtten zij het betrekkingaspect van het docentgedrag, zoals waargenomen door de leerlingen, te operationaliseren. Roosendaal, Vermunt en Thijssen, (1994) ontwikkelden de Inventaris Doceerstrategieën (IDS) waarmee het aansturen van leeractiviteiten door docenten in kaart kan worden gebracht en waarbinnen een tweetal doceerstijlen, namelijk een overnemende en een activerende, worden onderscheiden.

In het kader van het onderzoek 'De invloed van de kwaliteit van de instructie op de motivationele oriëntatie en de schoolcarrière van leerlingen' (SVO-project no. 0357) is de Vragenlijst InstructieGedrag (VIG) ontwikkeld en geëvalueerd (Van Amelsvoort, Bergen, Lamberigts & Setz, 1993; Setz, Bergen, Van Amelsvoort & Lamberigts, 1993; Bergen, Van Amelsvoort & Setz, 1994). De Vragenlijst InstructieGedrag (VIG) is ontwikkeld om het concept activerend instructiegedrag van docenten, zoals hierboven weergegeven, in kaart te brengen. Naast vormen van activering zijn in de VIG ook de kernvaardigheden duidelijkheid en controle opgenomen. De VIG is een 'high-inference' observatie-instrument waarmee patronen van instructiegedrag van docenten tijdens de les via de perceptie van de leerlingen in kaart kunnen worden gebracht.

In zijn oorspronkelijke vorm bestond de VIG uit 48 Likert-items met een vijfpuntschaal die betrekking hebben op activiteiten en gedragingen van docenten tijdens de les. De antwoordschalen lopen van 'bijna nooit' tot 'bijna altijd' en aan de leerlingen wordt gevraagd aan te geven wat zij het meest van toepassing achten op het gedrag van hun docent. Over de inhoud van de items geven de auteurs aan:

"Zo zijn er items geformuleerd met betrekking tot duidelijkheid van de docent op het gebied van het geven van uitleg, feedback en opdrachten; met betrekking tot controle van de docent van het lesverloop en de activiteiten van de leerlingen daarbinnen; en met betrekking tot het oriënteren van leer-

lingen op de leerstof en de les.(...) Daarnaast bevat de VIG items die verwijzen naar docentactiviteiten waarin het stimuleren en activeren van leerlingen tot uiting komt. Enerzijds gaat het hierbij om het actief betrekken van leerlingen (...), het aandacht besteden aan de aanpak van het schoolwerk en het aandacht besteden aan initiatieven van de leerlingen. Anderzijds gaat het hier om het bevorderen van zelfstandig werken, het stimuleren van eigen verantwoordelijkheid bij leerlingen, het geven van ruimte aan leerlingen om zelf activiteiten te ontplooiën en het laten samenwerken van de leerlingen." (Van Amelsvoort, Bergen, Lamberigts & Setz, 1993: 47-48).

Van de VIG bestaan verschillende versies (Van Amelsvoort et. al., 1993; Setz et. al, 1993). In deze bijdrage is gebruik gemaakt van een versie met 33 items, welke dikwijls ingezet wordt in het kader van scholingsinterventies (zie Bijlage 1).

3 Vraagstelling

In deze bijdrage, die zich beperkt tot de activeringsschalen van de VIG, gaat het enerzijds om de construct- en inhoudsvaliditeit van de meting van activerende instructie en anderzijds om de betrouwbaarheid van de meting.

De volgende drie vragen staan centraal:

1. Kan het theoretisch model van activering met de daarin veronderstelde onderverdeling van drie activeringsvormen zoals deze a priori aan de VIG ten grondslag ligt empirisch worden gerepliceerd?
2. Zijn de in de VIG te onderscheiden activeringsfactoren (schalen) betrouwbaar?
3. Op welke wijze dragen de in de VIG te onderscheiden activeringsfactoren (schalen) en activeringsindicatoren (items) bij aan het onderliggende latente concept activering?

4 Methode

4.1 Onderzoeksgroep

In het kader van interventies met betrekking tot het bevorderen van het actief en zelfstandig leren van leerlingen werd de VIG afgenomen op een zestal scholen voor voortgezet onder-

wijs. De VIG werd tweemaal afgenomen, aan het begin van de interventies (oktober) en aan het eind van de interventies (april). Voor de analyses is gebruik gemaakt van de gegevens van de eerste meting. In totaal ging het om 1504 leerlingen, die les kregen van 61 docenten. Van de respondentengroep is 48.5 procent van het mannelijk geslacht. Verder kan opgemerkt worden dat 67.3 procent van de leerlingen op het VWO zit, 22.7 procent op het HAVO. De grootste groep leerlingen is afkomstig uit de bovenbouw van het VO, namelijk het vierde (75.5 procent) en vijfde (9.3 procent) leerjaar; een kleinere groep (12.7 procent) komt uit de onderbouw. Er zitten geen leerlingen uit het eerste en zesde leerjaar in deze groep.

4.2 Procedure

Het analyseniveau is in feite het activerende instructiegedrag van docenten zoals gepercipieerd door de leerlingen die in de klas van de betreffende docent zitten. Er zijn dus twee niveaus: het klasniveau en het leerlingenniveau. Als input voor de analyses op klasniveau zou er gewerkt moeten worden met de gemiddelde leerlingenscore per klas voor de items. De verhouding tussen het aantal waarnemingen (61 docenten) en het aantal items (33) maakte zo een analyse technisch niet verantwoord. Daarnaast zijn wij geïnteresseerd in de vraag of

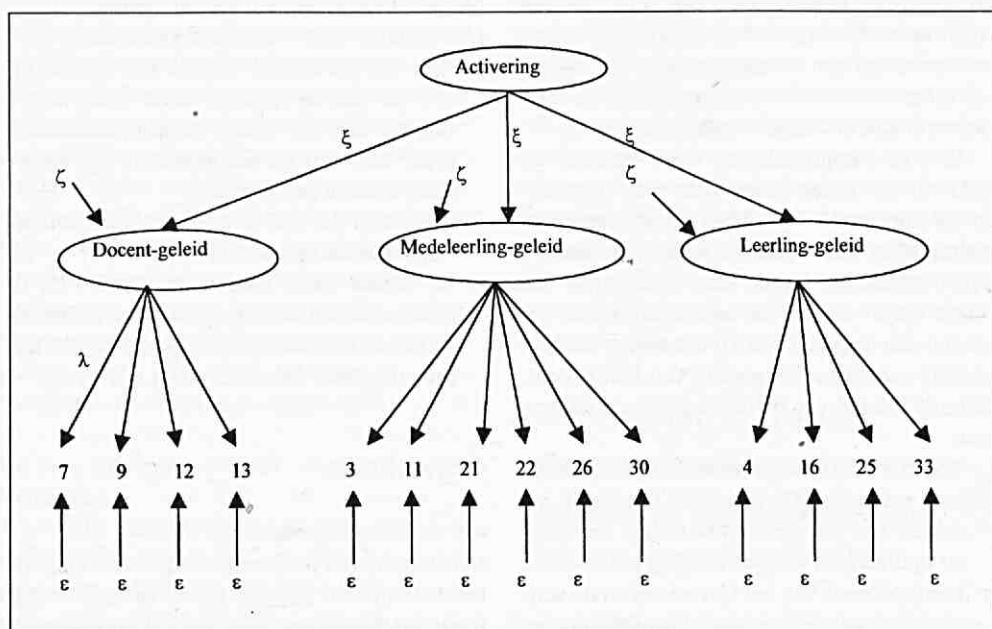
het model van activering ook aan individuele percepties van leerlingen ten grondslag ligt. Om die redenen hebben wij ervoor gekozen om de analyses op leerlingniveau uit te voeren.

Om vraag 1 te beantwoorden is gebruik gemaakt van LISREL-analyse. Vraag 2 is beantwoord door middel van betrouwbaarheidsanalyses en voor vraag 3 is opnieuw gebruik gemaakt van de uitkomsten van de LISREL-analyse.

4.3 Resultaten

Omdat de VIG items bevat die verwijzen naar de aspecten duidelijkheid van instructie en controle op leerlinggedrag is eerst met behulp van het statistische pakket SPSS een exploratieve factoranalyse uitgevoerd op alle 33 items. Deze factoranalyses leverden drie factoren op die niet op het concept activering betrekking hadden, namelijk twee duidelijkheidsfactoren en een controlefactor. De items van deze factoren zijn verwijderd. Daarnaast zijn alle items van de activeringsfactoren met een factorlading kleiner dan .40 verwijderd, evenals items met hoge (>.30) dubbele lading op meerdere factoren. Van de 33 items bleven na eliminatie volgens deze criteria 14 items over. De overige analyses zijn op deze 14 items (3, 4, 7, 9, 11, 12, 13, 16, 21, 22, 25, 26, 30 en 33) uitgevoerd.

Vervolgens zijn we nagegaan of het theore-



Figuur 2. Hypothetisch model voor de structuur van het concept activerende instructie binnen de VIG

tische model van activering met de daarin te onderscheiden activeringsfactoren in de VIG kon worden teruggevonden. Daartoe zijn de geselecteerde 14 items eerst op basis van een inhoudelijke validering toegekend aan telkens één van de drie veronderstelde activeringsschalen: de docent-geleide activeringsschaal (activering waarbij de docent zich vooral toelegt op het onderwijzen van strategieën en waarbij initiatieven van de leerlingen gestimuleerd en gewaardeerd worden); de medeleerling-geleide activeringsschaal (het activeren en stimuleren om leerlingen met elkaar samen te laten werken) en de leerling-geleide activeringsschaal (activering waarbij de leerling min of meer zelfstandig leert).

Deze itemdeling is vervolgens met een confirmatieve factoranalyse met behulp van LISREL getoetst. Daarbij zijn wij ervan uitgegaan dat de items niet alleen laden op de activeringsfactoren, maar tevens indirect op het achterliggende latente concept activering. Daarnaast leveren de activeringsfactoren een rechtstreekse bijdrage aan het latente activeringsconcept. Aldus ontstaat een tweede-orde factoranalyse (zie Figuur 2).

Bij zo'n tweede-orde factoranalyse laden items op schalen (λ -parameters), die op hun beurt weer op een achterliggende factor of latent concept laden (ξ -parameters). Als gevolg van het optreden van onzuiverheden in de meting, bijvoorbeeld door interpretatieproblemen bij de leerlingen of door het niet volledig omvatten van een factor door items, treden er meetfouten op, zowel op itemniveau (ϵ -parameters) als op schaalniveau (ζ -parameters). Al deze parameters kunnen door LISREL in kaart gebracht worden. Daarnaast is het mogelijk om met LISREL te kijken in hoeverre het model als geheel bij de data past en welke indirecte effecten er zijn (in ons geval bijvoorbeeld de effecten tussen het concept activering en de afzonderlijke items).

Ten slotte geeft LISREL mogelijkheden voor verbeteringen van het model aan. Aan deze verbeteringen hebben wij een aantal restricties opgelegd. Ten eerste willen wij vanwege de zuiverheid van het model geen ladingen toestaan van items op meerdere activeringsfactoren. Ten tweede houden wij rekening met de mogelijkheid dat, vanwege de overeenkomsten

in formulering van items binnen een schaal, meetfouten van items met elkaar correleren. Dergelijke correlaties tussen meetfouten bij items staan wij alleen toe tussen items die tot dezelfde activeringsfactor behoren. Ten derde mogen de meetfouten die optreden bij de schalen onderling niet correleren, het betreft hier immers verschillende aspecten van het concept activering.

Voor de procedure die ingezet kan worden bij het fitten van een tweede-orde factormodel verwijzen wij naar Verschuren (1991). Kort gezegd komt de analyse erop neer dat we eerst voor de drie afzonderlijke activeringsfactoren aparte LISREL-modellen hebben getoetst en de uitkomsten daarvan als vaste waarden (fixaties) hebben ingevoerd in het totale model zoals beschreven in Figuur 1. Vervolgens zijn de overige waarden binnen het model geschat.

Het model uit Figuur 1 blijkt goed bij de data te passen. Echter, LISREL geeft aan dat er nog enige afstand bestaat tussen het model en de data. Bovendien geeft het programma nog enkele verbeteringen aan, maar deze blijken in strijd te zijn met de eerdere restricties die we aan het model hadden opgelegd. Daarom hebben we van deze verbeteringen afgezien. In Bijlage 2 staan alle geschatte waarden weergegeven.

Voor de docent-geleide activeringsfactor blijkt het item "hij/zij besteedt aandacht aan de manier waarop je je huiswerk kunt aanpakken" de hoogste lading te hebben, terwijl het item "aan het eind van de les herhaalt hij/zij de belangrijkste punten van de lesstof" de laagste lading heeft. De docent-geleide activeringsfactor blijkt inhoudelijk het beste te duiden als het aandacht besteden door de docent aan strategieën om (zelfstandig) te leren. Daarbij speelt de docent nog een belangrijke rol, maar aan de leerlingen worden technieken aangeleerd om later meer zelfstandig te kunnen leren.

Voor de medeleerling-geleide activeringsfactor is het item met de beste fit: "als we samenwerken stimuleert hij/zij dat wij verantwoordelijkheid voor elkaar nemen". Het item met de zwakste fit is "hij waardeert het als we zelf initiatieven nemen". De items waarin expliciet het samenwerken aan bod komt, al dan niet onder begeleiding of op initiatief van

de docent, blijken voor deze factor het beste te passen.

Bij de leerling-geleide activeringsfactor blijken alle items naar verhouding sterk te laden. Het sterkst ladende item is het item "bij hem/haar krijg je de ruimte om zelf te plannen". Die items waarbij de leerling van de docent een grote ruimte krijgt om beslissingen te nemen binnen het leerproces blijken het beste bij deze factor te passen.

Uit de LISREL-analyse blijkt duidelijk dat er nog een aantal zwakke items in de VIG zit. Deze items worden gekenmerkt door een relatief lage lading, een relatief grote meetfout en het model kan bij deze items nauwelijks variantie verklaren (zie Bijlage 2). Voorbeelden van zulke items zijn item 12 ("aan het eind van de les behandelt hij/zij de belangrijkste punten van de lesstof"), item 22 ("hij/zij waardeert het als we zelf initiatieven nemen") en item 4 ("hij/zij maakt ons van tevoren duidelijk wat hij/zij tijdens een les wil gaan doen"). Het is opvallend dat al deze items een weinig 'activerend' karakter hebben.

Onderzoeksvraag 2 heeft betrekking op de betrouwbaarheid van de drie activeringsschalen in termen van de interne consistentie (Cronbachs alpha). De schaal docent-geleide activering blijkt een betrouwbaarheid van .75 te hebben, de factor medeleerling-geleide activering heeft eveneens een betrouwbaarheid van .75, maar de betrouwbaarheid van de leerling-geleide activering is met .59 matig betrouwbaar. Dit geeft aan dat met name deze laatste schaal nog verder verbeterd kan worden. Het is

overigens interessant om te zien dat de uitkomsten van de analyses naar de interne consistentie de gegevens van de LISREL-analyse lijken tegen te spreken. De daar optredende meetfouten bij de schalen (zie Bijlage 2) zijn het grootst bij de docent-geleide activeringsfactor, terwijl zij juist het kleinst zijn bij de leerling-geleide activeringsfactor.

De derde onderzoeksvraag heeft betrekking op de relatieve bijdrage van de activeringsfactoren en -indicatoren op het latente concept activering. Een eerste antwoord op die vraag kan worden afgeleid uit bijlage 2. Daaruit blijkt de docent-geleide activeringsfactor het sterkst te laden op het latente concept activering (.487), terwijl ook de medeleerling-geleide activeringsfactor tamelijk sterk laadt (.449). Opvallend is dat de leerling-geleide activeringsfactor een veel minder sterke bijdrage levert (.195). Het achterliggende activeringsconcept lijkt dus vooral gedekt te worden door docent-geleide en medeleerling-geleide activering en veel minder door leerling-geleide activering.

De LISREL-analyses leveren ook informatie op over de indirecte relaties uit het model, in ons geval de relatie tussen het activeringsconcept en de afzonderlijke items. In Tabel 1 staan de relaties vermeld tussen het activeringsconcept en de items met daarbij de meetfout die is opgetreden.

De sterkste relaties blijken te bestaan tussen het activeringsconcept en de items "hij/zij besteedt aandacht aan de manieren waarop je je huiswerk kunt aanpakken" en "als we samenwer-

Tabel 1
Relaties tussen de activeringsitems (indicatoren) en het (latente) activeringsconcept

Relatie van activering met...	Effectwaarde	Meetfout
Item 3	.517	.022
Item 4	.221	.016
Item 7	.492	.000
Item 9	.572	.000
Item 11	.440	.000
Item 12	.260	.000
Item 13	.457	.000
Item 16	.233	.017
Item 21	.568	.000
Item 22	.374	.000
Item 25	.284	.021
Item 26	.460	.000
Item 30	.421	.000
Item 33	.317	.023

ken stimuleert hij/zij dat we verantwoordelijkheid voor elkaar nemen", terwijl de zwakste relaties bestaan met de items "hij/zij maakt van tevoren duidelijk wat hij/zij wil gaan doen tijdens de les" en "aan het eind van de les behandelt hij/zij de belangrijkste punten van de lesstof". Ook uit deze resultaten blijkt dat het activeringsconcept vooral te duiden is aan de hand van het aandacht besteden aan strategieën en het stimuleren van samenwerken door de leerlingen. Veel minder blijkt het activeringsconcept te duiden in termen van meer directe instructie op het handelen van de leerlingen en het volledig loslaten van de verantwoordelijkheid. Het lijkt erop dat de VIG vooral het middengebied van activering in kaart brengt en veel minder de uiteinden van de activeringsdimensie. Dit betekent natuurlijk niet dat deze vormen van activering niet kunnen voorkomen, noch dat zij onwenselijk zouden zijn.

5 Conclusie en discussie

Uit deze onderzoeksresultaten concluderen wij dat het theoretische concept activerende instructie door middel van de VIG binnen de door ons geselecteerde itempool adequaat wordt geoperationaliseerd. Er is empirische evidentie, dat de VIG een valide en betrouwbaar instrument is om vormen van activerend docentgedrag te meten. Het conceptuele onderscheid tussen de drie activeringsvormen van docent-geleide, medeleerling-geleide en leerling-geleide activering wordt empirisch gerepliceerd.

De VIG biedt een goed uitgangspunt voor het verder doordenken van het construct activerende instructie vanwege de fit tussen het theoretisch model en de empirische operationalisatie. Daarnaast doen zich verschillende mogelijkheden voor vervolgonderzoek voor. Zo kunnen de inhoudsvaliditeit en de betrouwbaarheid van de drie vormen van activering verder verbeterd worden. Met name de schaal leerling-geleide activering is een kandidaat voor zo'n verbetering.

In dit onderzoek is geen sprake geweest van een representatieve respondentengroep. Het betrof een docentengroep die zich op vrijwillige basis heeft aangemeld voor nascholing op het gebied van zelfstandig en actief leren en die

als zodanig waarschijnlijk al meer open stond voor en kennis had van ideeën over hoe het actief leren van de leerlingen bevorderd zou kunnen worden. Het gevonden model geldt dan ook vooralsnog alleen voor deze groep. Onderzoek met een grotere en meer representatieve steekproef van docenten met hun klassen is in de toekomst nodig om de validiteit van het model meer definitief aan te tonen. Een grotere steekproef is tevens nodig om analyses te kunnen uitvoeren op klasniveau. Zoals reeds gezegd is dit het geëigende analyseniveau om het lesgedrag van docenten via de perceptie van leerlingen vast te stellen. Daarnaast is onduidelijk in hoeverre het gevonden model stabiel is over de tijd en wellicht kan veranderen onder invloed van interventies. Verder verdient het aanbeveling de stabiliteit van het model over subpopulaties, bijvoorbeeld over schooltypen, verschillende klassen van dezelfde docenten, verschillende leerjaren, et cetera, nader vast te stellen. Dergelijke analyses geven verder inzicht in de validiteit van het gevonden model, maar voeren hier gezien de beperkte respondentengroep te ver.

Wij willen er de nadruk op leggen dat de VIG een instrument is dat de perceptie van de leerlingen weergeeft in vrij algemene activeringstermen. Dit levert mogelijkheden en beperkingen op met betrekking tot de inhoud en de bruikbaarheid. In de huidige vorm van de VIG zijn items vak- en activiteitsoverstijgend geformuleerd. We merkten eerder al op dat activering gekoppeld is aan specifieke vakinhoud, i. e. dat zij bij elk vak op een specifieke wijze geïmplementeerd kan en moet worden. Daarbij lenen bepaalde les- of leeractiviteiten zich ook in meer of mindere mate voor bepaalde vormen van activering. Aangezien de VIG activering niet op een dergelijke specifieke wijze in kaart brengt is het de vraag of dit leerlingen in staat stelt een voldoende gedifferentieerde opinie te vormen.

Daarnaast zijn aan het gebruik van leerling-percepties verschillende voor- en nadelen verbonden (Brekelmans, 1989; d'Apollonia & Abrami, 1996). Eén van de voordelen is dat de docent feedback kan krijgen over activeringsvormen zoals hij/zij deze in de beleving van de leerlingen uitvoert. Omdat de belevingskwaliteit erg algemeen van aard is, geeft deze feedback een indruk van de mate van 'goed zijn'

van de docent met betrekking tot activering in de ogen van de leerlingen. Deze feedback zou vervolgens een plaats moeten krijgen naast de zelfevaluatie van de docent en de 'objectieve' evaluatie door middel van de observatie- of checklistmethode. Deze triangulering is nuttig voor zowel onderzoeks- als scholingsdoeleinden. Als op zichzelf staand instrument levert de VIG in het kader van de genoemde doelstellingen echter nog te beperkte informatie op. Daarvoor zijn checklists en observatieinstrumenten nodig die in de ogen van getrainde observatoren weergeven in hoeverre activeringsgedrag inderdaad wordt gerealiseerd in de praktijk. Omdat de VIG in empirische zin een representatie (en in zekere zin een bevestiging) van het beschreven conceptuele systeem van activeren-de instructie is, zelfs in de ogen van de leerlingen, denken wij dat het zinvol is om de item-pool binnen het aangegeven kader in kwantitatieve zin te verruimen en in kwalitatieve zin te verbeteren, zodat de vragenlijst naast observaties kan worden ingezet en op die manier bruikbaar gemaakt wordt voor interventiedoeleinden.

De resultaten stimuleren tevens het verder doordenken over hoe het actief en zelfstandig leren van leerlingen door docenten langs lijnen van geleidelijkheid bevorderd kan worden. Dit vooronderstelt dat er meer inzicht is in het soort gedrag dat docenten kunnen realiseren om vormen van activering tijdens hun lessen te implementeren. Dit brengt ons bij de vraag of er richtlijnen kunnen worden afgeleid voor de training van docenten in vormen van activeren-de instructie op basis van de VIG-scores (Derksen, Engelen, Bergen & Slegers, in press).

Literatuur

- Amelsvoort, J. van, Bergen, T., Lamberigts, R., & Setz, W. (1993). *De invloed van de kwaliteit van instructie op de motivationele oriëntatie en de schoolcarrière van leerlingen: deelrapport 1 'Docentgedrag, leerlingmotivatie en schoolprestaties'*. Nijmegen: VON/ITS.
- Apollonia, S. d', & Abrami, P. (1996, april). *Variables moderating the validity of student ratings of instruction: a meta-analysis*. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Baker, L., & Brown, A.L. (1984). Metacognitive skills and reading. In P.D. Pearson (Ed.), *Handbook of reading research* (vol.1.) (pp. 354-394). New York: Longman.
- Bergen, T. (1996). *Docenten scholen docenten. over de professionele ontwikkeling van docenten door middel van peer coaching*. Inaugurale rede Katholieke Universiteit Nijmegen.
- Bergen, T., Amelsvoort, J. van, & Setz, W. (1994). Het lesgedrag van docenten in relatie tot de vakspecifieke motivatie van leerlingen. *Pedagogische Studiën*, 71, 256-270.
- Bergen, T., & Derksen, K. (1995). Perspectief op zelfstandig leren. In *Handboek Tweede Fase (afl.3)*, (pp. 6.11.1-6.11.9). Houten: Bohn Stafleu van Loghum.
- Bergen, T., & Jong, F. de (1996). Het scholen van docenten in het actief en zelfstandig leren: een Nijmeegse benadering. *Studiehuisreeks*, 8, 49-63. Tilburg: Meso Consult.
- Brekelmans, M. (1989). *Interpersoonlijk gedrag van docenten in de klas*. Utrecht: W.C.C.
- Brown, A.L. (1987). Metacognition, executive control, self-regulation and other more mysterious mechanisms. In F. Weinert & R. Kluwer (Eds.). *Metacognition, motivation and understanding* (pp. 65-116). Hillsdale: Erlbaum.
- Créton, H., & Wubbels, H. (1984). *Ordeproblemen bij beginnende leraren*. Utrecht: W.C.C.
- Derksen, K. (in voorbereiding). *Tussen sturen en activeren*. Academisch proefschrift.
- Derksen, K., Engelen, A., Bergen, T., & Slegers, P. (in press.). *Het leren door docenten*. Meso-focus reeks.
- Dewey, J. (1960). *The quest for certainty*. New York: Capricorn.
- Flavell, J.H., Miller, P.H., & Miller, S.A. (1993). *Cognitive development*. Englewood-Cliffs: Prentice Hall.
- Glaser, R. (1991). The maturing of the relationship between the science of learning and cognition and educational practice. *Learning and Instruction*, 1, 129-144.
- Glaserfeld, E. von (1987). Preliminaries to any theory of representation. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 215-225). Hillsdale: Erlbaum.
- Glaserfeld, E. von (1995). *Radical constructivism: a way of knowing and learning*. London: The Falmer Press.
- Greeno, J.G. (1997). Response: on claims that answer the wrong question. *Educational Researcher*, 26 (1), 5-17.

- Johnson, D.W., Johnson, R.T., Holubee, E., & Roy, P. (1984). *Circles of learning*. Alexandria: ASCD.
- Johnson D.W., & Johnson, R.T. (1992). Positive interdependence: key to effective cooperation. In R. Hertz-Lazarowitz, & N. Millar (Eds.), *Interaction in cooperative groups* (pp. 174-202). Cambridge: University Press.
- Lamberigts, R. (1988). Coöperatief leren. *Onderwijskundig Lexicon* (C1300, 1-19). Alphen a/d Rijn: Samsom.
- Lamberigts, R. (1990). De helpende leerling. *Handboek Leerlingbegeleiding* (2345). Alphen a/d Rijn: Samsom.
- Lamberigts, R. (1993). De klas: een communicatieve context voor leren. *Onderwijskundig Lexicon* (C 1200, 1-24). Alphen a/d Rijn: Samsom.
- Lamberigts, R. (1996). *Activerende instructie als concept*. Paper gepresenteerd op Onderwijs Research Dagen, Tilburg.
- Laurillard, D. (1993). *Rethinking university teaching: a framework for the effective use of educational technology*. London: Routledge.
- Lodewijks, J. (1993). *De kick van het kunnen. Over arrangement en engagement bij het leren*. Tilburg: OKZ.
- Marzano, R.J. (1992). *A different kind of classroom: teaching with dimensions of learning*. Alexandria: ASCD.
- McCarthy, S.J.M., & McMahon, S.M. (1992). From convention to invention: three approaches to peer interactions during writing. In R. Hertz-Lazarowitz, & N. Miller (Eds.), *Interaction in cooperative groups*, (pp. 17-35). Cambridge: University Press.
- Philips, D.C. (1995). The good, the bad and the ugly: the many faces of constructivism. *Educational Researcher*, 24 (7), 5-12.
- Prawat, R.S., & Floden, R.E. (1994). Philosophical perspectives on constructivist views of learning. *Educational Psychology*, 29 (1), 37-48.
- Resnick, L.B. (1987). Learning in school and out. *Educational Researcher*, 16 (9), 13-20.
- Roosendaal, L.A., Vermunt, J.D.H.M., & Thijssen, A.P.J. (1994). *I.D.S.: inventaris doceerstrategieën*. Tilburg: Katholieke Universiteit Brabant/Theresialyceum.
- Rosenshine, B. & Stevens, R. (1986). Teaching functions. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.) (pp.376-391). New York: MacMillan Inc.
- Rosenshine, B., & Edmonds, J. (1990). New sources for improving instruction: the implicit skills study. *Journal of Personnel evaluation in education*, 4, 59-73.
- Rosenshine, B., & Meister, C. (1992). The use of scaffolds for teaching higher-level cognitive strategies. *Educational Leadership*, 49 (7), 26-33.
- Salomon, G. (1997, august). *Novel constructions learning environments and novel technologies: some issues to be concerned with*. Paper presented at the conference of the European Association for Research on Learning and Instruction, Athens.
- Setz, W., Bergen, T., Amelsvoort, J. van, & Lamberigts, R. (1993). *De invloed van de kwaliteit van instructie op de motivationele oriëntatie en de schoolcarrière van leerlingen: deelrapport 2 'Gepercipieerd en geobserveerd lesgedrag van docenten'*. Nijmegen: VON/ITS.
- Shuell T.J. (1988). The role of the student in learning from instruction. *Contemporary Educational Psychology*, 13, 276-295.
- Slavin, R. (1980). *Using student teamlearning*. Baltimore: The Johns Hopkins University.
- Slavin R. (1995). *Cooperative learning: theory research and practice* (2nd edition). Boston: Allyn & Bacon.
- Veenman, S. (1992). Effectieve instructie volgens het directe instructiemodel. *Pedagogische Studiën*, 69, 242-269.
- Verschuren, P.J.M. (1991). *Structurele modellen tussen theorie en praktijk*. Utrecht: Het Spectrum (aula-reeks).
- Vygotsky, L. (1987). Thinking and speech. In R.W. Lieber, & A.S. Carton (Eds.). *The collected works of L.S. Vygotsky: Problems of general psychology* (Vol. I) (pp. 37-285). New York: Plenum.
- Webb, N. (1982). Student interaction and learning in small groups. *Review of Educational Research*, 52, 421-455.

Auteurs

R. Lamberigts, coördinator student-activerend onderwijs, onderwijsinstituut PWO Nijmegen, Postbus 9104, 6500 HE Nijmegen.

P.J. den Brok, onderzoeker in opleiding, Interfacultair Instituut voor Lerarenopleiding, Onderwijsontwikkeling en Studievoordigheden (IVLOS) Utrecht.

K. Derksen, assistent in opleiding, Vakgroep Onderwijskunde Nijmegen.

T. Bergen, hoogleraar-wetenschappelijk directeur, Universitair Instituut voor de Lerarenopleiding Nijmegen (UNILO).

Correspondentie-adres: P.J. den Brok, IVLOS-Universiteit Utrecht, Postbus 80 127, 3508 TC Utrecht.

Abstract

The concept of teaching for active learning measured through pupil perceptions

R. Lamberigts, P.J. den Brok, K. Derksen & T. Bergen. Pedagogische Studiën, 1999, 76, 36-50.

When learning can be described as performing learning activities by pupils it follows that instruction

can be described as taking measures that make it possible for pupils to perform these activities, resulting in enduring changes in their knowledge, skills and attitudes. Teaching for active learning encompasses those behaviors of teachers that are aimed at realizing this. Within the concept of teaching for active learning a distinction is made between teacher-led, co-pupil led and pupil-led teaching for active learning. The concept of teaching for active learning finds its legitimation in different constructivist approaches on learning and knowledge. In this contribution, the operationisation of teaching for active learning by means of the Questionnaire of Instructional Behavior (QIB) is investigated. Linear Structural Analyses provide evidence for the assumed theoretical model behind the concept of teaching for active learning. Reliability in terms of internal consistence is found to be satisfying.

Items van de Vragenlijst Instructie Gedrag (VIG)

- 1 Hij/zij eist de volledige aandacht als hij/zij iets uitlegt.
- 2 Tijdens zijn/haar les kun je gemakkelijk andere dingen doen.
- 3 Hij/zij stimuleert ons elkaar te helpen bij het maken van opdrachten. (*)
- 4 Hij/zij maakt ons van tevoren duidelijk wat hij/zij tijdens een les wil gaan doen. (*)
- 5 Hij/zij gaat samen met ons na waarom een bepaald antwoord goed is.
- 6 Als je niet precies doet wat hij/zij zegt, dan corrigeert hij/zij je.
- 7 Hij/zij besteedt aandacht aan hoe je voor een proefwerk kunt studeren. (*)
- 8 Als het nodig is, is hij/zij bereid van zijn/haar lesopzet af te wijken.
- 9 Hij/zij besteedt aandacht aan manieren waarop je je huiswerk kunt aanpakken. (*)
- 10 Bij hem/haar weet je precies wat je wel en niet voor een proefwerk moet kennen.
- 11 Hij/zij laat ons samen met medeleerlingen nadenken over hoe je een bepaalde opdracht het beste aan kunt pakken. (*)
- 12 Aan het eind van de les herhaalt hij/zij de belangrijkste punten van de behandelde lesstof. (*)
- 13 Tijdens de les besteedt hij/zij aandacht aan hoe je je schoolwerk kunt plannen. (*)
- 14 Hij/zij geeft precies aan wanneer iets goed is.
- 15 Hij/zij herhaalt kort de lesstof die in eerdere lessen behandeld is.
- 16 We mogen bij hem/haar zelf bepalen hoe we aan opdrachten werken. (*)
- 17 Als hij/zij iets uitlegt is hij/zij moeilijk te volgen.
- 18 Bij hem/haar in de les heb je je precies aan de regels te houden.
- 19 Hij/zij stimuleert je om actief mee te doen.
- 20 Of je meedoet met de les, mag je bij hem/haar zelf uitmaken.
- 21 Als we samenwerken stimuleert hij/zij dat we verantwoordelijkheid voor elkaar nemen. (*)
- 22 Hij/zij waardeert het als we zelf initiatieven nemen. (*)
- 23 Hij/zij stelt prijs op de inbreng van leerlingen.
- 24 Als hij/zij een opdracht geeft, weten we precies wat de bedoeling is.
- 25 Hij/zij laat ons zelf het tempo bepalen waarin we opdrachten maken. (*)
- 26 Hij/zij stimuleert ons de resultaten van ons eigen werk met andere leerlingen door te praten. (*)
- 27 Hij/zij geeft overzichtelijk les.
- 28 Als hij/zij iets uitlegt dan stimuleert hij/zij om actief mee te denken.
- 29 Hij/zij geeft duidelijk aan waarom iets fout is.
- 30 Hij/zij stimuleert dat we verantwoordelijkheid nemen voor ons werk. (*)
- 31 Als hij/zij een opdracht geeft, dan heb je je daar precies aan te houden.
- 32 Als we in de klas aan opdrachten werken, dan houdt hij/zij in de gaten of we wel voldoende aan het werk zijn.
- 33 Bij hem/haar krijg je ruimte om je werk zelf te plannen. (*)

(* = opgenomen in de analyses)

Bijlage 2

Geschatte parameters van het model voor activering

Parameter	Lading ¹ (λ ; ξ)	Meetfout (ϵ ; ζ)	%-verklaarde variantie via model
Docent-geleid → 7	1.010 (.689)	.525	47.5
Docent-geleid → 9	1.174 (.801)	.358	64.2
Docent-geleid → 12	.534 (.364)	.867	13.3
Docent-geleid → 13	.939 (.641)	.589	41.1
Medeleerling-geleid → 3	1.152 (.642)	.588	41.2
Medeleerling-geleid → 11	.980 (.546)	.702	29.8
Medeleerling-geleid → 21	1.264 (.704)	.504	49.6
Medeleerling-geleid → 22	.834 (.464)	.784	21.6
Medeleerling-geleid → 26	1.024 (.570)	.675	32.5
Medeleerling-geleid → 30	.937 (.522)	.728	27.2
Leerling-geleid → 41	.136 (.428)	.817	18.3
Leerling-geleid → 16	1.197 (.451)	.796	20.4
Leerling-geleid → 25	1.460 (.551)	.697	30.3
Leerling-geleid → 33	1.628 (.614)	.623	37.7
Activering → docent-geleid	.487 (.767)	.192	58.9
Activering → medel.-geleid	.449 (.867)	.077	75.1
Activering → leerling-geleid	.195 (.555)	.098	30.8
Fittingsmaten totale model:			
Chi-kwadraat=880.42 ($df=85$; $p=.000$)			
GFI=.954; AGFI=.943			
RMSR=.076			

¹ De gestandaardiseerde schattingen staan tussen haakjes

Discussie

Naar een vollediger systeem van kwaliteitsbeoordeling van scholen

M.P.C. van der Werf en B.P.M Creemers

Rijksuniversiteit Groningen

Samenvatting

Op basis van gegevens van scholen over examenresultaten, zittenblijven en drop-out heeft Dronkers een overkoepelende maat geconstrueerd als index voor de kwaliteit van scholen. De validiteit van deze maat wordt in deze bijdrage ter discussie gesteld. Er wordt uiteengezet dat verschillende outputmaten, zoals numeriek rendement, efficiëntie, effectiviteit – waarin de doelen van het onderwijs worden bereikt in termen van geleerde kennis en vaardigheden – afzonderlijk beoordeeld dienen te worden, daarbij rekening houdend met achtergrond- en instroomkenmerken van leerlingen. Ook de selectiviteit van scholen, de mate waarin scholen in staat zijn een zo'n hoog mogelijk niveau te bereiken voor alle categorieën leerlingen, is een belangrijke kwaliteitsmaat. De kwaliteitsmaat van Dronkers houdt niet voldoende rekening met achtergrondkenmerken van leerlingen en contextkenmerken van scholen. De validiteit en betrouwbaarheid van de kwaliteitsmaat van Dronkers wordt getoetst door zijn resultaten te vergelijken met empirische gegevens van een grootschalig cohortonderzoek in het Voortgezet Onderwijs. Geconcludeerd wordt dat de relatie tussen de kwaliteitsmaat van Dronkers en alternatieve maten gering is. Er is meer onderzoek nodig om vast te stellen welke indicatoren van kwaliteit het meest valide zijn om de overall kwaliteit van scholen te evalueren.

1 Inleiding

Enige tijd geleden is er nogal wat commotie geweest over de gegevens van scholen die in *Trouw* zijn gepubliceerd en de rapportcijfers

van scholen die Dronkers op basis van deze gegevens had berekend. Op zichzelf is publicatie van gegevens, die als indicatie kunnen worden beschouwd voor de kwaliteit van scholen, verdedigbaar. Ouders hebben er recht op te weten of hun keuze voor een bepaalde school verantwoord is (geweest). De keuze voor één overkoepelende kwaliteitsmaat om scholen mee te beoordelen is echter niet verantwoord. Kwaliteit van scholen in termen van opbrengsten omvat meerdere aspecten die weliswaar niet los van elkaar staan, maar die ook niet zonder meer bij elkaar opgeteld kunnen worden tot een eindoordeel. Verder is in de beoordeling van de opbrengsten niet adequaat gecorrigeerd voor achtergrondkenmerken van leerlingen en scholen. Ten slotte is het van belang dat niet alleen gekeken wordt naar de opbrengsten van het onderwijs maar ook naar de wijze waarop deze opbrengsten gerealiseerd worden, dus naar de kwaliteit van het onderwijsproces in de school en in de klas.

Dit artikel handelt over de beoordeling van de opbrengsten van scholen. Op de beoordeling van de kwaliteit van processen wordt ingegaan in de conclusies. Na een korte uiteenzetting over de door Dronkers gehanteerde kwaliteitsmaat worden de bezwaren toegelicht, worden alternatieve maten voorgesteld en worden de uitkomsten van de beoordeling van scholen door Dronkers vergeleken met die van de beoordeling op andere kwaliteitsmaten.

2 Opbrengsten van scholen: rendement, efficiëntie, effectiviteit en selectiviteit

Dronkers heeft op basis van de percentages geslaagden voor het eindexamen, gemiddelde eindexamencijfers voor wiskunde, Nederlands en Engels en de percentages zittenblijvers en drop-outs in de leerjaren 2 tot en met het voorlaatste (leerjaar 5 VWO, leerjaar 4 HAVO, leerjaar 3 VBO en MAVO) een kwaliteitsmaat geconstrueerd. De score op deze maat is een gewogen optelling van de genoemde gegevens waarbij de gewichten zijn bepaald op grond van de factorladingen in een Principale Componenten Analyse. De resulterende score is vervolgens gecorrigeerd voor het percentage allochtone leerlingen door de score te vermin-

deren met het product van dit percentage en de ruwe regressiecoëfficiënt van de regressie van de score op het percentage allochtone leerlingen.

De wijze waarop de outputelementen zijn opgeteld tot een eindscore wordt door ons niet ter discussie gesteld. Bij de elementen die deel uitmaken van de eindscore kunnen echter vraagtekens worden geplaatst. De kwaliteit is in de eerste plaats beoordeeld in termen van *rendement*, het percentage geslaagden in een eindexamenjaar. Rendement van scholen is echter volgens ons meer. In productietermen gedacht is rendement het verschil tussen de opbrengsten en de kosten. In termen van opbrengsten en kosten kunnen de aantallen gediplomeerde leerlingen worden opgevat als opbrengsten en de aantallen leerlingen die uitvallen als de kosten. De aantallen uitvallers kunnen echter alleen worden bepaald als bekend is hoeveel leerlingen er in het begin waren. Rendement van onderwijs is dan dus niets anders dan het verschil tussen het aantal leerlingen dat in een bepaald jaar is ingestroomd in de eerste klas van een bepaald opleidingstype en het aantal leerlingen van die jaargang dat met een diploma de school verlaat. Om rendement goed te bepalen dient bekend te zijn hoeveel procent van de leerlingen die instromen in het eerste leerjaar van het voortgezet onderwijs met een diploma de school verlaat. Dit staat nog los van de efficiëntie. Bij *efficiëntie* gaat het om het tijdsaspect en dus om de vraag hoe lang leerlingen er over doen om het diploma te behalen. In de kwaliteitsmaat van Dronkers is efficiëntie bij benadering opgenomen, door rekening te houden met percentages zittenblijvers en drop-outs. Deze percentages zeggen echter niets over de tijd die het de eindexamenleerlingen heeft gekost om het diploma te behalen.

Rendement en efficiëntie zijn twee manieren om naar de opbrengsten van scholen te kijken. In het licht van de opvatting van onderwijs als productieproces zijn het ook consequente manieren. Maar voor de beoordeling van de kwaliteit van scholen zijn alleen rendement en efficiëntie onvolledig. Bovendien zal het hanteren van deze maten gemakkelijk kunnen leiden tot strategisch gedrag van scholen in die zin dat er sprake zal zijn van strenge selectie bij toelating en wellicht ook tot verlaging van de

prestatienormen, zowel gedurende de periode tot aan het eindexamen als ook bij het beoordelen van de examenprestaties (zie ook Bosker, Brandsma & Van Amelsvoort, 1994). Het huidige systeem van schoolonderzoeken, waarvan uit onderzoek bekend is dat de resultaten daarvan op sommige scholen grote discrepanties vertonen met de resultaten op het Centraal Schriftelijk Eindexamen (Pijl, 1991), sluit deze mogelijkheid zeker niet uit. Dit betekent niet dat eindexamenresultaten niet van belang zijn voor het bepalen van de kwaliteit van een school. Integendeel. Het bijbrengen van kennis en vaardigheden is de belangrijkste doelstelling van onderwijs en dus dient in de kwaliteitsbeoordeling betrokken te worden de mate waarin de leerdoelen zijn bereikt. De kwaliteit van onderwijs moet blijken uit wat leerlingen ervan hebben opgestoken (De Groot, 1993). Ook door de overheid wordt deze definitie van kwaliteit onderkend. In de nota 'De school als lerende organisatie' wordt gesteld dat de beoordeling van de kwaliteit van het onderwijs altijd een waardering inhoudt van de mate waarin de doelen die aan het onderwijs gesteld worden, bereikt worden (Netelenbos, 1995). De mate van doelbereiking wordt *effectiviteit* genoemd. De effectiviteit van scholen kan gemeten worden met behulp van toetsen tijdens en aan het eind van de schoolloopbaan, waarmee de prestaties van de leerlingen onafhankelijk beoordeeld worden, zodat er geen gesjoemel met de normen mogelijk is. Cito-toetsen en ook het Centraal Schriftelijk Eindexamen zijn geschikte middelen om de effectiviteit van het onderwijs te beoordelen, maar eindexamencijfers waarin zowel de resultaten van het Centraal Schriftelijk Eindexamen als die van de schoolonderzoeken zijn opgenomen zijn dat zeker niet. Als Cito-toetsscores of scores op het Centraal Schriftelijk Eindexamen echter worden gehanteerd als enige middel ter beoordeling van scholen wordt opnieuw strategisch gedrag uitgelokt, nl. een strenge selectie zowel bij toelating tot de school als gedurende 'de rit'. Alleen de beste leerlingen, die naar verwachting hoog zullen scoren op de toetsen blijven over. De rest stroomt af naar een lager onderwijstype, en zal daar dan een bijdrage leveren tot hogere scores op de toetsen, ofwel ze blijven zitten en/of raken zo gedemotiveerd dat ze de school verlaten.

Het zal duidelijk zijn dat het gebruik van alleen toetsscores als kwaliteitsmaat onmiddellijk gevolgen zal hebben voor het rendement en de efficiëntie van het onderwijs, zoals eerder gedefinieerd, evenzo als alleen beoordelen op grond van rendement en efficiëntie gevolgen zal hebben voor de effectiviteit van scholen (zie ook Hofstee, 1994). De consequentie hiervan is dat zowel de *effectiviteit* als het *rendement* en de *efficiëntie* in de beoordeling van scholen moeten worden betrokken, hetgeen Dronkers op grove wijze heeft gedaan. Wat Dronkers echter over het hoofd heeft gezien is dat het hanteren van deze drie criteria en het onderbrengen ervan in één overkoepelende kwaliteitsmaat het gevaar met zich mee brengt dat scholen in hoge mate selectief zullen worden bij toelating tot bepaalde typen onderwijs en bij beslissingen over bevordering naar hogere leerjaren, opstroom en afstroom en bij de keuze van het vakkenpakket of in de toekomst van profielen. Dit is strijdig met de taken en verantwoordelijkheden van het onderwijs. Het onderwijs heeft de verantwoordelijkheid om bij alle leerlingen 'het onderste uit de kan te halen'. Dit betekent dat een school zijn uiterste best moet doen om een zo hoog mogelijke opleiding en een zo hoog mogelijk kennisniveau te realiseren voor alle leerlingen. Het begrip 'zo hoog mogelijk' betekent uiteraard niet dat voor alle leerlingen dezelfde opleiding en hetzelfde kennisniveau gerealiseerd moet worden. Er zijn nu eenmaal individuele verschillen in mogelijkheden en motivatie van leerlingen die grenzen stellen aan de mogelijkheden van een school. Met 'zo hoog mogelijk' wordt bedoeld 'gegeven de aanleg, het beginniveau en de motivatie van de leerling'. Verschillen in kwaliteit (zowel m.b.t. rendement, efficiëntie en effectiviteit) tussen scholen komen in deze opvatting tot uitdrukking in de diploma's die leerlingen behalen, de tijd die ze daarover doen en in hun scores op onafhankelijke toetsen, *rekening houdend met de aanleg, het beginniveau en het verwachte diploma* bij intrede van de leerlingen in de school. Een school waarop meer leerlingen met een HAVO-advies een HAVO-diploma en hogere prestaties halen is een betere school dan een school die met qua achtergrondkenmerken vergelijkbare leerlingen lagere opbrengsten heeft. Maar als op de laatstgenoemde school ook

meer leerlingen met een MAVO-advies en relatief lage prestaties op de Cito-eindtoets een HAVO diploma halen alsook hogere cijfers dan op de eerstgenoemde school, dan is deze school tevens minder selectief – of positief uitgedrukt: heeft een beter compenserend vermogen (zie Brandsma & Knuver, 1988). Essentieel is dus dat, ter voorkoming van een hoge mate van selectiviteit door scholen, bij de bepaling van zowel het rendement en de efficiëntie als van de effectiviteit niet alleen gekeken wordt naar de overall resultaten van een school, gecorrigeerd voor instroomkenmerken van leerlingen. Ook dienen de drie typen opbrengsten voor afzonderlijke groepen leerlingen waarvoor verschillende verwachtingen bestaan bij entree op de school afzonderlijk te worden beoordeeld. Dronkers heeft weliswaar gekeken naar de opbrengsten van scholen voor VBO-, MAVO-, HAVO- en VWO-leerlingen, maar heeft hierbij geen rekening gehouden met de verwachtingen die voor de leerlingen bestonden bij hun instroom in het eerste leerjaar. Selectiviteit van scholen dient in de beoordeling van de opbrengsten van scholen te worden verdisconteerd door de verwachtingen (of het advies) te vergelijken met het bereikte eindniveau en door rekening te houden met opstroom en afstroom van leerlingen naar hogere respectievelijk lagere onderwijstypes.

3 Probleemstelling

In het voorgaande zijn de bezwaren uiteengezet tegen de door Dronkers gehanteerde maten voor het beoordelen van de opbrengsten van scholen. De kern van de bezwaren is dat in de door Dronkers gehanteerde maten geen rekening is gehouden met de instroomkenmerken van leerlingen. Het feit dat in de cijfers wel rekening is gehouden met het percentage allochtone leerlingen kan dit probleem niet ondervangen, omdat zo'n percentage geen goede indicator is voor de individuele instroomkenmerken van de leerlingen. Het is alleen een indicator voor de leerlingpopulatie van een school waarop een school zelf, door haar selectiebeleid veel invloed kan uitoefenen. Doordat geen rekening is gehouden met de instroomkenmerken van leerlingen zijn de examenresultaten (percentages geslaagden en

eindexamencijfers) eveneens meer een uitdrukking van de selectiviteit van scholen dan van hun kwaliteit. In het hiernavolgende wordt onderzocht of de kwaliteitsindex die Dronkers heeft gehanteerd samenhangt met kwaliteitsindexen die wel zijn gecorrigeerd voor individuele leerlingkenmerken. Deze alternatieve kwaliteitsindexen volgen uit wat eerder is opgemerkt over de taken van de school: kwaliteit wordt geïndiceerd door het gerealiseerde opleidings- en kennisniveau, gegeven een bepaalde verwachting op basis van aanleg, geadviseerde opleiding en beginniveau. De kwaliteitsindexen zijn afzonderlijk berekend voor VBO-, MAVO-, HAVO- en VWO-leerlingen. In de index is dus ook rekening gehouden met de selectiviteit van de school. De vraagstellingen luiden als volgt:

- 1 *Hoe groot zijn de verschillen tussen scholen op de alternatieve kwaliteitsindexen?*
- 2 *Hoe hangen de alternatieve kwaliteitsindexen onderling samen?*
- 3 *Hoe groot is de samenhang tussen de alternatieve kwaliteitsindexen en de kwaliteitsindex van Dronkers?*

4 Methode

4.1 Steekproef

De vraagstellingen worden beantwoord met behulp van de gegevens van de scholen die meedoen aan het VOCL '93 cohortonderzoek. De gegevens van dit cohort zijn geschikt dan die van cohort '89 omdat de meeste scholen die aan het cohort '89 meewerkten in 1996/1997, het jaar waarop de rapportcijfers van Dronkers betrekking hebben, niet meer in hun oorspronkelijke gedaante bestonden vanwege opheffing of fusies. Bovendien is het instroomjaar van het VOCL '93 cohort voor VBO- en MAVO-leerlingen hetzelfde als dat van de eindexamenleerlingen in 1996/1997. Het VOCL '93 cohort heeft echter als nadeel dat er nog geen gegevens beschikbaar zijn over eindexamencijfers en diploma's. We moeten volstaan met de gegevens tot en met het derde leerjaar.

Aan het VOCL '93 cohort werken ongeveer 350 scholen voor voortgezet onderwijs mee alsmede ongeveer 20.000 leerlingen die in 1993 in het eerste leerjaar zaten. Van ongeveer 200 scholen en 13.000 leerlingen zijn gegevens

tot en met het derde leerjaar beschikbaar. In Tabel 1 staat een overzicht van het aantal scholen per onderwijstype waarvan alternatieve kwaliteitsindexen berekend konden worden en waarvan tevens rapportcijfers van Dronkers beschikbaar waren. Ook de aantallen leerlingen per onderwijstype staan in deze tabel.

Tabel 1

Aantal scholen en leerlingen in de analyses, per schooltype

	VBO	MAVO	HAVO	VWO
Aantal scholen	79	44	83	44
Aantal leerlingen	2041	1048	2189	1146

4.2 Variabelen

Voor de berekening van de kwaliteitsindexen is gebruik gemaakt van de gegevens over het gerealiseerde onderwijstype (VBO, MAVO, HAVO en VWO) en de prestaties, beide in leerjaar drie. In deze indexen zit het efficiëntie-criterium verdisconteerd doordat alleen onvertraagde leerlingen in de analyses zijn opgenomen. Per onderwijstype zijn twee kwaliteitsindexen berekend, nl. één voor de prestaties voor wiskunde en één voor de prestaties voor Nederlands. De prestaties zijn gemeten met toetsen die speciaal door het Cito zijn geconstrueerd voor de evaluatie van de Basisvorming. De toets Nederlands was voor alle leerlingen in de afzonderlijke onderwijstypen hetzelfde, de toets wiskunde bestond uit een A-versie voor MAVO-, HAVO- en VWO-leerlingen en een B-versie voor IBO/VBO-leerlingen. De toets Nederlands bestond uit 40 items en de betrouwbaarheid was .80. De A-versie van de toets wiskunde bestond uit 33 items (betrouwbaarheid .85) en de B-versie van de toets wiskunde bestond uit 33 items met een betrouwbaarheid van .89.

Als instroomkenmerken van de leerlingen zijn beschikbaar: Intelligentie, gemeten met de PSB in leerjaar 1; Sociaal milieu, geïndiceerd door het opleidingsniveau van de ouders (7 niveaus); Etniciteit (allochtoon/ niet-allochtoon); Scores op Cito entreetoetsen in het eerste leerjaar voor rekenen (betrouwbaarheid .84) en Nederlandse taal (betrouwbaarheid .76); en Geadviseerd type voortgezet onderwijs (9 niveaus).

De laatste variabele in de analyses is uiteraard

het rapportcijfer van Dronkers.

4.3 Analyses

Voor de constructie van de kwaliteitsscores is gebruik gemaakt van multilevel analyses met behulp van het programma VARCL. Na specificatie van het lege model zijn eerst de leerlingvariabelen opgenomen. Dit om na te gaan hoe groot de verschillen zijn tussen scholen na correctie voor de instroomkenmerken van de leerlingen. Deze analyses zijn gedaan per onderwijstype afzonderlijk alsmede afzonderlijk voor wiskunde en Nederlands. De uit deze analyses resulterende posterior means voor elke school zijn de alternatieve kwaliteitsscores van scholen voor Nederlands en wiskunde. Deze zijn vervolgens met elkaar gecorreleerd om na te gaan hoe groot de onderlinge samenhang tussen de alternatieve kwaliteitsscores is. Ten slotte zijn ze gecorreleerd met de rapportcijfers van Dronkers.

5 Resultaten

Allereerst worden de verschillen tussen scholen op de kwaliteitsindexen gepresenteerd. Deze zijn afzonderlijk berekend voor Nederlands en wiskunde. Dit om na te gaan of de verschillen tussen scholen voor deze vakken ongeveer even groot zijn en om na te gaan of scholen die hoog scoren op Nederlands dit ook doen voor wiskunde. De verschillen zijn afzonderlijk voor de opleidingstypen geanalyseerd (Tabel 2).

Tabel 2
Percentages onverklaarde variantie in toetsprestaties in leerjaar 3 op leerling- en schoolniveau na controle voor instroomkenmerken, per vak en schooltype

Variantie	Onderwijstype			
	VBO	MAVO	HAVO	VWO
Wiskunde				
leerling	60	56	68	68
school	12	22	14	11
Nederlands				
leerling	75	69	66	74
school	15	23	20	17

De percentages variantie op schoolniveau laten zien dat er aanzienlijke verschillen tussen scholen zijn, ook nadat rekening is gehouden met de instroomkenmerken van de leerlingen. Dit betekent dat scholen verschillen in effectiviteit, zoals afgemeten aan de prestaties van de leerlingen in het derde leerjaar op onafhankelijke toetsen. Op sommige scholen hebben leerlingen na drie jaar meer geleerd dan op andere. Dit geldt voor elk gevolgd onderwijstype, maar meer voor MAVO en HAVO en meer voor Nederlands dan voor wiskunde.

Ter beantwoording van vraagstelling 2 is per school een gemiddelde score voor Nederlands en wiskunde berekend waarin rekening is gehouden met de instroomkenmerken van de leerlingen (de zogenaamde posterior means). Dit zijn de kwaliteitsscores van de scholen voor Nederlands en wiskunde. Vervolgens zijn correlaties berekend tussen deze twee kwaliteitsscores. In Tabel 3 staan de resultaten.

Tabel 3
Correlatie tussen de gemiddelde scores van Nederlands en wiskunde in leerjaar 3, per schooltype

	VBO	MAVO	HAVO	VWO
	.45	.34	.24	.24

Uit deze tabel blijkt dat de correlaties niet erg hoog zijn. Scholen die hoog scoren voor wiskunde scoren niet tevens hoog voor Nederlands. De correlatie is het hoogst voor het VBO. Een verklaring hiervoor kan zijn dat in het VBO vaak door een en dezelfde docent zowel wiskunde als Nederlands wordt gegeven. Docenten in het VBO zijn vaak meervoudig bevoegd op basis van oude LO aktes (Den Boer, Kuyper & Rekvelde, 1996). In de andere schooltypen is meestal sprake van vakdocenten met een tweede- of eerstegraads bevoegdheid.

De derde vraagstelling heeft betrekking op de vergelijking van de scores van de scholen op de alternatieve kwaliteitsmaten en de rapportcijfers van Dronkers. Voor het maken van deze vergelijking zijn de rapportcijfers van Dronkers van de scholen die participeren aan VOCL '93 gecorreleerd met de posterior means voor Nederlands en wiskunde van die scholen. De resultaten staan in Tabel 4.

Tabel 4

Correlatie tussen de gemiddelde scores van Nederlands en wiskunde in leerjaar 3 en rapportcijfers Dronkers, per schooltype

	Rapportcijfer			
	VBO	MAVO	HAVO	VWO
Nederlands	-.13	.10	.06	-.04
Wiskunde	-.08	.24	.41	.05

Uit de correlaties blijkt dat de samenhang tussen de rapportcijfers van Dronkers en de kwaliteitsscores op basis van toetsprestaties in het derde leerjaar over het algemeen zeer laag en soms zelfs negatief is. De samenhang met de rapportcijfers is het hoogst voor de wiskunde-prestaties in HAVO-klassen, maar ook deze correlatie is veel lager dan men zou mogen verwachten. In het algemeen mag geconcludeerd worden dat de kwaliteitsbeoordeling op basis van de overall kwaliteitsscore niet samenhangt met andere kwaliteitsindexen die aangeven wat onvertraagde leerlingen in een bepaald onderwijstype aan kennis van bepaalde vakken hebben opgedaan aan het eind van het derde leerjaar, rekening houdend met hun beginniveau, sociale en etnische achtergrond en advies. In de probleemstelling is geopperd dat dit te maken heeft met het feit dat Dronkers niet adequaat heeft gecorrigeerd voor instroomkenmerken van de leerlingen. Deze verklaring wordt sterker indien de rapportcijfers van Dronkers beter zouden correleren met de ongecorrigeerde scores op de alternatieve kwaliteitsmaten, dat wil zeggen de gemiddelde toetsprestaties waarin geen rekening is gehouden met de instroomkenmerken van de leerlingen. Analyse van de correlaties met de ongecorrigeerde gemiddelden (niet in tabel) laat echter zien dat deze in het algemeen weliswaar iets hoger zijn, maar nog steeds niet erg hoog. De hoogste correlatie is nu .45, wederom voor wiskunde in HAVO.

In Tabel 5 staan de correlaties tussen de ongecorrigeerde gemiddelde eindexamencijfers, zoals in *Trouw* gepubliceerd en de ongecorrigeerde alternatieve kwaliteitsscores. Voor de volledigheid zijn in deze tabel ook de correlaties opgenomen tussen de rapportcijfers van Dronkers en de gemiddelde eindexamencijfers.

Tabel 5

Correlatie tussen de gemiddelde scores van Nederlands en wiskunde in leerjaar 3 en gemiddelde eindexamencijfers van Dronkers en tussen de rapportcijfers van Dronkers en gemiddelde eindexamencijfers, per schooltype

	Gemiddelde eindexamencijfer			
	VBO	MAVO	HAVO	VWO
Nederlands	.30	.23	-.08	.08
Wiskunde	.14	.06	.05	.30
Rapportcijfer	-.16	.79	.07	.34

De correlaties tussen de ongecorrigeerde gemiddelden voor Nederlands en wiskunde en de ongecorrigeerde eindexamencijfers zoals gepubliceerd in *Trouw* zijn laag tot zeer laag. Het feit dat ook de correlaties tussen eindexamencijfers en de rapportcijfers van Dronkers laag zijn (behalve in MAVO) en in VBO zelfs negatief, terwijl de rapportcijfers mede gebaseerd zijn op de eindexamencijfers is zeer vreemd. Het doet vermoeden dat de constructie van de rapportcijfers op basis van een gewogen optelling van eindexamencijfers, rendement en zittenblijven/uitval geen valide manier is om een overall index voor kwaliteit te construeren, althans niet zoals door Dronkers is verricht. Overigens blijkt dit ook reeds uit de resultaten van de factoranalyses van Dronkers zelf.

6 Conclusies

Met de analyses zoals hierboven gepresenteerd is aangetoond dat het beoordelen van scholen op basis van een overkoepelende kwaliteitsindex een riskante aangelegenheid is, zeker wanneer daarin onvoldoende rekening wordt gehouden met de instroomkenmerken van de leerlingen. De kwaliteitsindex van Dronkers hangt nauwelijks samen met andere indexen van kwaliteit, zowel indien deze worden gecorrigeerd voor instroomkenmerken van leerlingen als indien deze ongecorrigeerd worden gehanteerd. De overkoepelende index hangt zelfs nauwelijks samen met een afzonderlijke index die van de overkoepelende index deel uitmaakt. Uit deze resultaten kan geconcludeerd worden dat de validiteit van de beoordeling van scholen door Dronkers zeer twijfelachtig is. Dit ligt niet alleen aan het feit dat er slechts grove maten beschikbaar waren, maar

ook aan de wijze waarop de overkoepelende index is geconstrueerd.

Dit leidt tot de vraag of het eigenlijk wel mogelijk is een overkoepelende index te construeren. Voorlopig denken we dat het antwoord hierop negatief moet zijn. In de eerste plaats is duidelijk gemaakt dat het bij de kwaliteit van scholen gaat om meerdere aspecten, zoals rendement, efficiëntie, effectiviteit en selectiviteit, die kunnen worden beschouwd als verschillende dimensies van kwaliteit. Zolang onduidelijk is hoe die dimensies zich tot elkaar verhouden, is een optelling van de scores van scholen op deze dimensies – gewogen dan wel ongewogen – onmogelijk. Dit geldt ook voor het optellen en middelen van eindexamenvakken tot gemiddelde eindexamenscores. In de gepresenteerde analyses hebben we laten zien dat scholen weinig consistent zijn in hun output over vakken heen. Een school die hoog scoort voor wiskunde kan tegelijkertijd laag scoren voor Nederlands. Een kwaliteitsindex kan dus niet volstaan met scores voor slechts enkele vakken en bovendien zullen de vakken gewogen moeten worden bij het berekenen van een gemiddelde. In plaats van een empirische weging op grond van factorladingen zoals Dronkers heeft gedaan zal hierbij gekeken moeten worden naar de zwaarte van de vakken en naar de voorspellende waarde van de vakken voor verdere opleiding of studie. Op dit moment is hierover weinig bekend.

In de tweede plaats is ook bekend dat scholen weinig stabiel zijn in hun opbrengsten over schooljaren heen. Een school die het ene jaar hoog scoort op een kwaliteitscriterium hoeft een jaar later niet per definitie ook hoog te scoren. Criteria die meer stabiel zijn over de tijd zouden misschien in de algemene index zwaarder moeten wegen dan minder stabiele criteria. Een gevaar hiervan is echter dat wellicht de minder essentiële criteria het meest stabiel blijken te zijn.

Afgezien van bovengenoemde problemen is het echter voor een eerlijke vergelijking tussen scholen ook van belang dat correctiefactoren worden ingebouwd waarvan bekend is dat die in belangrijke mate verschillen in output verklaren maar die los staan van de invloed van de school. Aan de andere kant moet niet gecorrigeerd worden voor factoren die duidelijk geen invloed hebben. In dit verband wijzen we erop

dat in onze analyses de etniciteit van leerlingen er niet toe doet voor de verklaring van verschillen in prestaties tussen leerlingen in het VWO, nadat reeds rekening is gehouden met SES, intelligentie en beginniveau. Bij de berekening van kwaliteitsscores hoeft voor VWO scholen dan ook geen rekening te worden gehouden met de etniciteit van de leerlingen. Correctie voor het percentage allochtonen op schoolniveau is alleen nodig als dit, na correctie voor etniciteit en andere instroomkenmerken op leerlingniveau nog steeds een zelfstandige bijdrage levert aan de verschillen in output tussen scholen. Het zonder meer corrigeren voor het percentage allochtonen op schoolniveau zoals Dronkers dat heeft gedaan zou wel eens kunnen hebben geleid tot een onterecht hoge waardering van VWO-afdelingen in een scholengemeenschap met veel allochtone leerlingen en een onterecht lage waardering van scholen met weinig van deze leerlingen. Voor de goede orde: ook wij hebben niet de pretentie in onze analyses afdoende te hebben gecorrigeerd voor instroom- en achtergrondkenmerken van de leerlingen. Het bewijs daarvoor is dat er nog zeer veel variantie op leerling- en schoolniveau onverklaard is. We vermoeden dat veel van de variantie op leerlingniveau verklaard kan worden door de motivatie en inzet van de leerlingen. Nadere analyses zullen dit moeten uitwijzen. Daarnaast zullen er vermoedelijk op schoolniveau en in de context van de school factoren zijn die niets te maken hebben met de kwaliteit van het onderwijsleerproces van de school maar die wel deels de verschillen tussen scholen kunnen verklaren. Als deze factoren er zijn zal daarmee in de kwaliteitsindex ook rekening moeten worden gehouden. Te denken valt bijvoorbeeld aan de buurt of regio waarin de school staat, opleiding en leeftijd van leraren en dergelijke.

Samenvattend kan geconcludeerd worden dat het thans niet mogelijk is om scholen te beoordelen op basis van een kwaliteitsindex. Dat neemt niet weg dat wel resultaten van scholen op beschikbare opbrengstmaten kunnen worden gepubliceerd, zoals bijvoorbeeld percentages geslaagden, eindexamencijfers en dergelijke, overeenkomstig de bestaande plannen tot het uitgeven van 'kwaliteitskaarten' door de inspectie. Voorlopig zullen daarop alleen nog cijfers omtrent rendement en exa-

menresultaten komen te staan en vanaf 1999 ook doorstroom naar en resultaten in het vervolgonderwijs. Op zich is daar niets op tegen, onder de conditie dat het bij eindexamenresultaten gaat om de cijfers op het Centraal Schriftelijk Eindexamen. Het moet echter duidelijk zijn dat hiermee niet een soort keurmerk aan scholen kan worden gegeven op grond waarvan De Kwaliteit van de school is vastgesteld. De reactie van Dronkers (1997) op de kwaliteitskaarten onderschrijven we dus niet waar hij stelt dat de gegevens daarvan onvoldoende zijn voor ouders omdat een samenvattend oordeel ontbreekt. Ouders zullen volgens Dronkers door de bomen het bos niet meer zien. Het verdient echter de voorkeur dat ouders op basis van meerdere gegevens zelf uitmaken waar 'het bos' is dan dat hun een 'bos' wordt beloofd op basis van een invalide kwaliteitsmaat. Dit betekent niet dat een overall kwaliteitsmaat per definitie onmogelijk is. Om dit te kunnen bepalen is nader onderzoek nodig in de vorm van analyses op bestaand materiaal. De gegevens die de inspectie in haar bezit heeft zijn hiervoor eigenlijk te beperkt gebleken. Momenteel zijn er wel grote databestanden beschikbaar die van twee cohorten leerlingen op ruim 400 scholen voor voortgezet onderwijs gegevens bevatten over hun schoolloopbaan. Op basis van deze gegevens kunnen de opbrengsten van scholen worden berekend voor diverse criteria, zoals behaalde diploma's, bereikt kennisniveau in leerjaar 3 en op het moment van eindexamen (eindexamencijfers), opstroom, afstroom en zittenblijven en kan worden gecorrigeerd voor instroom en achtergrondkenmerken van leerlingen en contextkenmerken van de school. Een mooi opbrengstcriterium, waarin zowel de dimensie rendement als de dimensie efficiëntie is verdisconteerd, is de score op de leerjarenladder (Bosker, Van der Velden & Hofman, 1985). De resultaten van scholen op deze afzonderlijke opbrengstcriteria kunnen onderling worden vergeleken evenals de resultaten over twee cohorten heen als indicatie voor de stabiliteit. Op basis van deze analyses kunnen diverse kwaliteitsindexen worden geconstrueerd en met elkaar vergeleken. De indexen die het meest met elkaar samenhangen en het meest stabiel zijn over de tijd kunnen wellicht uiteindelijk deel gaan uitmaken van een overkoepelende kwaliteitsbeoordeling van scholen.

Zoals in de inleiding echter reeds is opgemerkt is het beoordelen van de kwaliteit van scholen in alleen termen van opbrengsten nogal beperkt. Het gaat er ook om hoe die opbrengsten worden bereikt. In haar reactie op het *Trouw*-onderzoek geeft de Staatssecretaris te kennen dat kwaliteitsoordelen zo veel mogelijk over de volle breedte dienen te worden bepaald (Netelenbos, 1997). Zeker indien kwaliteitsoordelen voor ouders bedoeld zijn, is eenzijdige nadruk op output van scholen onvoldoende. Ouders zijn niet alleen geïnteresseerd in de resultaten van scholen maar ook of hun kinderen het op school naar hun zin (zullen) hebben (Mertens, Van Dommelen & De Rijcke, 1994). Ook leerlingen zelf hebben daar baat bij. Uit onderzoek naar de ontwikkeling van motivatie van leerlingen in de eerste drie jaar van het voortgezet onderwijs blijkt dat er een positief verband is tussen de mate waarin leerlingen het leuk vinden op school en waardering hebben voor de leraren enerzijds en motivatie en prestaties anderzijds (Kuyper & Swint, 1996). Voor de beoordeling van de kwaliteit van scholen dient daarom ook gekeken te worden naar de kwaliteit van de schoolorganisatie, het schoolklimaat en de kwaliteit van het instructieproces in de klas. Het ligt voor de hand dat op scholen die hoger scoren op deze proceskenmerken leerlingen met meer plezier naar school gaan, een positievere houding tegenover leren hebben, meer gemotiveerd zijn en dergelijke en dat leraren met meer enthousiasme lesgeven, meer hun vak bijhouden, de leertijd effectiever gebruiken en dus minder huiswerk hoeven te geven. Uitgezocht zal moeten worden of scholen op deze kenmerken verschillen en of dit samenhangt met hun opbrengsten. Met nadruk merken we op, aansluitend bij Hofstee (1992), dat het hierbij alleen dient te gaan om procesindicatoren die voorspellend zijn voor de output en waarop scholen zelf invloed kunnen uitoefenen. Scholen zijn zelf immers het meest verantwoordelijk voor de kwaliteit van het onderwijsproces en de output daarvan (Mertens, 1995). Tot dusver heeft empirisch onderzoek nog maar weinig procesvariabelen opgeleverd die substantieel en consistent samenhangen met de output van scholen. De meeste van deze variabelen hebben te maken met tijd en gelegenheid tot leren (Cremers, 1995). Vooralsnog pleit dit voor een

beperkte set van indicatoren voor de beoordeling van de kwaliteit van het onderwijsproces op scholen. Wellicht leidt dit ook tot opruiming van een groot aantal procesindicatoren die deel uitmaken van de thans in zwang zijnde instrumenten voor zelfevaluatie door scholen (Van de Grift, 1996). Veel van de variabelen in deze instrumenten hebben met de kwaliteit van het onderwijs in termen van output niets van doen en zijn derhalve niet van belang in het kader van de kwaliteitszorg. Verder onderzoek zal moeten uitwijzen welke procesvariabelen wel relevant zijn om in de toekomst in een beoordelingssysteem op te nemen.

Literatuur

- Boer, P. den, Kuiper, H., & Rekveld, I. (1996). *Vakkencombinaties in de Basisvorming*. Groningen: GION.
- Bosker, R.J., Brandsma, J.T., & Amelsvoort, H.W.C.H., van. (1994). Kwaliteitsverbetering door output-financiering? In B.P.M. Creemers (red.), *Deregulering en de kwaliteit van het onderwijs* (pp. 149-161). Groningen: RION monografieën 17.
- Bosker, R.J., Velden, R.K.W. van der, & Hofman, W.H.A. (1985). *Een generatie geselecteerd. Deel 1: de schoolloopbanen*. Groningen: RION.
- Brandsma, H.P., & Knuver, J.W.M. (1988). Organisatorische verschillen tussen basisscholen en hun effect op leerlingprestaties. *Tijdschrift voor Onderwijs Research*, 13, 201-222.
- Creemers, B. (1995). Process indicators on school functioning and the generalisability of school factor models across countries. In *Measuring the quality of schools* (pp. 105-118). Parijs: OECD.
- Dronkers, J. (1997). Netelenbos komt met volstrekt onvoldoende kwaliteitskaart van scholen. *Trouw*, 31 december 1997, p.11.
- Grift, W. van de (1996). Autonomie, deregulering en kwaliteit. *Basisschool Management*, mei 1996, 12-17.
- Groot, A.D. de (1993). *Denken over onderwijs. Analyses en kritieken van A.D. de Groot*. Den Haag: SVO.
- Hofstee, W.K.B. (1992). Onderwijsindicatoren: seizoensopruiming. In B.P.M. Creemers & W.Th.J.G. Hoebe (red.), *Indicatoren van onderwijs-effectiviteit* (pp. 77-81). Groningen: ICO.
- Hofstee, W.K.B. (1994). Hoe bestuurbaar is onderwijs? In B.P.M. Creemers (red.), *Deregulering en de kwaliteit van het onderwijs* (pp. 59-67). Groningen: RION monografieën 17.
- Kuiper, H., & Swint, F.E. (1996). *Microscopisch schoolloopbaanonderzoek. De eerste drie jaren in het voortgezet onderwijs*. Groningen: GION.
- Mertens, F. (1995). De school van de toekomst zorgt zelf voor kwaliteit. In P. Appelfhof & F. Mertens, *Eenheid strategisch beleid* (pp.27-32). Zoetermeer: Ministerie van OC&W.
- Mertens, F., Dommelen, J. van, & Rijke, F.J.M. (1994). Onderwijskwaliteit als recht en als opdracht. In B.P.M. Creemers (red.), *Deregulering en de kwaliteit van het onderwijs* (pp. 1-13). Groningen: RION monografieën 17.
- Netelenbos, T. (1995). *De school als lerende organisatie. Kwaliteitsbeleid op scholen voor primair en voortgezet onderwijs*. Zoetermeer: Ministerie van OC&W.
- Netelenbos, T. (1997). Kwaliteit. *Uitleg*, 26 (5), p.97.
- Pijl, Y.J. (1991). *Discrepanties tussen het schoolonderzoek en het Centraal Schriftelijk Eindexamen*. Groningen: RION.

Auteurs

M.P.C. van der Werf is senior onderzoeker bij het GION en coördinator van de VOCL cohort studies in het voortgezet onderwijs.

B.P.M. Creemers is hoogleraar bij de vakgroep Onderwijskunde van de Rijksuniversiteit Groningen en directeur van het GION.

Adres: GION, Rijksuniversiteit Groningen. postbus 1286, 9701 BG Groningen.
E-mail m.p.c. van.der.werf@ppsw.rug.nl

Abstract

Towards a more complete system of quality assessment of schools

M.P.C. van der Werf & B.P.M. Creemers. Pedagogische Studiën, 1999, 76, 51-60.

On basis of school data about examination results, retention rate and drop-out rate one output measure was constructed by Dronkers, indicating the quality of schools. The validity and reliability of this quality indicator is questionable. It is made clear

that different output measures, such as numerical profit, efficiency, effectiveness – the degree of goal attainment in terms of learned knowledge and skills – have to be evaluated separately, taking into account student background and intake characteristics. Also the selectivity of schools, the degree to which schools are able to reach the highest as possible educational level for all categories of students, given their intellectual abilities and prior knowledge, is an important quality indicator. The quality indicator of Dronkers does not take into account sufficiently the student background characteristics and school context variables. The validity and reliability of the quality indicator of Dronkers is tested by comparing his results with empirical data of a large scale cohort study in secondary education. It is concluded that the relationship between the quality indicator of Dronkers and alternative indicators is very low. More research is needed to establish which indicators are the most valid and reliable to evaluate the overall quality of schools.

Openbare prestatiegegevens van scholen; nuttigheid en validiteit

R.J. Bosker en J. Scheerens

Universiteit Twente

Samenvatting

Het publiceren van de opbrengsten van scholen voor voortgezet onderwijs wordt belicht vanuit drie toepassingscontexten en een conceptuele en technische beschouwing over wat toegevoegde waarde nu precies is.

De toepassingscontexten die besproken worden zijn: ondersteuning van schoolkeuze-processen, het afleggen van rekenschap (accountability) door scholen en schoolverbetering. Vanwege de relatief beperkte omvang van de verschillen tussen de scholen, onvolledige kennis over wat nu de kern is van de onderwijs-productiefunctie, en de beperkte besturingsmogelijkheden van hogere besturingsniveaus in het onderwijs, wordt beargumenteerd dat het publiceren van opbrengstgegevens van scholen slechts van beperkte waarde is voor elk der toepassingscontexten.

In de discussie over het begrip toegevoegde waarde en zijn meting, wordt betoogd dat de definitie in elk geval als uitgangspunt dient te hebben, de mate waarin individuele leerlingen boven of beneden de (statistische) verwachting presteren. Ouders hebben bovendien een ander soort toegevoegde waarde op het oog als ze een school kiezen dan andere gebruikers van de opbrengstinformatie: ouders willen weten wat de school voor hun kan betekenen, anderen willen een zuivere maat voor de inspanning die de school als zodanig geleverd heeft. Daar komt bij dat ouders met name geïnteresseerd zijn in differentiële toegevoegde waarde: de specifieke betekenis van de school voor het eigen kind gegeven diens eigenschappen. Het artikel besluit met een uiteenzetting over de mogelijke negatieve gevolgen van het openbaar maken van schoolresultaten, maar de slotsom is dat de

praktijk als zodanig vermoedelijk opbrengstverhogend zal werken.

1 Inleiding

De tweede *Trouw*-publicatie (in juni 1998) over de opbrengsten van scholen voor voortgezet onderwijs lokte aanmerkelijk minder discussie uit dan de eerste, die in het najaar van 1997 verscheen. Tot nu toe waren de twee belangrijkste elementen in de discussie of het wel verstandig was opbrengstgegevens van scholen openbaar te maken en of de door Dronkers gehanteerde techniek voor het becijferen van scholen wel adequaat was (Veenstra, Dijkstra, Peschar & Snijders, 1998; Van der Werf & Creemers, 1999; Van Putten & Van der Kamp, 1998; Dronkers, 1998). Langzamerhand verstomt de discussie over het eerste element en ten aanzien van het tweede element is een Babylonische spraakverwarring rond het begrip 'toegevoegde waarde' opgetreden. In de onderhavige bijdrage zullen we onze standpunten ter zake uiteenzetten.

Hoewel de *Trouw*-publicatie en ook de kwaliteitskaarten van scholen die door de inspectie van het onderwijs uitgegeven worden primair bedoeld zijn om schoolkeuzeprocessen te ondersteunen, zijn er nog twee andere beweegredenen voor het openbaar maken van prestatiegegevens van scholen, namelijk enerzijds zelfevaluatie en schoolverbetering en anderzijds externe controle op de kwaliteit van het functioneren van scholen, een toepassingscontext die wordt aangeduid met de term 'accountability'. We zullen eerst op de diverse toepassingscontexten ingaan, vervolgens de meer technische aspecten van het bepalen van de door scholen gerealiseerde opbrengsten behandelen en ten slotte beide constituerende elementen van deze bijdrage met elkaar verbinden. Slechts zijdelings, en alleen daar waar het de betoogtrant niet stoort, zullen we op de technische onvolkomenheden van de werkwijze van Dronkers ingaan.

2 De toepassingscontext

2.1 Hoe ouders kiezen en wat er te kiezen valt

Uit onderzoek van Derriks, De Kat en Deckers (1995) blijkt dat in Nederland drie motieven een belangrijke rol spelen bij de schoolkeuze: kwaliteit, afstand en denominatie. Prestatiegegevens worden overigens niet genoemd als kwaliteitsindices, maar wel, bijvoorbeeld, de studiebegeleiding, de grootte van de school en het schoolregime. Dit beeld komt overeen met de resultaten van een casestudy-onderzoek van Lam (1998), die met 20 ouders interviews voerde met als vraag welke informatie ze graag wilden hebben ter ondersteuning van de schoolkeuze. De resultaten sporen met die van Derriks c.s.: procesindicatoren als pedagogisch klimaat, de sfeer op de school, didactische werkvormen, klassengrootte, aansluiting op het vervolgonderwijs e.d. worden belangrijker gevonden dan opbrengstgegevens.

Als ouders zich zouden laten leiden door opbrengstgegevens van scholen bij hun schoolkeuze, hebben ze dan inderdaad iets te kiezen? De rapportcijfers die Dronkers in de eerste *Trouw*-publicatie uitdeelde suggereren een enorme kwaliteitsvariatie, maar dat is natuur-

lijk een artefact van de door hem gehanteerde methodiek, waarbij per definitie altijd enen en tien worden uitgedeeld¹. Hebben ouders veel te kiezen? In Tabel 1 geven we een overzicht van het aantal scholen per schoolsoort dat zich binnen een gemeentegrens bevindt.

De meerderheid der gemeenten biedt per schoolsoort aan ouders geen alternatief: er is maar één school. Ruime keuzemogelijkheden (3 of meer scholen), althans als we voorbij gaan aan keuzes die zich langs denominatieve lijnen voltrekken, zijn slechts in ca. 20% van de gemeenten aanwezig. Hebben ouders dan ook nog iets te kiezen, in de zin dat de scholen die in aanmerking komen verschillen in hun opbrengsten? Tabel 2 bevat daartoe gegevens van enkele kengetallen die de inspectie op de zogenaamde kwaliteitskaarten van scholen vermeldt (Inspectie van het onderwijs, 1998).

Met de interkwartiele range geven we het verschil weer tussen de scholen die tot de onderste 25% slechtst presterende behoren en de scholen die tot de groep der 25% best presterende scholen behoren. Ouders en hun kinderen die binnen het agrarisch onderwijs een school moeten kiezen hebben feitelijk geen keuzemogelijkheden. De meeste variatie in gemiddelde examencijfers vinden we in het

Tabel 1
Aantal scholen per schoolsoort per gemeente

	VBO	MAVO	HAVO	VWO
1 school	64%	56%	56%	54%
2 scholen	19%	22%	20%	20%
3 scholen	9%	9%	10%	10%
4 scholen	3%	4%	6%	6%
5 of meer scholen	5%	9%	8%	8%
totaal aantal gemeenten	246	340	216	214

(bron: Bosker, 1998)

Tabel 2
Variatie in opbrengsten tussen scholen binnen gemeenten; geschatte interkwartiele range

	VBO	MAVO	HAVO	VWO
Onvertraagd in de bovenbouw	11%	6%	14%	16%
Gemiddeld cijfer centraal schriftelijk examen	n.v.t.	0.22	0.17	0.29
% vakken op C-niveau				
Techniek	24%	n.v.t.	n.v.t.	n.v.t.
Zorg & welzijn	28%	n.v.t.	n.v.t.	n.v.t.
Economie	28%	n.v.t.	n.v.t.	n.v.t.
Landbouw	0%	n.v.t.	n.v.t.	n.v.t.

(bron: Bosker, 1998)

VWO met verschillen van 6.36 tegenover 6.65, terwijl het bij het HAVO gaat om 6.25 tegenover 6.42. Dat zijn uiterst smalle marges. In het AVO vinden we de ruimste marges qua onvertraagde doorstroom wederom in het VWO met 50% tegenover 66%, met als betekenis dat op de 25% laagst scorende VWO-scholen slechts 1 op de 2 leerlingen zonder zittenblijven vanaf leerjaar 3 het VWO-diploma behaalt, terwijl dit op de 25% hoogst scorende scholen 2 op de 3 leerlingen betreft. Dat zijn relevante verschillen die bovendien ook in het HAVO gevonden worden, maar in het MAVO gaat om een zeer geringe bandbreedte met 83 tegenover 89%. Grote variaties doen zich voor in het VBO waar het de keuze van vakken op C-niveau betreft. Bij de sector 'zorg & welzijn', bijvoorbeeld, gaat het om een geschat verschil van 38 tegenover 66%.

Dronkers roert zelf het probleem van de normering aan: wanneer presteert een school nu onvoldoende en wanneer voldoende? We zouden als standaard kunnen stellen dat een school onvoldoende opbrengsten heeft als de rendementsgegevens onder de 50% liggen (meer dan 1 op de 2 leerlingen blijft in de bovenbouw in één der leerjaren zitten en voor het VBO bovendien: minder dan 1 op de 2 leerlingen slaagt met minimaal 3 vakken op C-niveau), of als de gemiddelde examencijfers onder de 6 liggen. We zouden ook een standaard kunnen formuleren waarboven scholen ruim voldoende presteren. Voor de rendementsgegevens zouden we die op 90% kunnen stellen, en voor de examencijfers op een 7 of hoger gemiddeld. Wat valt er dan te kiezen c.q. te vermijden (zie Tabel 3)?

Met de, overigens vrij willekeurig geformuleerde, standaarden in de hand kunnen ouders en hun kinderen bij hun keuze in elk geval binnen het AVO de minder presterende scholen vermijden. In het VBO is dat welhaast onmogelijk, althans als we kijken naar het criterium dat minstens de helft der leerlingen het eind-examen aflegt met minimaal 3 vakken op het C-niveau. Op zoek naar excellentie zullen velen echter ontdekken dat er in hun omgeving geen school van hun gading te vinden is.

We zien dat het beeld nogal divers is, en in sommige gevallen is er dus iets te kiezen, en in andere gevallen leidt het publiceren van de opbrengstgegevens van scholen tot de geruststelling dat elke keuze (binnen de mogelijkheden, en alleen dan als er iets te kiezen is) even goed of even slecht is. In sommige gevallen hebben ouders en hun kinderen alleen een serieus alternatief als ze bereid zijn ook buiten de eigen gemeentegrenzen te zoeken. Dit laatste speelt inderdaad een rol in de Verenigde Staten van Amerika en in Engeland, waar gezinnen die om economische redenen moeten verhuizen, zich soms bij het zoeken van een huis laten leiden door het gegeven of er in de buurt een school gevestigd is die er in de 'rankings' goed voor staat (Lam, 1998).

2.2 Het 'accountability' perspectief

Het openbaar maken van de resultaten van scholen wordt ook verdedigd vanuit 'accountability' overwegingen, namelijk dat scholen rekenschap dienen af te leggen over met welk succes de gemeenschapsgelden zijn besteed. Dat rekenschap afleggen geschiedt dan zowel

Tabel 3
'Extreem' scorende scholen

	rendement bovenbouw		examencijfer	
	<50%	>90%	<6	>7
VBO	0.2%	30.0%	n.v.t.	n.v.t.
MAVO	0.8%	28.1%	5.2%	3.4%
HAVO	33.7%	0.4%	7.5%	0.2%
VWO	27.0%	0.2%	4.4%	2.4%
3 of meer vakken op C-niveau				
Techniek	84.5%	1.3%	n.v.t.	n.v.t.
Zorg & welzijn	48.3%	5.5%	n.v.t.	n.v.t.
Economie	18.8%	24.8%	n.v.t.	n.v.t.
Landbouw	71.9%	3.1%	n.v.t.	n.v.t.

(bron: Bosker, 1998)

naar hogere bestuursorganen, als naar de zittende ouders en de lokale gemeenschap.

Veel staten in de VS hebben hun 'school report cards' ingevoerd als onderdeel van beleid gericht op accountability van schooldistricten. Het afleggen van rekenschap door scholen past bij een andere besturingsfilosofie van het onderwijs, waarin de overheid meer afstand bewaart en scholen autonomer kunnen opereren. De overheid kan zich vervolgens beperken tot het stellen van de doelen ('management by objectives'). Vervolgens kan ze nagaan of de scholen die doelen inderdaad bereiken en in voorkomende gevallen sancties opleggen. In het verleden is in Nederland met die gedachte gespeeld toen de overheid overwoog een vorm van outputfinanciering in het voortgezet onderwijs in te voeren (Van Amelsvoort, Bosker & Brandsma, 1993). Het publiceren van opbrengstgegevens van scholen is een alternatief voor outputfinanciering. De beloning of sanctie wordt nu door de marktwerking geregeld: goed presterende scholen trekken meer leerlingen aan, slecht presterende scholen zullen uiteindelijk door een gebrek aan aanmeldingen de tent moeten sluiten.

De vraag in hoeverre besturing vanuit hogere niveaus als een toepassingscontext van belang is, hangt samen met de autonomie van scholen. In Tabel 4, ontleend aan een onderzoek dat is uitgevoerd in het kader van het OECD onderwijsindicatorenproject (Scheerens, 1998) wordt aangegeven welk percentage van beslissingen in een bepaald domein, door de school zelf, of door hogere bestuursorganen wordt genomen.

Tabel 4
Autonome besluitvorming door scholen in het voortgezet onderwijs in vijf landen (als proportie beslissingen genomen door de school versus door hogere bestuursorganen)

	autonoom binnen kader	volledig autonoom
Denemarken	31	10
Duitsland	37	11
Engeland	60	40
Frankrijk	29	9
Nederland	75	39

(bron: Scheerens, 1998)

In deze tabel zijn de gegevens van vijf landen opgenomen: Denemarken en Duitsland, waar geen enkele vorm van openbaar maken van schoolresultaten bestaat, Engeland en Frankrijk, landen die een lange traditie op dit gebied hebben en ten slotte Nederland, waar dit sinds vorig jaar gebeurt. Er is een onderscheid gemaakt tussen beslissingen die scholen kunnen nemen binnen door hogere bestuursorganen vastgestelde kaders en beslissingen die scholen volledig autonoom kunnen nemen. In vergelijking met andere landen zijn Nederlandse scholen voor voortgezet onderwijs zeer autonoom. Opmerkelijk is overigens ook dat in Denemarken en Duitsland, waar geen publicatie van opbrengstgegevens plaats vindt, scholen minder autonoom zijn dan in Engeland en Nederland, waar dat wel gebeurt. Merkwaardig is de positie van Frankrijk, waar scholen nauwelijks autonoom beslissingen kunnen nemen, maar waar desalniettemin over de opbrengsten van scholen ruim gepubliceerd wordt.

Ervaringen in landen waarin het publiceren van opbrengstgegevens duidelijk vanuit een accountability-perspectief plaatsvindt, wijzen overigens op een nogal beperkte daadwerkelijke benutting. Zo komen Cibulka en Derlin (1995, pp. 493-503) tot de conclusie dat "SPR (school performance reporting; rjb en js) is not considered very important by policy-makers or the general public", en is nog lang niet aangehouden dat "SPR can become a potent, effective policy lever". Wel zien zij het als één van de krachtigste symbolen waar de overheid mee voor de dag kan komen om aan het publiek duidelijk te maken dat de beleidsmakers erg hun best doen om de kwaliteit van het onderwijs te verbeteren.

2.3 De marges voor schoolverbetering

De derde toepassingscontext ligt in het gebruik van de gegevens door scholen zelf, in het kader van zelfevaluatie en schoolverbetering. Of dit een realistische toepassingscontext is, hangt af van de mate van 'maakbaarheid' van de opbrengsten door de scholen. Een eerste inzicht in dit vraagstuk kunnen we ontleenen aan onderzoek van Luyten (1994), die examengegevens van een aantal jaren analyseerde om de potentiële bijdrage van scholen in dezen af te schatten (zie Tabel 5).

Tabel 5
De compositie van variantie in examenprestaties

	variantiebronnen	percentage
leerling	85	
school	15	
- per vak		40
- per jaar		8
- per vak per jaar		27
- over vakken en jaren heen		25

(bron: Luyten, 1994)

Prestatieverschillen tussen leerlingen blijken voor 15% te herleiden te zijn op de school die ze bezoeken. Nadere analyse leert echter dat de stabiele (over de jaren heen) en consistente (over de examenvakken heen) schoolbijdrage echter slechts een kwart van die 15% uitmaakt. Er treden enorme variaties tussen de vakken op (een school die het bijvoorbeeld goed doet voor de exacte vakken hoeft het niet ook goed te doen voor de moderne vreemde talen), per jaar zijn de examenresultaten weer anders, en bovendien wisselen de resultaten per jaar per vak. De marges voor algehele schoolverbetering zijn dus beperkt en het lijkt belangrijker

dat scholen elk jaar een goed inzicht krijgen voor welke vakken de prestaties achter blijven. Een tweede indicatie voor de maakbaarheid van de opbrengsten door de scholen kan ontleend worden aan het bepalen van die variabelen die variantie kunnen binden. Op basis van het onderzoek van Brandsma (1993) naar de rekenprestaties van leerlingen op de basisschool kan de variantie in de volgende componenten uiteengelegd worden (zie Tabel 6).

Tabel 6
Het binden van variantie in leerlingprestaties

	binnenschools	tussenschools
leerling covariabelen als ses, sekse, pretest, IQ, e.d.	50%	16%
schoolvariabelen	0%	10%
onverklaard	22%	2%

Van de oorspronkelijke 28% tussenschoolse variantie blijkt meer dan de helft te herleiden te zijn op verschillen in instroom van leerlingen met betrekking tot sociaal-economisch milieu van herkomst, sekse, intelligentie en parate

Tabel 7
De impact (als correlaties) van school- en klaskenmerken op leerlingprestaties in taal en rekenen/wiskunde (bron: Scheerens en Bosker, 1997, p. 305; lege cellen wijzen op het ontbreken van overzichtsstudies ter zake)

	kwalitatieve synthesen	internationaal vergelijkend onderzoek	meta-analyses/ 'best-evidence' synthesen
'resource' factoren			
leerling-leerkracht ratio		-0.03	0.02
scholing leerkracht		0.00	-0.03
ervaring leerkracht			0.04
salaris leerkracht			-0.07
uitgaven per leerling			0.20
schoolorganisatiefactoren			
productief klimaat	+		
prestatiedruk m.b.t. 'basics'	+	0.02	0.14
onderwijskundig leiderschap	+	0.04	0.05
monitoring/evaluatie	+	0.00	0.15
samenwerking en consensus	+	-0.02	0.03
betrokkenheid ouders	+	0.08	0.13
stafontwikkeling	+		
hoge verwachtingen	+	0.20	
orderlijk klimaat	+	0.04	0.11
instructiefactoren			
gelegenheid om te leren	+	0.15	0.09
taakgerichte leertijd	+	0.00 (n.s.)	0.19
huiswerk		-0.01 (n.s.)	0.06
directe instructie	+	-0.01 (n.s.)	0.11 (n.s.)
coöperatief leren			0.27
feedback			0.48
reinforcement			0.58
differentiatie/adaptieve instructie			0.22

schoolgeschiktheid. Ongeveer eenderde deel van de tussenschoolse variantie kan herleid worden tot verschillen in de organisatie en het onderwijs van de scholen. Welke factoren dat zijn, is te zien in Tabel 7.

Te zien is een patroon waarbij factoren die verder weg liggen van het primaire proces ('resources' en schoolfactoren) een minder sterke samenhang met de opbrengsten te zien geven dan factoren die er rechtstreeks mee te maken hebben (instructiekenmerken). Interessant met het oog op schoolverbetering is de vraag of scholen zelf weten hoe ze hun opbrengsten kunnen verhogen. Uit het eerder genoemde onderzoek van Van Amelsvoort et al. (1993) blijken echter met name de scholen met de lagere opbrengsten meer geneigd hun toevlucht te zoeken in oneigenlijke maatregelen als verscherpen van het doublure-, doorstroom- en aannamebeleid. De scholen met de hogere opbrengsten daarentegen noemen vaak meer eigenlijke maatregelen als het optimaliseren van het functioneren van de schoolorganisatie, de leerlingbegeleiding, en het selectie- en determinatiebeleid alsmede vakinhoudelijke verbeteringen.

De onderwijscononoom Monk (1992), het door economen uitgevoerde onderzoek naar onderwijsproductiefuncties overziend (onderwijskundigen zouden dat onderwijseffectiviteitsonderzoek noemen), komt tot de slotsom dat het bewijs voor het bestaan van een dergelijke functie mager is. In dat opzicht lijkt het standpunt van de schoolleiders die oneigenlijke maatregelen voorstellen om de opbrengsten te verbeteren begrijpelijk. Maar Monk komt met een alternatief. Wellicht, zo stelt hij, bestaat er een quasi-productiefunctie, die maar één productiviteitsverhogende factor kent: de probleemoplossende capaciteit van de school. Scholen met een voldoende probleemoplossende capaciteit zijn in staat gegeven de eigen omstandigheden een optimale mix van maatregelen te treffen om de productiviteit te verhogen. Om de eigen opbrengsten naar waarde te kunnen schatten is het dan niet alleen verstandig dat scholen zich meten aan een benchmark (bijvoorbeeld het nationaal gemiddelde), maar dat zij zich ook kunnen oriënteren op benchmark scholen die in dezelfde omstandigheden verkeren.

2.4 Tussenbalans

In het voorafgaande is de betekenis van het openbaar maken van resultaten van scholen nogal gerelativeerd. Ouders gebruiken de gegevens die beschikbaar komen zelden als doorslaggevende informatie bij de schoolkeuze. Bovendien verschillen scholen zo weinig, dat er in veel gemeenten geen betekenisvolle verschillen in prestaties zullen zijn. Wat betreft schoolverbetering is er ook slechts een beperkte betekenis, namelijk vooral door een basis te creëren voor benchmarking (de school kan zich vergelijken met andere scholen en met het landelijk gemiddelde). Voor schooldiagnose en zelfevaluatie is meer gedetailleerde informatie nodig. Verder zijn de marges voor schoolverbetering smal en is de instrumentele kennisbasis ('wat werkt er?') nog onvolledig. Een wellicht niet bedoeld effect kan zijn dat scholen zich, vooral in situaties van harde concurrentie, door selectie gaan richten op verbetering van de gemiddelde prestaties. Het publiceren van resultaten van scholen in het kader van accountability (het ter verantwoording houden van scholen door een hoger bestuurlijk niveau) lijkt in Nederland tot nu toe geen duidelijk motief. Ervaringen uit de VS met deze toepassing wijzen uit dat de daadwerkelijke benutting tegenvalt en dat het vooral een symbolische functie heeft.

3 De Techniek

3.1 Technische onvolkomenheden

Van der Putten en Van der Kamp (1998) hebben reeds gewezen op onzuiverheden in de door Dronkers gebruikte gegevens en op enkele merkwaardigheden rond het toepassen van de factoranalyse die Dronkers uitvoert om een schaal voor het interne rendement te kunnen ontwikkelen. Het verweer van Dronkers dat de onvolkomenheden het gevolg zijn van fouten door de inspectie gemaakt, snijdt onzes inziens geen hout. Die gegevens zijn door de inspectie met een bepaald doel eind verzameld, namelijk om in het kader van haar taken een gesprek op gang te brengen tussen de school en de toezichthoudend inspecteur over in-, door-, uitstroompatronen en examenresultaten. Dat stelt andere eisen aan de gegevensverzameling dan wanneer dergelijke gegevens gebruikt zouden

worden voor evaluatieve doeleinden². In de tweede *Trouw*-publicatie gebruikt Dronkers zelfs uitstroomgegevens met de suggestie dat het hier om schoolverlaters gaat, hetzij harde drop out, hetzij leerlingen die om redenen die met het genoten onderwijs te maken hebben van school wisselen. Dergelijke vormen van uitstroom zijn in de inspectiegegevens echter niet te scheiden van schoolwisseling vanwege verhuizing. Met andere woorden: hier zijn ernstige validiteitsproblemen, zo ernstig dat het onverantwoord is deze uitstroomgegevens bij de rendementsberekeningen te betrekken. Validiteitsproblemen ontstaan voorts door het veronachtzamen van de bekende niveauverschillen in de eindexamens in het VBO en MAVO en de verschillen tussen de vakken in becijfering (zie ook: Luyten, 1998). Nergens houdt Dronkers bovendien rekening met de (aanzienlijke) sectorverschillen die zich in het VBO voordoen³. Een validiteitsprobleem ontstaat ook door de datagestuurde wijze waarop Dronkers zowel de eerste als de tweede keer zijn rendementsmaat bepaalt. Weliswaar komt een bijdrage door doublures aan de factorscore in de verkeerde richting in de tweede publicatie niet meer voor, maar wat blijft is natuurlijk dat de factoroplossing afhankelijk is van de in de gegevens aanwezige structuur, en dat die oplossing noch uniek noch in de tijd stabiel noch constant over de schoolsoorten heen is⁴.

3.2 Toegevoegde waarde

Het presenteren van gemiddelde examencijfers of doublurepercentages per school als indicatoren voor de opbrengst is een gevaarlijke onderneming. Men kan zich namelijk de vraag stellen: Representeert zo'n indicator de inspanning die een school geleverd heeft? Deze vraag heeft sinds de jaren zestig centraal gestaan in het onderzoek naar effectieve scholen. Daar is immers het probleem aan de orde of, bijvoorbeeld, schoolleiderschap invloed heeft op de prestaties van leerlingen. Om dit oorzakelijk verband te kunnen leggen, dienen we ons ervan te vergewissen dat hogere prestaties niet veroorzaakt worden door de gezinsachtergrond van de leerlingen in plaats van door schoolleiderschap. In dit kader werken onderzoekers op het gebied van schooleffectiviteit met het begrip toegevoegde waarde: wat voegt de

school toe aan het niveau dat de leerlingen bij instroom in de school reeds hebben?

Achtergrondkenmerken van leerlingen als etniciteit, sociaal-economisch milieu van herkomst en geslacht worden in het onderzoek naar effectieve scholen vaak als proxies (d.w.z.: benaderingen) gebruikt voor het startniveau wanneer directe metingen ter zake (zoals scores op een Cito-entreetoets) niet voorhanden zijn. In de praktijk van het publiceren van schoolresultaten, zowel in Nederland als in andere landen (bijvoorbeeld Australië, Frankrijk en sommige staten in de VS), wordt bij ontstentenis van gegevens over individuele leerlingen zelfs een groepsniveau-proxy gehanteerd. Als kenmerk van de leerlingpopulatie wordt bijvoorbeeld het aantal leerlingen dat in aanmerking komt voor 'free meals' genomen (of in Nederland het percentage leerlingen waarvoor de school faciliteiten heeft ontvangen in het kader van de culturele minderheden regeling). Vervolgens worden de opbrengsten die een school gerealiseerd heeft voor dit kenmerk gecorrigeerd. De vraag is echter of dit een adequate manier van werken is, en of we met de aldus gecorrigeerde opbrengstgegevens de toegevoegde waarde te pakken hebben. Veenstra et al. (1998) hebben dit onderwerp in hun kritiek op Dronkers ook reeds aan de orde gesteld, onder meer verwijzend naar de dissertatie van Bosker (1990), maar Dronkers snapt kennelijk het punt niet. Waar het om gaat is dat het werken met geaggregeerde gegevens "dangerous at best, and disastrous at worst" is (Aitkin & Longford, 1986, p. 42). Waar het om draait bij het bepalen van de door scholen toegevoegde waarde is de vaststelling van de gemiddelde mate waarin individuele leerlingen boven de statistische verwachting presteren op een school. Dat is iets geheel anders dan de vaststelling of een schoolpopulatie als geheel boven de statistische verwachting presteert.

Dit probleem heeft een conceptuele en een statistische achtergrond. Conceptueel kan het percentage leerlingen uit culturele minderheden voor iets anders staan dan een simpele sommatie van individuele omstandigheden. Het geeft voor leerlingen weer in welke schoolpopulatie ze zich bevinden, en als zodanig is het dus een kenmerk van de context van het onderwijs. Is het gemiddelde instroomniveau

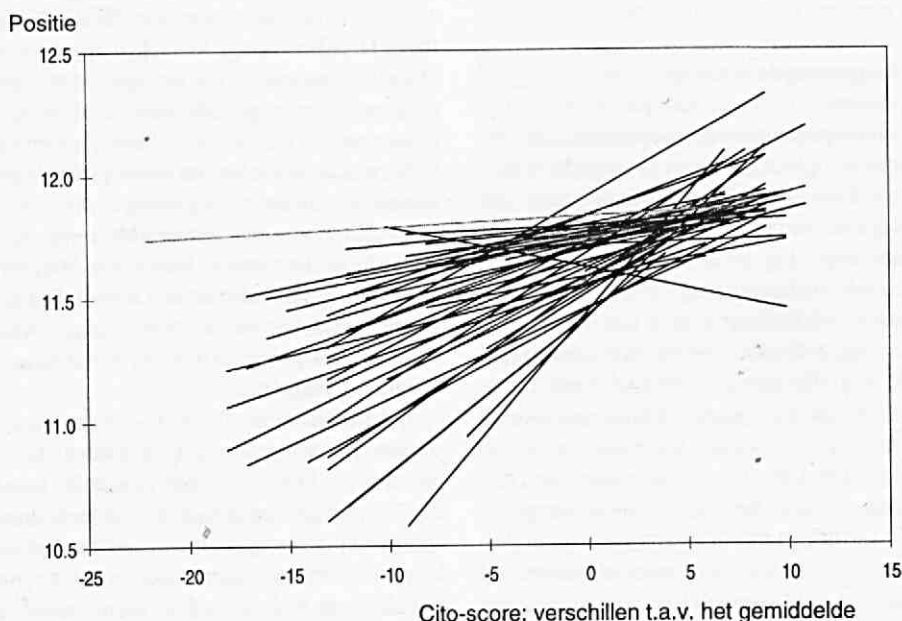
bijvoorbeeld aan de orde, dan kan dit geïnterpreteerd worden als een maat voor de rijkheid van de leeromgeving (Dar & Resh, 1986). Statistisch is er ook iets anders aan de orde als de relatie op leerlingniveau (relatie binnen scholen) tussen het instroomniveau en de examen-cijfers een andere sterkte en/of richting heeft dan de analoge relatie op het geaggregeerde niveau van de school (relatie tussen scholen). Correcties op leerling- dan wel op schoolniveau voor instroomverschillen vallen dan anders uit en scholen die bij de ene correctie positief uit de boot komen, komen dat dan in andere mate bij de andere correctie. In elk geval resulteert het uitvoeren van een correctie op geaggregeerd niveau zelden tot identieke resultaten als een correctie op leerlingniveau en in dat opzicht is het gebruik van de term toegevoegde waarde derhalve verwarrend.

3.3 Differentiële effecten

Tot nu toe hebben we het, zij het impliciet, alleen over de gemiddelde leerling gehad. Een school heeft hoge opbrengsten of juist niet. Maar binnen een school is er een enorme variatie: voor sommige leerlingen wordt wel en voor anderen juist niet een hoge opbrengst gerealiseerd. Dit is nu juist iets wat bij het schoolkeuzeproces een belangrijke rol zou moeten spe-

len: ouders vragen zich af of de school van hun voorkeur bij het kind past. Bijvoorbeeld: wat is de kans dat een leerling met een 'twijfel'-advies (HAVO-VWO) de betreffende school uiteindelijk met een VWO-diploma zal verlaten? De vraag is dus niet hoe goed de school het gemiddeld doet, maar wat de specifieke opbrengsten van de school voor dergelijke leerlingen zijn. Deze vraag verwijst naar wat we noemen differentiële schooleffecten: is een school gelijkelijk effectief voor de diverse te onderscheiden groepen leerlingen? De techniek om dit te onderzoeken ('slopes as outcomes') is wat lastig uit te leggen (de lezer zij verwezen naar Bosker & Snijders, 1990), maar we kunnen e.e.a. illustreren aan de gepresenteerde grafiek (zie Figuur 1).

De grafiek bevat voor elke school de lijn voor de regressie van de uiteindelijke schoolprestaties van een leerling op zijn of haar entreescore. Gemakshalve hebben we van elke entreescore het nationale gemiddelde afgetrokken, zodat de score 0 de score voor de gemiddelde leerling is. De grafiek maakt duidelijk dat een school die het goed doet voor de gemiddelde leerling het niet per se ook goed hoeft te doen voor leerlingen binnen die school die een van het gemiddelde afwijkende entreescore hadden. In dit voorbeeld zijn de verschillen tus-



Figuur 1. Differentiële opbrengsten van scholen

sen de scholen voor leerlingen die relatief laag op de entreetoets scoren het grootst. Voor een ouder is nu de vraag: gegeven de entreescore (bijvoorbeeld de score op de Cito-eindtoets), op welke school heeft mijn kind de beste vooruitzichten?

4 Toepassingscontexten en toegevoegde waarde

In het voorgaande hebben we stilgestaan bij de begrippen 'toegevoegde waarde' en 'differentiële toegevoegde waarde'. Het probleem dat Dronkers en ook de inspectie diende op te lossen was, dat er geen gegevens aanwezig waren over het instroomniveau van de leerlingen. In de eerder beschreven werkwijze hebben we aangegeven dat het eindniveau (examencijfers, behaald diploma, e.d.) te voorspellen is uit de parate en potentiële schoolgeschiktheid die de leerlingen hebben bij entree in de school voor voortgezet onderwijs. Het gaat dus om het cognitief niveau bij aanvang van de loopbaan in het secundair onderwijs. Dronkers gebruikt als proxy voor dat instroomniveau van een school het percentage leerlingen uit de culturele minderheden⁵. Daar is geen discussie over, zolang als er geen andere redenen zijn om deze variabele op te nemen. Dronkers echter verwijst naar het aan Bourdieu ontleende begrip 'cultureel kapitaal' (zie: De Graaf, 1987). Dat is echter een kenmerk van het gezin waarin het kind opgroeit, en daarmee lijkt het erop dat Dronkers deze variabele gebruikt niet alleen om te corrigeren voor het instroomniveau maar ook voor de onderwijsondersteuning die leerlingen in de thuissituatie genieten.

Dat is het essentiële probleem als er meer achtergrondgegevens van leerlingen bij de bepaling van de verwachte uitkomsten worden betrokken: een indicatie voor het sociaal-economisch milieu, bijvoorbeeld, naast een Cito-score kan alleen gelegitimeerd worden als men kan volhouden dat niet gemeten gezinsinvloeden (maar door het sociaal-economisch milieu van herkomst geïndiceerd) op de loopbaanontwikkeling van de leerling de meting van het werkelijke schooleffect verstoren. We raken hiermee aan een schier onoplosbaar probleem: de onmogelijkheid te komen tot een sluitende bewijsvoering over causale attributie, ofwel:

komt het door de school die ze bezoeken dat leerlingen het zo goed of slecht doen? Dit probleem is eerder opgemerkt door Raudenbush en Willms (1996) en door Grisay (1996). Het probleem ontstaat door keuzegedrag van leerlingen en hun ouders en door selectiegedrag van scholen (in plaats van dat leerlingen aselekt aan scholen worden toegewezen). Stel dat een school kwalitatief hoogwaardig onderwijs te bieden heeft (we doelen nu op het primaire proces en niet zozeer op de uitkomsten), en stel dat de leerlingen met een hoger schoolgeschiktheidsniveau en hun ouders bewust deze school kiezen⁶. Dan is het onderliggende model niet meer simpel input → proces → output, maar het wordt gecompliceerd tot proces → input → proces → output. Het statistisch model voor het berekenen van de toegevoegde waarde is echter gebaseerd op het eerste model, en de consequentie is, dat, als het tweede model geldig is, de correctie voor instroomverschillen tevens een onbedoelde correctie voor de kwaliteit van het onderwijsproces is. De implicatie is, dat men er verstandig aan doet zowel de bruto (bijv. ruwe examencijfers) als de netto opbrengstgegevens (de 'toegevoegde waarde', in casu examencijfers gecorrigeerd voor Cito-entreetoetsscores) in de beschouwing te betrekken want daar ergens tussenin ligt de echte edoch onbekende bijdrage van de school⁷.

We komen nu bij een ander probleem, en dat is dat het begrip toegevoegde waarde en zijn operationalisering afhankelijk is van de toepassingscontext. In de context van 'accountability' en schoolverbetering is het streven erop gericht zo zuiver mogelijk de inspanning die de school gepleegd heeft boven tafel te krijgen. De correcties kunnen in dit geval niet streng genoeg zijn: alle gegevenheden (onveranderbare omstandigheden) dienen bij de correctie betrokken te worden, niet alleen de Cito-entreetoets, maar ook de gezinsachtergrond, de omgeving waarin de school staat, kenmerken van de schoolpopulatie als geheel e.d. Wat resteert is dan een goede benadering voor wat vermoedelijk door inspanningen van de kant van de school gerealiseerd is. Voor ouders is maar één ding interessant, en om dat duidelijk te maken keren we terug naar het onderscheid in binnen- en tussenschoolse regressie. We hebben betoogd dat de relatie tussen het eindniveau en

het startniveau de basis vormt voor het berekenen van toegevoegde waarde. Normaliter impliceert het de volgende regressievergelijking (we zien gemakshalve even af van differentiële schooleffecten):

$$\text{eind}_{is} = \text{intercept} + \beta_1 \text{start}_{is} + R_{is} + U_s \quad (1)$$

Het subscript i staat voor leerling en s voor school. U_s is het schooleffect, ofwel de toegevoegde waarde. Het probleem zit hem nu in de regressiecoëfficiënt β_1 . Dit is een mixture van de binnen- en van de tussenschoolse regressie van het eindniveau op het startniveau. Die tussenschoolse regressie geeft het verband weer tussen het schoolgemiddelde eindniveau en het schoolgemiddeld startniveau en zoals we reeds hebben aangegeven kan men dit interpreteren als het effect van de leeromgeving. Maar ouders willen daar geen correctie voor: als hun kind goed leert omdat de schoolpopulatie gunstige karakteristieken heeft dan is dat mooi meegenomen.

Derhalve is het verstandiger dit effect expliciet te modelleren, zodat β_1 niet langer vervuild wordt door contextuele effecten.

$$\text{eind}_{is} = \text{intercept} + \beta_2 \text{start}_{is} + \beta_3 \text{start}_{is} + R_{is} + U_s \quad (2)$$

Als β_3 significant afwijkt van 0, dan is er inderdaad een contextueel effect. De toegevoegde

waarde waar ouders nu geïnteresseerd in zijn, bevat tevens het effect van de rijke leeromgeving:

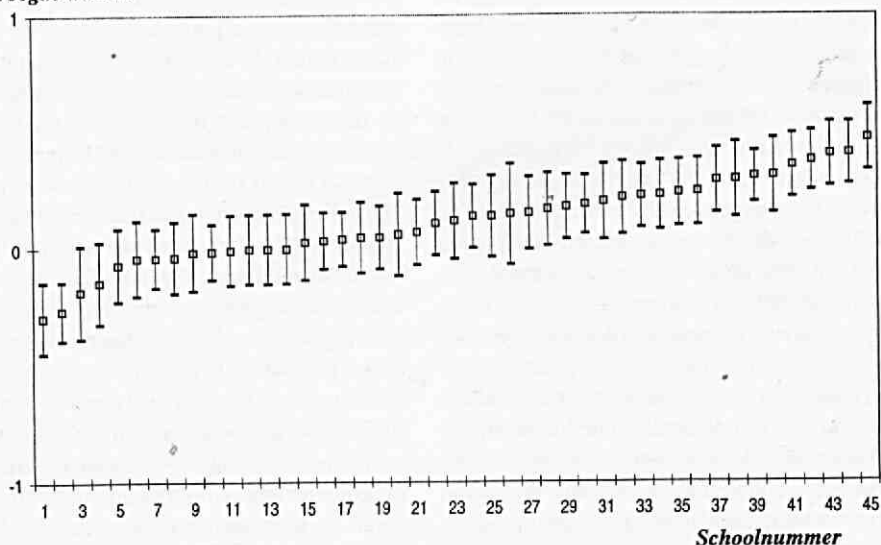
$$\beta_3 \text{start}_{is} + U_s$$

In het algemeen is het overigens zo dat na correctie voor instroomverschillen de verschillen in netto opbrengsten van scholen nog eens aanmerkelijk kleiner zijn dan de toch al beperkte verschillen die we in het begin van deze bijdrage hebben laten zien. Om een indruk van de omvang van die verschillen in toegevoegde waarde te geven, is Figuur 2 opgenomen, waarin de prestaties van leerlingen in het MAVO na 3 jaar onderwijs zijn gemeten. Helaas hebben zittenblijvers de betreffende toetsen voor wiskunde en Nederlands niet gemaakt zodat dit een door 'attrition bias' aangetaste en dus niet representatieve onderzoeksgroep is⁸. Alleen scholen zijn opgenomen met minimaal 25 leerlingen.

We zien dat de overlap tussen de scholen groot is: er zijn maar weinig scholen die significant van elkaar verschillen. Bedenk daarbij dat in die gevallen waar het zich toch voordoet, het vaak voor een kiezende ouder geen reële alternatieven betreft omdat de scholen niet beide bereikbaar hoeven te zijn.

Er zitten overigens nog enkele adders onder het gras. Opbrengstgegevens gebaseerd op

Toegevoegde waarde



Figuur 2. Verschillen in toegevoegde waarde tussen scholen inclusief 90%-betrouwbaarheidsintervallen

onder andere eindexamenresultaten betreffen een lichte leerlingen die vier tot zes jaar geleden hun onderwijsloopbaan in het voortgezet onderwijs zijn begonnen. Het kind en zijn ouders die voor de schoolkeuze staan, hebben met de toegevoegde waarde dus een gegeven in handen waarvan ze maar moeten hopen dat dit ook voor de huidige instroomlichting gerealiseerd zal worden. Het is in dat opzicht verstandiger de toegevoegde waarde apart te berekenen voor de brugperiode en dan nog eens voor de bovenbouw. De tweede adder betreft de schoolwisselingen: als een redelijk deel van de eindexamenleerlingen op een bepaalde school hun eerste jaren in het voortgezet onderwijs op andere scholen heeft doorgebracht, welke school krijgt dan de credits?

5 Ten slotte

In deze bijdrage zijn de diverse toepassingscontexten voor het openbaar maken van schoolresultaten alsmede de technische kant van de berekening daarvan de revue gepasseerd. Geconstateerd kan worden dat de communis opinio is, dat het openbaar maken van schoolresultaten om uiteenlopende redenen verstandig is. In het publieke debat wordt meermalen gewezen op de mogelijke negatieve effecten, met name dan het strategische gedrag dat één en ander kan uitlokken van de zijde van de scholen. Smith (1995) heeft alle gevaren op een rij gezet. De voornaamste die in Nederland een rol spelen, en een mogelijke remedie, zijn:

1. Kortzichtige gerichtheid op gunstige scores voor de eerstvolgende publicatie (iedereen gaat over en er wordt tot in de treure geoeffend op examens). Oplossing: vooruitzicht dat met de praktijk van het publiceren lang doorgaan zal worden, en een redelijk omvangrijke reeks kengetallen in plaats van slechts één of twee.
2. De school richt zich alleen op de kengetallen en vergeet niet-kwantificeerbare relevante uitkomsten (zoals sociale vaardigheid, zelfstandig werken). Oplossing: produceren een kengetal voor extern rendement.
3. De school richt zich op selectie. Wanneer dat streng selecteren betekent, is toegevoegde waarde het kengetal dat dit probleem oplost. Wanneer scholen de systematiek van

toegevoegde waarde begrijpen, kunnen ze zich richten op de minst vaardige leerlingen qua startniveau: met deze groep is immers de grootste winst te boeken. Oplossing: vermeld ook steeds bruto scores op kengetallen.

4. Inperking van het innovatief vermogen. Dat is wellicht de grootste bedreiging, en alleen adequaat op te lossen door ervoor te zorgen dat de set kengetallen flexibel is en steeds een aantal nieuwe bevat.

In het algemeen zijn er voor de problemen die Smith aansnijdt dus wel oplossingen denkbaar. Belangrijk is dat men het publiceren van opbrengstgegevens niet als alleen zaligmakend ziet, en hiervoor is aangegeven dat binnen de diverse toepassingscontexten de nodige relativeringen gemaakt dienen te worden over de nuttigheid van deze praktijk.

Ondanks deze relativeringen moet het toch beschouwd worden als een nuttige praktijk die bijdraagt aan de rationaliteit van keuzegedrag, besturingsprocessen op scholen zelf en van het overheidsbeleid. Naar onze mening is vooral het feit dat de aandacht op prestatie-indicatoren wordt gelegd een belangrijke prikkel voor scholen om meer opbrengstgericht te gaan opereren. Een dergelijke wijziging in de schoolcultuur zal naar verwachting leiden tot prestatieverbetering.

Noten

1. Welk een revolutie zou er uitbreken als een leraar eens zo normloos zou beoordelen. Met de toegepaste methodiek – 'grading on the curve', met een curve die loopt van 1 tot en met 10 – heeft Dronkers schromelijk onderschat welk een maatschappelijke implicaties het uitdelen van rapportcijfers (want die begrijpt iedereen) aan scholen kan hebben. In dit opzicht is de tweede Trouw-publicatie een vooruitgang.
2. In die context is het niet bezwaarlijk dat de gegevens enigszins vervuild zijn. Men moet zich natuurlijk wel realiseren dat het hier om een kosten-batenafweging gaat: het construeren van een volledig betrouwbaar en valide gegevensbestand kost nu eenmaal menskracht.
3. Sommige vakken hebben om onduidelijke redenen een hoger gemiddelde voor het centraal

- eindexamen dan andere vakken. In 1997 deed zich dit bijvoorbeeld bij wiskunde A in het VWO voor, dat met 6.9 zeer hoog scoorde. Als een school nu 'toevallig' veel leerlingen heeft die wiskunde A volgen, dan wordt het schoolgemiddeld cijfer (Dronkers middelt over wiskunde A, Engels en Nederlands) kunstmatig hoog.
- 4 Het volledig bouwen op de resultaten van de factoranalyse leverde de eerste keer zelfs een resultaat op in het VWO waarin de richting van een bijdrage aan de rendementsscore afhankelijk was van het leerjaar!
 - 5 In de kwaliteitskaarten van de inspectie is dat ook precies de reden voor de selectie van een aantal kenmerken (percentage leerlingen uit culturele minderheden, percentage leerlingen waarvan de ouders een tegemoetkoming in de studiekosten ontvangen, vier grote steden versus overig Nederland en denominatie) die tezamen als 'stand in' voor het onbekende instroomniveau fungeren (Inspectie van het onderwijs, 1998, p. 8).
 - 6 Mutatis mutandis ontstaat hetzelfde probleem als kwalitatief hoogwaardige scholen hun leerlingen gericht selecteren.
 - 7 Zulks onder de aanname dat niet één der betrokken partijen tegenovergestelde doelen nastreeft.
 - 8 De redenering van Van der Werf en Creemers dat dit juist goed is, is volstrekt onjuist en misleidend.

Literatuur

- Aitkin, M., & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society, Series A (General)* 149, (Part 1), 1-43.
- Amelsvoort, H.W.C.H. van, Bosker, R.J., & Brandsma, T.F. (1993). Rendementsbekostiging in de bovenbouw van HAVO en VWO; overwegingen en reacties. *Pedagogische Studiën*, 70, 283-297.
- Bosker, R.J. (1990). *Extra kansen dankzij de school* (Academisch proefschrift). Nijmegen: ITS/OoMO.
- Bosker, R.J. (1998). Kwaliteit als toegevoegde waarde. In R.J. Bosker e.a., *Het vergelijken van scholen* (pp. 85-102). Enschede: OCTO/Universiteit Twente.
- Bosker, R.J., & Snijders, T.A.B. (1990). Statistische aspecten van multi-niveau onderzoek. *Tijdschrift voor Onderwijsresearch*, 15, 317-329.
- Brandsma, H.P. (1993). *Basisschoolkenmerken en de kwaliteit van het onderwijs*. Groningen: RION.
- Cibulka, J.G., & Derlin, R.L. (1995) State educational performance reporting policies in the U.S.: accountability's many faces. *International Journal of Educational Research*, 23, 479-492.
- Dar, Y., & Resh, N. (1986). Classroom intellectual composition and academic achievement. *American Educational Research Journal*, 23, 357-374.
- Derriks, M., Kat E. de, & Deckers, P. (1995). *Schoolkeuzemotieven van kinderen bij de overgang bo/vo* (SCO-rapport 417). Amsterdam: SCO-Kohnstamm Instituut.
- Dronkers, J. (1998). Het betere is de vijand van het goede. Een reactie op de commentaren over het Trouw rapportcijfer. *Pedagogische Studiën*, 75, 142-150.
- Graaf, P.M. de (1987). *De inhoud van financiële en culturele hulpbronnen in onderwijsloopbanen*. Nijmegen: ITS/OoMO.
- Grisay, A. (1996). *Evolution des acquis cognitifs et socio-affectifs des élèves au cours des années de college*. Liège: Université de Liège.
- Inspectie van het Onderwijs (1998). *De kengtallen op de kwaliteitskaart. Een technische toelichting*. Leiden: SDU-servicecentrum.
- Lam, J.F. (1998). Publicatie van schoolgegevens: een internationale vergelijking. In R.J. Bosker e.a., *Het vergelijken van scholen* (pp. 11-33). Enschede: OCTO/ Universiteit Twente.
- Luyten, H. (1994). Stability of school effects in Dutch secondary education: The impact of variance across subjects and years. *International Journal of Educational Research*, 21, 197-216.
- Luyten, H. (1998). Moeilijke en minder moeilijke examens. In R.J. Bosker e.a., *Het vergelijken van scholen* (pp. 37-51). Enschede: OCTO/Universiteit Twente.
- Monk, D.H. (1992). *Education productivity research: An update and assessment of its role in education finance reform*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.
- Putten, C.M. van, & Kamp, L.J.T. van der, (1998). Dronkers' rapportcijfers voor de middelbare scholen: nauwkeurigheid, geldigheid en betekenis. *Pedagogische Studiën*, 75, 135-141.
- Raudenbush, S.W., & Willms, J.D. (1996). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20, 307-335.
- Scheerens, J. (1998). *Decentralization, school improvement and school evaluation*. Copenhagen:

invited address for the European Conference on School Improvement and Quality Development of Schools.

- Scheerens, J., & Bosker, R.J. (1997). *The foundations of educational effectiveness*. Oxford: Elsevier Science.
- Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration*, 18, 277-310..
- Veenstra, D.R., Dijkstra, A.B., Peschar, J.L., & Sijders, T.A.B. (1998). Scholen op rapport. Een reactie op het Trouw-onderzoek naar schoolprestaties. *Pedagogische Studiën*, 75, 121-134.
- Werf, M.P.C. van der, & Creemers, B.P.M. (1999). Naar een volledig systeem van kwaliteitsbeoordeling van scholen. *Pedagogische Studiën*, 76, 51-60.

Auteurs

R.J. Bosker en J. Scheerens zijn als hoogleraren verbonden aan de faculteit der Toegepaste Onderwijskunde (afdeling Onderwijsorganisatie en -management en het OCTO) van de Universiteit Twente.

Adres: Universiteit Twente, Faculteit der Toegepaste Onderwijskunde, Afdeling Onderwijsorganisatie en -management, Postbus 217, 7500 AE Enschede

Abstract

Publishing performance data of schools; issues of usefulness and validity

R.J. Bosker & J. Scheerens. *Pedagogische Studiën*, 1999, 76, 61-73.

The immediate motive for this article is the recently started practice in the Netherlands to publicly report performance indicators on secondary schools. Three contexts of application of performance reporting are discussed: supporting school choice, school improvement and external accountability. Given the uncertainty and relatively small size of output differences between schools, incomplete knowledge about education production functions, and limited discretion of higher administrative levels over secondary schools in the Netherlands, performance reporting is seen as being of only limited relevance in each of these contexts of application.

In a discussion on the meaning and measurement of 'value added', it is argued that any measure should start with a definition based on the degree of overachievement or underachievement of individual students. Parents, quite obviously, have a different kind of value added in mind than other stakeholders: they want to know what a school can mean for their child, whereas others want to have a 'clean' as possible measure for the true impact of the school (i.e. by efforts of the teachers). Next to this parents are interested in differential value added: the specific meaning a school can have for their own child given its cognitive and other relevant abilities. The article closes with a discussion of possible unintended consequences of publishing school performance reports. The final conclusion is, nevertheless, that the increased consciousness about output performance is likely to act as an achievement incentive to schools thus leading to increased performance of secondary schools.

Het betere blijft nog steeds de vijand van het goede.

Een antwoord op Van der Werf en Creemers, en Bosker en Scheerens

J. Dronkers

SCO-Kohnstamm Instituut, Universiteit van Amsterdam

Samenvatting

In deze bijdrage wordt ingegaan op de kritiek van Bosker en Scheerens en van Van der Werf en Creemers op mijn berekening van de toegevoegde waarde voor het intern rendement door scholen in het voortgezet onderwijs, uitgevoerd in het kader van de *Trouw*-publicaties. Zes aspecten komen in mijn antwoord aan de orde: de doelen van het openbaar maken van prestatiegegevens van scholen; de verschillende dimensies in de prestaties van scholen; betekenis en berekening van toegevoegde waarde; longitudinaal versus cross-sectioneel meten van intern rendement; het belang van differentiële opbrengsten van scholen; technische kritiek. De conclusie van deze reactie is dat mijn benadering de best mogelijke oplossing biedt voor een aantal bezwaren tegen het publiceren van de inspectiegegevens.

In de artikelen van Bosker en Scheerens, en Van der Werf en Creemers, die in dit nummer van *Pedagogische Studiën* zijn opgenomen, komen interessante kwesties aan de orde met betrekking tot de twee *Trouw*-publicaties (Agerbeek, Hageman, Kreulen & Lakmaker, 1997; Agerbeek, 1998) van de inspectie-gegevens over scholen. Een aantal foutieve interpretaties door de tweede groep auteurs van mijn werkwijze bij deze *Trouw*-publicaties laat ik buiten beschouwing, omdat een herhaalde weerlegging daarvan mij niet erg relevant lijkt voor het debat over de verdere aanpak van kwaliteitsmeting. Ik volsta

op dat punt te verwijzen naar mijn vorige bijdrage daarover in *Pedagogische Studiën* (Dronkers, 1998b).

1 De doelen van openbare prestatiegegevens van scholen

Bosker en Scheerens noemen terecht drie doelen van het openbaar maken van inspectiegegevens over de prestaties van scholen: helpen van ouders bij schoolkeuze; helpen van scholen bij zelfevaluatie en schoolverbetering; het uitoefenen van externe controle op het functioneren van scholen. Elk van deze doelen kan zijn eigen logica hebben en uiteenlopende eisen stellen aan de wijze van openbaar maken. Maar Bosker en Scheerens beoordelen echter het belang van dit openbaar maken onjuist, omdat zij te sterk leunen op Amerikaanse literatuur.

Ten aanzien van het helpen van ouders bij de schoolkeuze betogen zij aan de hand van hun Tabellen 1 en 2 dat in de meerderheid van de Nederlandse gemeenten de ouders per schoolsoort geen keuze hebben. Zij veronderstellen dat ouders in hun schoolkeuze in het voortgezet onderwijs zich beperken tot hun eigen woongemeente. Dat is echter niet waar. Veel ouders laten hun kinderen voortgezet onderwijs volgen in een andere dan hun woongemeente. Dat kan ook goed, omdat het openbaar vervoer in Nederland goed is en de afstanden tussen gemeenten klein. Dat geldt niet alleen voor het platteland richting stad, maar ook voor grote steden richting forensengemeenten (bijvoorbeeld Amsterdam en Amstelveen). Door de Nederlandse verhoudingen valt er voor ouders meer te kiezen dan Bosker en Scheerens suggereren.

Ten aanzien van het uitoefenen van externe controle op het functioneren ('accountability') onderschatten zij het belang van het openbaar maken. Het uitoefenen van externe controle en het publiceren van de uitkomsten van die externe controle vloeit voort uit het gebruik van nationaal verzameld belastinggeld voor de financiering van een nationaal onderwijsstelsel, zoals die in Nederland onder de Franse invloed is gevormd. Angelsaksische onderwijsstelsels zijn daarentegen niet opgezet als een nationaal onderwijssysteem, maar op basis van de markt of de lokale gemeenschap (Archer,

1984). Door de daaruit voortvloeiende Angelsaksische autonomie van scholen is het publiceren allereerst een kwestie van het functioneren van die markt of die lokale gemeenschap. Maar in nationaal opgezette onderwijsstelsels zijn de uitgangspunten bij publicatie van schoolgegevens het controleren of de nationaal vastgestelde doelen gehaald worden en de publieke verantwoording van die controle. Daarom is het logisch dat in het Franse onderwijsstelsel met relatief weinig autonomie voor scholen prestatiegegevens van scholen door de overheid openbaar gemaakt worden, en is dat geen merkwaardige uitzondering zoals Bosker en Scheerens ten onrechte menen. Gezien deze controlerende taak van de onderwijsinspectie is het merkwaardig dat Bosker en Scheerens beweren dat de Nederlandse inspectiegegevens alleen verzameld werden om een gesprek op gang te brengen tussen de school en de toezichthoudende inspecteur, maar dat deze inspectiegegevens geen evaluatieve doeleinden zouden hebben. Als dat waar is, zou het woord 'toezichthoudend' verwijderd moeten zijn en zou een andere overheidsdienst belast moeten zijn met het namens de Nederlandse burgers en overheid controleren van het functioneren van het onderwijs. Dat woord is niet verwijderd en die andere dienst bestaat niet. Dus mogen Nederlandse burgers er op vertrouwen dat de onderwijsinspectie namens hen goede gegevens over scholen verzamelt waarmee die inspectie in staat is het onder haar toezicht staande onderwijs op zijn kwaliteit te evalueren¹. Het recht van de Nederlandse burger op informatie over het handelen van zijn overheid (inclusief de controle die de overheid namens zijn burgers uitoefent) was dan ook de grond van het gerechtelijk vonnis op grond waarvan *Trouw* de inspectiegegevens kreeg. Als de inspectie haar taak goed uitvoert, mogen de inspectiegegevens niet ernstig vervuild zijn. Dat bleken die gegevens wel te zijn, bijvoorbeeld omdat sommige scholen verzuimden uitzonderlijke groepen leerlingen (asielzoekers, allochtone zij-instromers die voor het eerst in Nederland zijn, drop-out projecten) uit hun opgaven te verwijderen (Dronkers, 1998a). Met vervuilde gegevens wordt de inspecteur door de scholen te gemakkelijk met een kluitje het riet in gestuurd. Bovendien vergeten Bosker en Scheerens dat gemeente- en schoolbe-

sturen voor het uitoefenen van hun taken als bevoegd gezag van scholen inzicht moeten hebben in het functioneren van hun scholen, juist in vergelijking met andere scholen en gemeenten. Naar aanleiding van de *Trouw*-publicaties is mij uit reacties van openbare en bijzondere schoolbestuurders gebleken hoe weinig informatie zij op dit punt kregen van hun scholen en de onderwijsinspectie². Het lijkt mij dus onjuist dat Bosker en Scheerens in hun Tussenbalans stellen dat het ter verantwoording houden van scholen door een hoger bestuurlijk niveau geen duidelijk motief was. Bij het uitoefenen van externe controle binnen een nationaal onderwijsstelsel is het overigens onlogisch alle onveranderbare omstandigheden (zoals denominatie, vier grote steden; zie de kwaliteitskaart) in de correcties te verwerken. Het Nederlandse onderwijsstelsel garandeert nu juist zijn burgers dat alle scholen min of meer gelijkwaardig functioneren, ongeacht denominatie, de gemeente of andere onveranderbare omstandigheden. Het is daarom merkwaardig dat in beide artikelen gepleit wordt voor het controleren voor alle omstandigheden waarop een school geen invloed zou hebben. Als het openbaar onderwijs een minder effectieve bestuursvorm heeft die negatieve consequenties heeft voor de studieresultaten (zie Dijkstra, Dronkers & Hofman, 1997), dan moet dat in het kader van externe controle juist zichtbaar blijven en niet worden weggecorrigeerd zoals in de kwaliteitskaart van de inspectie gebeurt³. Interessant is het op te merken dat Van der Werf & Creemers in hun bijdrage geen oog hebben voor het doel externe controle.

Ten aanzien van de mogelijkheden van schoolverbetering met het middel van publicatie van prestatiegegevens zeggen Bosker en Scheerens niets nieuws. Het is ook mij reeds lang bekend dat schoolkenmerken niet veel individuele variantie verklaren. Maar dat is geen reden om scholen niet aan te spreken op datgene wat ze wel kunnen, en Tabel 7 van Bosker en Scheerens laat zien dat er genoeg mogelijkheden voor scholen zijn. Inderdaad is het vanuit het doel schoolverbetering wel logisch meer rekening te houden met omstandigheden waarop de school geen invloed heeft. Maar men zal niet snel allerlei omstandigheden voor scholen onveranderbaar moeten verklaren, omdat ook dat het innovatief vermogen van scholen kan aantasten.

2 Kwaliteitsdimensies

In alle *Trouw*-publicaties kan het voor de goede lezer duidelijk zijn geweest dat de inspectiegegevens alleen betrekking hadden op het intern rendement van scholen. Meer was er niet en scholen, onderwijsorganisaties en OC&W hadden het ook niet nodig gevonden andere kwaliteitsindicatoren systematisch te laten verzamelen door de inspectie. Uiteraard zijn er meer kwaliteitsdimensies dan alleen het intern rendement. Naast het intern rendement zie ik nog drie andere kwaliteitsdimensies, die afgeleid zijn uit de maatschappelijke functies van het onderwijs: extern rendement (succes in het vervolgonderwijs of op de arbeidsmarkt), sociaal-culturele vorming (waaronder culturele participatie, bepaalde waarden en normen) en maatschappelijke integratie (waaronder het voorkomen van misdadig gedrag)⁴. Uiteraard kunnen binnen deze kwaliteitsdimensies nadere verfijningen gemaakt worden. Zo onderscheiden Van der Werf en Creemers binnen de dimensie van het intern rendement effectiviteit, rendement en efficiëntie. Dat is hun goed recht. Maar zij overdrijven het belang van hun onderscheid. Voor de tweede *Trouw*-publicatie heb ik de onderlinge samenhang tussen de verschillende aspecten van het intern rendement (uitstroom, zittenblijven, gemiddeld cijfer voor centraal schriftelijk deel, percentage bèta-pakketten, percentage geslaagden) onder andere met een factoranalyse nagegaan (Dronkers, 1998c) (zie Tabel 1).

De hoogte van deze factorladingen, de omvang van de verklaarde variantie en de eigen values wijzen steeds op een tamelijk robuuste dimensie, die een overkoepelende

maat voor het intern rendement rechtvaardigt. Dat was overigens geen unieke vondst van mij. Reeds in 1989 concluderen Bosker, Guldemond, Hofman en Hofman (1989, p. 91) dat "scholen die het op de ene indicator⁵ goed doen, ook vaker hogere scores behalen op andere indicatoren: dus hoge slaagpercentages gaan samen met minder doublures en een hogere retentiviteit van de school. Aangezien wij op diverse bestanden steeds tot deze resultaten geraken, zelfs als we een en ander uitsplitsen naar diverse opleidingstypes, kunnen wij concluderen dat deze bevindingen redelijk betrouwbaar zijn". Op grond daarvan komen deze auteurs dan ook tot één kwaliteitsindicator: een combinatie van de genoemde loopbaanindicatoren (Bosker et al., 1989, p. 92). In het licht van mijn analyses en deze Groningse studie is het dus maar de vraag of de effectiviteit, het rendement en de efficiëntie van scholen erg veel van elkaar zullen verschillen. Bovendien belemmert het gebruik van een overkoepelende maat strategisch gedrag van scholen, want scholen die met het oog op gunstige slaagpercentages of hoge eindexamencijfers sterk gaan selecteren (hoge uitval, veel zittenblijven), worden door mijn overkoepelende maat juist niet beloond, in tegenstelling tot wat Van der Werf en Creemers beweren.

Bosker en Scheerens menen dat deze factoranalyse niet valide is omdat zij noch in de tijd stabiel, noch over de schoolsoorten constant is. Waarom het gelijk zijn over schoolsoorten een validiteitseis zou zijn, ontgaat mij. De verschillende onderwijssoorten zijn te verschillend in hun functioneren om te veronderstellen dat de verschillende aspecten van het intern rendement steeds dezelfde betekenis en hetzelfde

Tabel 1

De factorladingen, eigen values en verklaarde varianties van het totaal intern rendement van VWO-, HAVO-, MAVO- en VBO-vestigingen

	VWO	HAVO	MAVO	VBO
Slaagpercentage	.89	.91	.87	.83
Gem.eindexamen cijfer	.84	.86	.83	.41
Gem. zittenblijven	-.74	-.72	-.61	-.72
Gem. uitval	-.66	-.74	-.73	-.80
% bèta vakken	.37	.35	n.v.t.	n.v.t.
% examen D-niveau	n.v.t.	n.v.t.	.70	.32
Eigen value	2.6	2.8	2.9	2.1
% variantie	52.3	55.1	57.2	42.3
Aantal vestigingen	498	478	691	414

Bron: Dronkers, 1998c

gewicht zouden hebben. Ik heb nooit voorgesteld elk jaar een nieuwe factoroplossing te maken. Het mooiste zou zijn op grond van een aantal schooljaren de gewichten van de verschillende aspecten vast te stellen. Daarvoor ontbreken tot op dit moment de noodzakelijke inspectiegegevens, want die zijn niet voor meerdere jaren beschikbaar. Mijn oplossing in de *Trouw*-publicatie was daarom een pragmatische, die gezien de resultaten in Tabel 1 zo gek nog niet is.

3 Toegevoegde waarde

Bosker en Scheerens wijzen terecht op de conceptueel en statistisch belangrijke verschillen tussen het werken met geaggregeerde of individuele gegevens. Dat betekent onder meer dat men met de twee typen gegevens verschillende vragen kan beantwoorden. Met geaggregeerde gegevens kan men de vraag beantwoorden "hoe hoog het intern rendement van een school is vergeleken met die van scholen met een vergelijkbare leerlingpopulatie". Met individuele leerlinggegevens kan men de vraag beantwoorden wat "de gemiddelde vooruitgang is van vergelijkbare leerlingen van verschillende scholen". Maar Bosker en Scheerens zeggen daar niets nieuws mee. Bovendien heb ik nooit gepretendeerd met de geaggregeerde inspectiegegevens de tweede vraag goed te beantwoorden. Wel heb ik steeds betoogd dat het antwoord op de eerste vraag ook interessant is, juist ook voor de doelen schoolverbetering en externe controle (Dronkers, 1998a, 1998b). Bosker en Scheerens reserveren op eigen gezag de term 'toegevoegde waarde' van een school voor de gemiddelde waarde waarin individuele leerlingen boven de statistische verwachting presteren op school (de tweede vraag), en verwijten mij vervolgens dat ik de term 'toegevoegde waarde' verwarrend gebruik. Dat is echter gezien de internationale praktijken een particuliere opvatting van Bosker en Scheerens. In andere landen waar schoolresultaten worden gepubliceerd en waar meestal ook alleen geaggregeerde gegevens beschikbaar zijn, wordt de term 'toegevoegde waarde' (value added; valeur ajoutée) gebruikt als men daar de eerste vraag beantwoordt. Derhalve meen ik dat mijn gebruik van de term 'toege-

voegde waarde' meer in lijn is met de internationale gebruiken dan de particuliere opvatting van Bosker en Scheerens, die deze term alleen willen reserveren voor individuele gegevens. Ten slotte wil ik nog opmerken dat Bosker geaggregeerde gegevens blijkbaar niet zo gevaarlijk en rampzalig vindt, dat hij het aanbod heeft afgeslagen om met geaggregeerde gegevens voor de kwaliteitskaart van de inspectie prestaties van vergelijkbare scholen te berekenen. Ook bij geaggregeerde data vindt Bosker dus dat het betere de vijand van het goede kan zijn.

Van der Werf en Creemers hebben niet begrepen dat ik in beide *Trouw*-publicaties achtergrondkenmerken van leerlingen zoals etniciteit en zij-instroom gebruikt heb als een redelijke indicator van het startniveau van de leerlingen. Meer was niet voorhanden bij inspectie of OC&W. Ook in andere landen worden deze indicatoren gebruikt, als het aanvangsniveau ontbreekt⁶. Analyses van de VOCL-cohorten hebben in het verleden reeds laten zien dat deze achtergrondkenmerken een redelijke indicator vormen van het aanvangsniveau van leerlingen en Veenstra, Dijkstra, Peschar en Snijders (1998) hebben dat nogmaals laten zien. Als Van der Werf en Creemers hun eigen analyse van het VOCL-cohort nog eens nader bekijken, zullen zij zien dat het percentage allochtone leerlingen een redelijke indicator van het aanvangsniveau van leerlingen is. Uiteraard is het het beste als zo spoedig mogelijk een goede indicator van het aanvangsniveau (bijvoorbeeld gemiddelde score op de Cito-toets) beschikbaar komt. Maar waarschijnlijk zal dat nog wel enkele jaren kosten. Tot dat moment zal men het met indicatoren van het aanvangsniveau moeten doen. Het is dan echter van belang daarbij voorzichtig te zijn, omdat indicatoren van het aanvangsniveau gedeeltelijk samenvallen met schoolkenmerken. Zo zijn in de kwaliteitskaart de denominatie, de vier grote steden en de breedte van de scholengemeenschappen, tezamen met het percentage allochtone leerlingen en leerlingen met een tegemoetkoming van studiekosten, gebruikt om het aanvangsniveau van leerlingen te schatten. Dit is een mooi voorbeeld waarin kenmerken van effectieve scholen ten onrechte gelijkgesteld worden met goede indicatoren van aanvangsniveau van leerlingen, waardoor

verschillen tussen het functioneren van scholen uit het zicht verdwijnen. Bosker en Scheerens leiden uit mijn gebruik van de term 'cultureel kapitaal' als samenvatting van relevante achtergrondkenmerken af dat ik voorstander zou zijn van het corrigeren voor de onderwijsondersteuning naast de correctie voor het aanvangsniveau van leerlingen. Dat is onjuist: alleen correctie voor het cognitieve aanvangsniveau lijkt mij legitiem bij de vaststelling van de toegevoegde waarde van scholen. Correcties voor alle andere (leerling)kenmerken (inclusief etniciteit, sociaal-economische achtergrond) leidt tot verwarring van leerling- en schoolkenmerken. Van der Werf en Creemers pleiten voor het gebruik van het advies van de basisschool als meting van het aanvangsniveau in het voortgezet onderwijs. Door de overadviesgeving in bepaalde regio's en bij bepaalde denominaties (Dronkers, Van Erp, Robbijn & Roelvel, 1998) lijkt mij dat onverstandig, omdat scholen dan afgerekend worden op geïnflateerde verwachtingen (adviezen) in plaats van hun functioneren.

4 Longitudinaal versus cross-sectioneel meten

Zowel Bosker en Scheerens als Van der Werf en Creemers lijken er voor te pleiten alleen met longitudinale gegevens te werken bij het openbaar publiceren van prestatiegegevens van scholen. Bosker en Scheerens doen dit vooral op grond van hun voorkeur voor een multi-level aanpak, Van der Werf en Creemers vooral op grond van hun definities. Dat pleidooi lijkt verstandig, maar aan het gebruik van longitudinale gegevens voor dit doel kleven zowel praktische als conceptuele bezwaren.

Het praktische bezwaar is simpel: longitudinale gegevens over alle scholen bestaan er op dit moment niet en zullen er de eerstkomende jaren ook nog niet zijn. Het VOCL-cohort is maar een steekproef van alle scholen in het voortgezet onderwijs en is dus onbruikbaar voor het openbaar publiceren van prestatiegegevens van scholen: te veel scholen ontbreken of komen maar gedeeltelijk in het cohort voor. Bovendien mogen de VOCL-gegevens van het Centraal Bureau voor de Statistiek niet zodanig gebruikt worden dat scholen met naam en toe-

naam herkend kunnen worden. Beide artikelen pleiten dus voor iets wat voorlopig niet kan: het betere blijft dan nog steeds de vijand van het goede. Gelukkig is Bosker flexibel genoeg dat zijn voorkeur voor longitudinale data hem niet verhinderd heeft mee te werken aan de kwaliteitskaart van de inspectie, waarin alleen cross-sectionele gegevens gebruikt worden. Hij maakt daarin dezelfde aanname als ik voor de *Trouw*-publicaties heb moeten maken, maar die Van der Werf en Creemers zo krachtig zonder nadere empirie verwerpen: "percentages zittenblijvers en drop-out zijn redelijke indicatoren van de tijd die het leerlingen kost om het diploma te halen". Opnieuw blijkt Bosker het haalbare goede te verkiezen boven het onbereikbare betere.

Het conceptuele bezwaar is lastiger. Voor de verschillende doelen van het openbaar maken van prestatiegegevens over scholen gaat het vooral om het huidige functioneren van de school, niet om het functioneren van de school vier tot acht jaar geleden. Prestatiegegevens over scholen, gebaseerd op longitudinale gegevens, verwijzen echter per definitie naar het functioneren van de school vier tot acht jaar geleden. Men kan immers pas met longitudinale gegevens een uitspraak over het functioneren van scholen doen als de laatste zittenblijvers van dat cohort de scholen verlaten hebben. Kortom, het openbaar publiceren van prestatiegegevens van scholen uitsluitend op basis van longitudinale gegevens komt voor de externe controle, voor het kiezen van ouders en voor schoolverbetering als mosterd na de maaltijd. Scholen zullen dan begrijpelijk zeggen dat het slechte functioneren te wijten is aan de moeizame fusie of de moeilijkheden van een aantal jaren geleden, maar dat die problemen nu overwonnen zijn. Bosker en Scheerens vermoeden dit conceptuele bezwaar als zij erop wijzen dat longitudinale gegevens gebaseerd zijn op een lichte leerlingen die vier tot zes⁷ jaar geleden hun onderwijsloopbaan begonnen zijn. Cross-sectionele data⁸ zijn daarom beter om het huidige functioneren van scholen te meten: het gevolg van een moeizame fusie of van andere moeilijkheden die het goed functioneren van scholen belemmeren, zijn duidelijker en sneller zichtbaar: in een hoger percentage zittenblijvers, in lagere eindexamencijfers, etc.

Het pleidooi van Van der Werf en Creemers

voor het gebruik van de leerjarenladder bij prestatiemeting van scholen lijdt aan ditzelfde conceptuele bezwaar tegen longitudinale data. Bosker en Scheerens wijzen daarom terecht op de adder van schoolwisselingen: welke school krijgt dan de 'credits'? De analyse van Veenstra et al. (1998) is een goed voorbeeld van dit probleem. Zij hebben in hun analyse ook de leerjarenladder gebruikt en bovendien hebben zij leerlingen ook na schoolwisselingen gevolgd. Dat betekent dat de plaats op de onderwijsladder na acht jaar hun criterium was, ongeacht op welke scholen onderwijs was gevolgd. Door deze aanpak krijgen bij Veenstra et al. (1998) scholen met veel ongediplomeerde schoolverlaters toch de 'credits' voor de verdere schoolloopbaan die alsnog afgelegd is op andere scholen. Dit betekent bijvoorbeeld dat de verdere schoolloopbaan van een VWO-leerling op het HAVO door hen toch op conto van de VWO-school wordt geschreven. Door dit gebruik van de leerjarenladder wordt het verschil tussen de uitkomsten met een longitudinale aanpak met multi-levelanalyse en mijn aanpak onnodig vergroot⁹. Als men de leerjarenladder zou willen gebruiken, zou hij 'stilgezet' moeten worden wanneer een leerling de betrokken school verlaat. De (onderwijs)loopbaan die de leerling na het verlaten van de betrokken school alsnog volgt, is immers het extern rendement van de school. Van der Werf en Creemers hebben duidelijk niet de consequenties doordacht van hun pleidooi voor de leerjarenladder.

5 Differentiële effecten

Terecht wijzen Bosker en Scheerens op het belang van differentiële opbrengsten van scholen: wat geldt voor de gemiddelde leerling op een school, hoeft niet te op te gaan voor een specifieke leerling. Dat is niets nieuws: bij elke *Trouw*-publicatie is dit voorbehoud gemaakt. Dat betekent dat ouders bij de schoolkeuze niet alleen moeten kijken naar de kwaliteit van een school voor de gemiddelde leerling, maar dat die ouders ook de specifieke kenmerken van hun kind daarbij moeten betrekken. Maar het belang van die differentiële opbrengsten moet niet overdreven worden. Inderdaad past een model zonder differentiële opbrengsten slech-

ter bij de verschillen in loopbanen dan een model met differentiële opbrengsten. Maar de verbetering in de passing van het model is niet zo ingrijpend¹⁰ dat daarmee resultaten zonder differentiële opbrengsten misleidend zouden zijn. Dat blijkt ook als men Figuur 1 uit het artikel van Bosker en Scheerens nauwkeurig beziet. Slechts bij enkele scholen wijkt de regressielijn ernstig af van die van de meeste andere scholen. Het sterk benadrukken van het belang van differentiële opbrengsten brengt daarom te veel scholen in de waan dat zij, ondanks een laag intern rendement, het voor specifieke leerlingen wel goed zouden doen.

6 Technische kritiek

Bosker en Scheerens verwijten mij dat ik in de tweede *Trouw*-publicatie (Agerbeek, 1998; Dronkers, 1998c) invalide gegevens over de ongediplomeerde uitstroom van scholen heb gebruikt, omdat bij deze uitstroomgegevens schoolwisseling vanwege ongediplomeerde uit- of afstroom niet onderscheiden kan worden van schoolwisseling vanwege verhuizing. Maar het feit dat dit onderscheid niet gemaakt kan worden is alleen ernstig wanneer scholen onderling sterk verschillen in het percentage schoolwisselingen vanwege verhuizen. Het is onwaarschijnlijk dat dit voor de reguliere leerlingen het geval is. En indien scholen toch sterk zouden verschillen in het verhuisgedrag van hun reguliere leerlingen, wordt hiervoor in de tweede *Trouw*-publicatie gecorrigeerd door middel van het percentage zij-instromers¹¹. Daarom meen ik dat de feitelijke uitstroomverschillen tussen scholen een goede indicatie zijn voor de omvang van de schoolwisselingen vanwege ongediplomeerde af- of uitstroom. Bovendien door rekening te houden met deze ongediplomeerde uitstroom, voorkom ik dat selectieve scholen waar een hoge af- of uitstroom plaatsvindt onterecht bevoordeeld worden op hun hoge eindexamencijfers of slaagpercentages¹². Met mijn keuze voor het kleinere kwaad van de aanname van gelijk verhuisgedrag op alle scholen, heb ik het grotere kwaad van het onterecht belonen van strenge selectie door scholen voorkomen. Ook hier blijft het betere de vijand van het goede. Bovendien, als Bosker en Scheerens toch gelijk

zouden hebben dat de ongediplomeerde uitstroom invalide gemeten is, hoe verklaren zij dan dat de gemiddelde uitstroom in de factoranalyses, weergegeven in Tabel 1, de te verwachten lading en richting heeft?

In de beide *Trouw*-publicaties zijn de verschillende sectoren in VBO helaas niet nader onderscheiden omdat de noodzakelijke gegevens over het intern rendement niet allemaal beschikbaar waren over de afzonderlijke sectoren in het VBO.

Van der Werf en Creemers menen dat ouders zelf wel alle gegevens zullen kunnen verwerken en dat een samenvattende kwaliteitsmaat, naast de bruto kengetallen, niet nodig is. Een vergelijking tussen de tweede *Trouw*-publicatie en de kwaliteitskaart van de inspectie laat duidelijk zien dat een samenvattende kwaliteitsmaat geen overbodige luxe is.

7 De analyse van Van der Werf en Creemers

De lezers van de twee *Trouw*-publicaties, de eerdere discussie in *Pedagogische Studiën* en de bijdrage van Bosker en Scheerens zullen reeds begrepen hebben dat de analyse van Van der Werf en Creemers met de data van het VOCL-cohort op te veel misverstanden en foute veronderstellingen rust om een betrouwbare validering te zijn van mijn samenvattende maat voor het intern rendement van scholen. Daarom som ik de twee belangrijkste maar kort op.

1. Van der Werf en Creemers gebruiken als indicator voor kwaliteit een score op twee toetsen (Nederlands en wiskunde), afgenomen halverwege het voortgezet onderwijs. Ik gebruik de cijfers van het centraal schriftelijk examen als een van de indicatoren van het intern rendement, naast zittenblijven, uitstroom en slagen. In Tabel 1 laat ik zien dat deze indicatoren goed samenhangen, maar dat ze elk wel een eigen betekenis hebben. Dat is ook precies de reden, waarom het niet nodig is dat mijn samenvattende maat voor intern rendement hoog samenhangt met het eindexamencijfer. De indicatoren van Van der Werf en Creemers zijn daarom veel eenzijdiger dan mijn samenvattende maat voor het intern rendement.

2. De twee toetsen die Van der Werf en

Creemers gebruiken zijn alleen afgenomen bij die leerlingen die de eerste drie jaar in het voortgezet onderwijs niet zijn blijven zitten, die niet ongediplomeerd uitgestroomd zijn, of die niet afgestroomd zijn naar een lager schooltype. Met andere woorden, de toetsen meten alleen de prestaties bij de succesvolle leerlingen. Mijn samenvattende maat voor intern rendement neemt de uitgevallen leerlingen en de zittenblijvers juist wel mee. In feite belonen Van der Werf en Creemers door hun twee indicatoren scholen met een hoge uitval en veel zittenblijvers, omdat zij alleen de prestaties van de overgebleven leerlingen meten. Daarom is het niet verbazingwekkend maar eerder geruststellend dat mijn maat nauwelijks samenhangt met die van Van der Werf en Creemers.

Noten

- 1 Ik ben mij ervan bewust dat in de jaren zeventig en tachtig vaak anders werd gehandeld, maar dat was niet op een wijziging in de wettelijke taken van de inspectie gebaseerd.
- 2 Een goed voorbeeld van deze behoefte van schoolbestuurders aan vergelijkende informatie over hun scholen is de rechtszaak van een ouder Mevr. Schaapman tegen de gemeente Amsterdam over het slecht functioneren van een openbare basisschool in die stad. De gemeente Amsterdam verloor dit proces terecht, onder meer omdat zij het functioneren van deze school door gebrek aan gegevens niet kon vergelijken met andere basisscholen in Amsterdam.
- 3 Bosker is bij de opzet en inrichting van de kwaliteitskaart van de inspectie betrokken.
- 4 De Onderwijsraad meent in haar advies over de kwaliteitskaart dat deze drie kwaliteitsdimensies niet onder 'de wettelijke deugdelijkheidseisen' vallen en dus niet in een verplichte kwaliteitsmeting thuishoren. Ten aanzien van sociaal-culturele vorming en maatschappelijke integratie heeft de Onderwijsraad gelijk. Maar omdat de gelijkwaardigheid van de eindexamens wel een deugdelijkheidseis is, de overheid die gelijkwaardigheid maar gedeeltelijk via het centraal schriftelijk deelt en het schoolonderzoek over scholen niet gelijkwaardig beoordeeld wordt (Pijl, 1991; Dronkers, 1998a), mag het extern rendement wel tot een deugdelijkheidseis gerekend worden.

- 5 De 'eigen values' van de tweede factoren lagen vlak bij of onder de 1.00.
- 6 Bosker et al. (1989) gebruikten als indicatoren slaagpercentages, afstroom, drop-out, double-ren, herexamens en schoolonderzoeken.
- 7 Ook de kwaliteitskaart van de inspectie gebruikt soortgelijke achtergrondkenmerken van leerlingen als indicator van het aanvangsniveau.
- 8 Zes jaar is een onderschatting, omdat Bosker en Scheerens het zittenblijven vergeten te verdisconteren.
- 9 Nog beter is een gemiddelde over de laatste drie jaar om toevallige schommelingen te vermijden. Van der Werf en Creemers overdrijven overigens de instabiliteit van de opbrengsten over schooljaren (Dronkers, 1998a).
- 10 Omdat de mogelijkheden voor een verdere schoolloopbaan van een ongediplomeerde VWO-schoolverlater groter zijn (HAVO, MAVO, VBO) dan voor een ongediplomeerde MAVO- of VBO-schoolverlater zijn de discrepanties tussen de uitkomsten van Veenstra et al. (1998), die de leerjarenladder na acht jaar gebruikten, en van de eerste *Trouw*-publicatie het grootst voor het VWO en het kleinst voor de MAVO en het VBO.
- 11 Bosker deelde dit desgevraagd mee op een studiebijeenkomst van de Sociaal-Wetenschappelijk Raad op 9 mei 1998 over de *Trouw*-publicaties.
- 12 Zij-instromers omvatten opstroomers na behalen van een lager diploma, ongediplomeerde afstroomers uit een hoger onderwijstype, binnenlandse verhuizingen en overplaatsingen tussen scholen. Het intern rendement van scholen met veel zij-instromers werd daarvoor gecorrigeerd (Agerbeek, 1998; Dronkers, 1998c).
- 13 Doordat de kwaliteitskaart geen gegevens bevat over deze ongediplomeerde uitstroom, belooft zij onterecht wel selectieve scholen.

Literatuur

- Agerbeek, M. (1998). Schoolprestaties 1998. *Trouw*, 13 juni, ZENZ.
- Agerbeek, M., Hageman, E., Kreulen, E., & Lakmaker, H. (1997). *Trouw*-onderzoek schoolprestaties. *Trouw*, 25 oktober, ZENZ.
- Archer, M.S. (1984). *Social origins of educational systems*. London/Beverly Hills: Sage.
- Bosker, R.J., Guldmond, H., Hofman, R.H., & Hofman, W.H.A. (1989). *Kwaliteit in het voortgezet*

onderwijs. Groningen: Instituut voor Onderwijs-onderzoek, Rijksuniversiteit Groningen.

- Dijkstra, A. B., Dronkers, J., & Hofman, R. (1997). *Verzuiling in het onderwijs. Actuele verklaringen en analyse*. Groningen: Wolters-Noordhoff.
- Dronkers, J. (1998a). Het Trouw-rapportcijfer van scholen in het voortgezet onderwijs; een analyse van de inspectiegegevens over de schooljaren 1995/96 en 1996/97. *Tijdschrift voor Onderwijs-research*, 23, 159-176.
- Dronkers, J. (1998b). Het betere is de vijand van het goede. Een reactie op de commentaren over het Trouw rapportcijfer. *Pedagogische Studiën*, 75, 142-150.
- Dronkers, J. (1998c). *Verschildt de toegevoegde waarde van het intern rendement van scholen in het voortgezet onderwijs? Enkele kanttekeningen bij het Trouw-onderzoek Schoolprestaties 1998*. Samenvatting van een openbaar college gegeven op 19 juni 1998. Amsterdam: SCO-Kohnstamm Instituut.
- Dronkers, J., Erp, M. van, Robijn, M., & Roeleveld, J. (1998). Krijgen leerlingen in de grote steden en met name in Amsterdam te hoge adviezen? *Tijdschrift voor Onderwijsresearch*, 23, 17-30.
- Pijl, Y.J. (1992). *Discrepanties tussen het school-onderzoek en het Centraal Schriftelijk Eindexamen*. Groningen: RION.
- Veenstra, D.R., Dijkstra, A.B., Peschar, J.L., & Snijders, T.A.B. (1998). Scholen op rapport. Een reactie op het Trouw-onderzoek naar schoolprestaties. *Pedagogische Studiën*, 75, 121-134.

Auteur

J. Dronkers is hoogleraar onderwijskunde (in het bijzonder stratificatie en mobiliteitsvraagstukken) aan de Faculteit der Pedagogische en Onderwijskundige Wetenschappen van de Universiteit van Amsterdam.

Adres: Universiteit van Amsterdam, Wibautstraat 4, 1091 GM Amsterdam

Abstract

The better is still the enemy of the good. An answer to Bosker & Scheerens and Van der Werf & Creemers

J. Dronkers. *Pedagogische Studiën*, 1999, 76, 74-82.

In this contribution the critical points of Bosker & Scheerens and Van der Werf & Creemers on my computation of the added value of the internal performance of secondary Dutch schools, made for the Dutch paper *Trouw* is answered. Six aspects are discussed: the goals of publishing inspection data on school performance; the different dimensions of performance of schools; the meaning and computation of added value; longitudinal versus cross-sectional measurement of internal performance; the importance of differential performances of schools; technical points. The conclusion of this contribution is that my approach is still the best solution to a number of problems around publishing inspection data on school performance.

Boekbespreking

M. Gemmeke

Politieke betrokkenheid van kinderen op de basisschool

Academisch proefschrift Universiteit van Amsterdam

Thesis Publishers, 1998, 282 pagina's

ISBN 90 9011 668 0

In dit proefschrift beschrijft de politicologe Mireille Gemmeke de resultaten van haar onderzoek naar de ontwikkeling van politieke betrokkenheid bij leerlingen in de basisschoolleeftijd (7 t/m 13 jaar). De belangrijkste vragen waren hoe die betrokkenheid zich ontwikkelt, en welke factoren daarop van invloed zijn.

Het thema politieke socialisatie lijkt terug van weggeweest. In de jaren vijftig en zestig deden vooral Amerikaanse politicologen en sociologen onderzoek naar politieke opvoeding, vaak vanuit de gedachte dat de voor de democratie noodzakelijke burgerschapskennis en burgerschapshouding al in de vroege jeugd gedetermineerd worden. Deze 'persistence'-opvatting verloor vanaf de jaren zeventig geleidelijk terrein, eerst aan onderzoekers die ervan uitgingen dat politieke opvattingen en betrokkenheid juist tijdens de adolescentie beïnvloed worden (*impressionable years*), en weer later, vanaf de jaren tachtig, aan de idee dat politieke beïnvloeding eigenlijk nooit stopt (*lifelong openness*). Met die verschuivingen is het onderzoek naar politieke socialisatie van jongere kinderen steeds meer in de wetenschappelijke marge terecht gekomen. In de eerste plaats nam binnen de psychologische theorievorming het geloof in het doorslaggevend belang van vroege ervaringen af. In de tweede plaats bleek de stabiliteit van politieke houdingen die in de kindertijd worden verworven empirisch gering. En in de derde plaats verloor het politieke socialisatie-onderzoek in zijn algemeenheid terrein doordat men geleidelijk aan ontdekte dat politieke oriëntaties (bijvoorbeeld het anti-Vietnamoorlog protest) niet slechts het resultaat zijn van opvoeding, maar ook van de politieke gebeurtenissen zelf. Wat zou dan de reden kunnen zijn om het onderzoek naar de politieke socialisatie van kinderen nieuw leven in te blazen? Gemmeke noemt er een aantal. Ze wijst op grote

methodologische tekortkomingen uit het vroegere onderzoek, op de grote hoeveelheid (politieke) informatie die tegenwoordig op kinderen is gericht en op de enorme aardverschuivingen die zich voordoen in het politieke landschap waarin kinderen van nu opgroeien: het ineenstorten van het communisme, de eenwording van Europa en het ontstaan van multiculturele samenlevingen. Daarmee is het probleem van de politieke betrokkenheid van karakter veranderd. Moderne democratieën hebben niet meer genoeg aan burgers die bereid zijn om eens in de zoveel tijd te gaan stemmen, maar kunnen slechts overleven wanneer die burgers bereid zijn zich actief in te zetten, wanneer zij beseffen dat zij zelf deel zijn van de politieke gemeenschap. Politieke betrokkenheid is daarmee een burgerschapsoriëntatie geworden. Deze wordt door Gemmeke omschreven als "het hebben van een binding met de politieke gemeenschap door kennisverwerving en belangstelling met betrekking tot allerlei aspecten van die politieke gemeenschap". Zulke bindingen verwijzen naar een fundamentele en algemene waardeoriëntatie die naar alle waarschijnlijkheid stabiel is dan een specifieke politieke voorkeur. Juist daarom is onderzoek naar burgerschapsvorming bij kinderen weer relevant, aldus de auteur.

Voor haar onderzoek onderscheidde Gemmeke drie aspecten van politieke betrokkenheid: een cognitieve component (kennis en bewustzijn van politieke problemen), een emotionele component (interesse, bezorgdheid) en een gedragsmatige component (naar het jeugdjournaal kijken, de krant lezen en dergelijke). Enkele honderden kinderen van Amsterdamse basisscholen, verdeeld over de groepen 4 t/m 8 werden ondervraagd over problemen als het milieu, oorlog, hongersnood, discriminatie, bezuinigingen en werkloosheid, over de wijze waarop ze daar kennis van nemen en over de vraag hoe elk van deze problemen hen persoonlijk raakt.

Juist het onderscheiden van de verschillende componenten van politieke betrokkenheid leverde interessante resultaten op. Kort gezegd: tussen groep 4 en groep 8 neemt de kennis over politieke figuren, instellingen en processen duidelijk toe. Ook vormen politieke problemen vaker onderwerp van gesprek en

wordt er actiever naar informatie gezocht. In termen van emotionele betrokkenheid verandert er in deze periode echter maar weinig. Zelfs de jongste groep die aan het onderzoek meedeed bleek al behoorlijk in verschillende politieke problemen geïnteresseerd te zijn. Het onderzoek bracht geen verschillen tussen meisjes en jongens aan het licht. Dat mag op zijn minst opvallend genoemd worden, want bij onderzoek naar de politieke betrokkenheid van volwassenen wordt vrijwel altijd een sterk gender-effect gevonden. Volgens de onderzoekster kan dit twee dingen betekenen: ofwel de politieke socialisatie is veranderd, (en wordt de slimme meid nu ook op de politiek voorbereid), ofwel de gender-differentiatie ontstaat pas op wat latere leeftijd, bijvoorbeeld in de adolescentie. Alleen longitudinaal onderzoek lijkt het antwoord op deze vraag te kunnen geven.

Misschien wel de belangrijkste bevinding van dit onderzoek is het aanzienlijke verschil tussen allochtone en autochtone kinderen in termen van hun politieke ontwikkeling. Zoals kon worden verwacht blijven allochtone kinderen op het cognitieve aspect achter bij hun autochtone leeftijdgenoten. De kenniskloof neemt echter niet af naarmate de (basis)schoolloopbaan vordert, maar groeit zelfs. Met de emotionele betrokkenheid ligt het anders: allochtone leerlingen zijn in dit opzicht sterker betrokken dan autochtone leerlingen, vooral bij onderwerpen die hen direct raken zoals werkloosheid en immigratie. Het onderwerp 'discriminatie' blijkt beide groepen even sterk bezig te houden. Uit een koppeling van leerkrachtkarakteristieken en de ontwikkelingsgegevens van de kinderen komt naar voren, dat het onderwijzend personeel wel degelijk invloed op de politieke betrokkenheid kan hebben. Hoe sterker onderwijzers zelf in politiek geïnteresseerd zijn, en hoe meer les over politiek zij geven, hoe beter hun leerlingen scoren op het cognitieve domein. De sfeer in de klas blijkt nog een bredere invloed te hebben: leerkrachten die prijs stellen op participerende, zelfstandige leerlingen in hun klas bevorderen in positieve zin de ontwikkeling van politieke betrokkenheid op elk van de drie onderscheiden domeinen. Overigens rapporteert de onderzoekster ook een 'redundancy-effect': de positieve invloed die leerkrachten

kunnen uitoefenen op de ontwikkeling van politieke kennis is aanzienlijk groter voor allochtone dan voor autochtone kinderen. Kennelijk wordt deze laatste groep in de thuis-situatie al zoveel geïnformeerd, dat de onderwijzer er nog maar weinig aan kan toevoegen.

Wat zijn van dit alles nu de (onderwijs-) pedagogische implicaties? Ondanks de grote emotionele betrokkenheid van allochtone leerlingen bij een aantal voor hen relevante politieke problemen, lijkt het erop of de kennisachterstand ten opzichte van autochtone leerlingen gedurende de basisschoolperiode alleen maar groter wordt. Als nu tegelijkertijd blijkt dat leerkrachten via de lesinhoud en de sfeer in de klas juist de politieke betrokkenheid van allochtone leerlingen kunnen versterken, ligt een pleidooi voor 'politieke bijscholing' voor deze groep – niet in termen van waarden, maar van kennis en inzicht – voor de hand. Men zou het 'inburgeringsles' op de basisschool kunnen noemen. Daarbij is echter een belangrijke kanttekening op zijn plaats. Gemmeke definieert in haar onderzoek politieke betrokkenheid als kennis, belangstelling en (informatiezoek)gedrag ten aanzien van de politieke gemeenschap. Maatschappelijke betrokkenheid, bijvoorbeeld ten aanzien van situaties die voor kinderen in hun eigen directe leefomgeving van belang zijn (buurt, school), valt daarmee buiten de boot. Juist over deze maatschappelijke participatie van kinderen is de laatste jaren veel geschreven. In veel gemeenten wordt, vaak met succes, geëxperimenteerd om ook allochtone kinderen actief te betrekken bij zaken als het inrichten van speelgelegenheden, het verhogen van de veiligheid in de buurt en de organisatie van vrijetijdsvoorzieningen. Mijns inziens is het onderscheid tussen politieke en maatschappelijke betrokkenheid dan ook conceptueel wel juist, maar vanuit een pedagogisch en sociaal oogpunt is het eerder contraproductief. Maatschappelijke participatie, zoals hiervoor omschreven, zou juist voor kinderen in achterstandssituaties wel eens een heel belangrijk middel en een belangrijke voorwaarde kunnen zijn om uiteindelijk het doel dat ook Gemmeke voor ogen heeft: participierend burgerschap, dichterbij te brengen. De speeltuin is voor de meeste kinderen een veel zichtbaarder politieke arena dan het

parlement in Den Haag. En een participatoire sfeer in de klas, een belangrijke oefenplaats voor maatschappelijke participatie, draagt bij aan de belangstelling voor de 'grote' politiek, zo blijkt ook weer uit dit onderzoek. Het had mijns inziens dan ook erg voor de hand gelegen om in dit ontwikkelingsgerichte onderzoek beide vormen van betrokkenheid in elkaars verlengde te bekijken.

Ik vind het erg lovenswaardig dat een politicologe het 'gat' dat pedagogen en psychologen de laatste jaren hebben laten vallen op het terrein van onderzoek naar sociaal-politieke vorming en opvoeding, heeft proberen op te vullen met dit belangwekkende en soms baanbrekende onderzoek. Het woord is nu weer aan de kinderwetenschappers, om de ontstane politicologische vertekening te compenseren.

M. de Winter

Mededelingen

ORD' 99

Op donderdag 20 en vrijdag 21 mei 1999 organiseren de Afdeling Onderwijskunde, het Instituut voor Toegepaste Sociale Wetenschappen (ITS), de Universitaire Lerarenopleiding (UNILO) en het Instituut voor Onderzoek van het Wetenschappelijk Onderwijs (IOWO) van de Katholieke Universiteit Nijmegen de 26ste Onderwijs Research Dagen. De ORD staan dit jaar speciaal in het teken van samenwerkend leren.

Alle informatie over deelname is te vinden op de ORD'99-website:

www.socsci.kun.nl/ped/owk/ord99.

Voor nadere informatie kan men zich wenden tot het congressecretariaat:

Katholieke Universiteit Nijmegen, Afdeling Onderwijskunde, bezoekadres: Montessorilaan 3, 4e verdieping, K. 15, telefoon: 024-3612585 Fax: 024-3615978, postadres: Postbus 9104, 6500 HE Nijmegen

Landelijke Pedagogendag 1999

Op zaterdag 27 maart 1999 wordt de negende Landelijke Pedagogendag 1999 georganiseerd in het 'Pieter de la Courtgebouw' van de Faculteit der Sociale Wetenschappen, Wassenaarseweg 52 te Leiden, met als thema 'Gezin, morele opvoeding en antisociaal gedrag'. De organisatie is in handen van de Afdeling Algemene en Gezinspedagogiek van de Universiteit Leiden in samenwerking met het Werkverband Wijsgerige en Historische Pedagogiek van de Nederlandse Vereniging van Pedagogogen en Onderwijskundigen.

De Pedagogendag is vooral bedoeld voor academisch geschoolden die beroepsmatig werkzaam zijn op het terrein van opvoeding, onderwijs, opleiding en vorming. Dit congres biedt hun een platform om eigen onderzoeksbevindingen te presenteren en met collega's hierover in discussie te treden.

Congressecretariaat:

Universiteit Leiden, Faculteit der Sociale Wetenschappen, Afdeling Algemene en Gezinspedagogiek, t.a.v. mevr. J.M.C. Scholten, Postbus 9555, 2300 RB Leiden
tel.: 071 - 5273434; fax: 071 - 5273945;

E-mail: Secrpel@Rulfsfsw.Fsw.Leidenuniv.nl
Verdere informatie op de volgende website:
<http://www.fsw.leidenuniv.nl/www/w3-pdalp/pedag.htm>

Afscheid hoogleraren

Afscheid Prof. de Ruyter

Op 16 april 1999 neemt Prof. dr. P.A. de Ruyter, hoogleraar orthopedagogiek van de Vrije Universiteit te Amsterdam afscheid. Bij die gelegenheid wordt een symposium georganiseerd over 'Het perspectief van de orthopedagogiek'. Deelname aan het symposium, inclusief het symposiumboek, kost f 125,-. Het symposium wordt gehouden in de Aula van het Hoofdgebouw van de VU, Boelelaan 1 105 te Amsterdam.

Voor informatie kunt U zich richten tot mevr. C. Logger, Vakgroep Pedagogiek VU, tel. 020 - 4448901; fax: 020 - 4448745; E-mail: j.logger@psy.vu.nl.

Afscheid Prof. dr. J.D. Imelman

Op 9 februari 1999, nam Prof. dr. J.D. Imelman afscheid als hoogleraar Pedagogiek aan de faculteit Sociale Wetenschappen te Utrecht. De titel van de afscheidsrede luidde: 'De afgebroken jaargang'.

Promoties

Universiteit Utrecht

Promovendus: Drs. P.W.J. Huistinx

Titel proefschrift: Milieu, sekse, etniciteit en schoolloopbanen. Een onderzoek onder jongeren aan het begin van de jaren negentig.

Promotor: prof. dr. W.J.H. Meeus

Co-promotor: dr. G.H. Maassen

Datum: 4 december 1998

Inhoud andere tijdschriften

Tijdschrift voor Orthopedagogiek

37e jaargang, nr. 9, 1998

Nauwkeurigheid L-HETS, door W.J.P.J.M. Steerneman, H.J. Pelzer en E.E.J. De Bruyn Orthopedagogiek. Vragen en Visies, door E. Broeckaert e.a. met een reactie door B. Levering.

LSOB in Zuid-Afrika, door P. Du Toit

Tijdschrift voor Orthopedagogiek
37e jaargang, nr. 10, 1998

Dominant gedrag bij jonge kinderen, door J. Bokhorst, F.A. Goossens, P.H. Dekker en P.A. de Ruyter
Arbeidsatisfactie groepsleiders, door J.D. van der Ploeg en E.M. Scholte
Hechtingsstoornissen, door N. Polderman

Tijdschrift voor Orthopedagogiek
37e jaargang, nr. 11, 1998

Diagnostische criteria Regionale Verwijzings Commissie, door M.J.G. de Wit
De Nijmeegse observatieschaal voor Kleuters, door M. Huizing, R. Van der Krol en H. Rost
Protocol toegangsfuncties: coördinatie van crisisinterventie, door L.T. van der Linden en E.C. Mellink

Pedagogisch Tijdschrift
23e jaargang, nr. 1, 1998

Leerkrachtengroepen en een projectmatige aanpak van onderwijsinnovatie, door A. Gheysen, J.C. Verhoeven, E. Van Assche en H. Van den Broeck
De professionele oriëntatie van docenten en hun betrokkenheid bij onderwijsvernieuwingen, door C.T. Jongmans, H.J.A. Biemans, P.J.C. Slegers en F.P.C.M. de Jong
'Leren leren', vereist consistentie in de leerstof, door A.T.J. Vernooij
Zuilvervorming en katholiek onderwijs. Een vergelijkend historisch-sociologische studie, door R. Verstraeten.

Pedagogisch Tijdschrift
23e jaargang, nr. 2/3, 1998

Themanummer: Onderwijsvernieuwingen in vakken in het voortgezet onderwijs, onder redactie van J. Terwel en M. Volman

Pedagogisch Tijdschrift
23e jaargang, nr. 4/5, 1998

Themanummer: Interpretatief onderzoek in pedagogiek en onderwijskunde, onder redactie van P. Smeyers en B. Levering

Pedagogisch Tijdschrift
23e jaargang, nr. 6, 1998

Een programma voor het coachen van beginnende leraren, door S. Veenman, H. De Laat en C. Staring
Straffen en opvoeden, door I. Weijers
The punishment of children, door B. Levering

Ontvangen boeken

- Akker, J. van den, Pieters, J., Visscher-Voerman, I., Wald, A. (Red.). *Studiehuis en onderwijsonderzoek*. Uitgeverij Garant, Apeldoorn, 1998, f 39,-.
- Busato, V.V., *Leerstijlen nader geanalyseerd* (dissertatie). Universiteit van Amsterdam, 1998.
- Bouwmeester, M., Deković, M., & Groenendaal, H., *Opvoeding in Somalische vluchtelinggezinnen in Nederland*. Van Gorcum, Assen, 1998, f 29,50.
- Dillemans, R., Lowyck, J., Peer, G. Van der, laeys, C., Elen, J., *New technologies for learning: contribution of ICT to innovation in education*. University Press, Leuven, 1998, Bfrs. 1.450,-.
- Geense, P., & Pels, T., *Opvoeding in Chinese gezinnen in Nederland*. Van Gorcum, Assen, 1998, f 45,-.
- Hoogeveen, M., & Bonset, H., *Het schoolvak Nederlands onderzocht*. Uitgeverij Garant, Apeldoorn, 1998, f 49,90.
- Klaassen C., & Leeferink, H., *Partners in opvoeding in het basisonderwijs*. Van Gorcum, Assen, 1998, f 29,50.
- Koops, W., & Slot, W. (Red.), *Van lastig tot misdadig*. Bohn Stafleu Van Loghum, Houten, 1998, f 52,-.
- Lazonder, A.W., *Aspecten van flexibiliteit in de kwalificatiestructuur*. Leerstoelgroep Agrarische Onderwijskunde, Wageningen, 1998.
- Lens, W., & Depreeuw, E., *Studiemotivatie en faalangst nader bekeken*. Universitaire Pers, Leuven, 1998, Bfrs. 595,-.
- Meijer, P., *Teachers' practical knowledge. Teaching reading comprehension in secondary education* (dissertatie). Universiteit Leiden, 1998.
- Nijsten, C., *Opvoeding in Turkse gezinnen in Nederland*. Van Gorcum, Assen, 1998, f 45,-.
- Rupp, J.C.C., *Van oude en nieuwe universiteiten*. Sdu Uitgevers, Den Haag, 1997.
- Smaling, A., & Hijmans, E., *Kwalitatief onderzoek en levensbeschouwing*. Uitgeverij Boom, Meppel, 1998, f 41,50.
- Veugelers, W., en Kat, E. de, *Opvoeden in het voortgezet onderwijs*. Van Gorcum, Assen, 1998, f 37,50.

In 1998 werd redactionele medewerking verleend door:

Dr. C.A.J. Aarnoutse Dr. (K.U. Nijmegen)	Prof.dr. N.A.J. Lagerweij (U.U. Utrecht)
Prof.dr. van den Akker (U.T. Enschede)	Prof.dr. J.M.G. Leune (E.U. Rotterdam)
Prof.dr. T.Bergen (K.U. Nijmegen)	Prof.dr. J.Lowyck (K.U. Leuven)
Dr. Th. Boland (SLO, Enschede)	Prof.dr. E.C.H. Marx (Em. hoogleraar, Leiden)
Prof.dr. R.J. Bosker (U.T. Enschede)	Dr. J. Onstenk (U.v.A Amsterdam)
Dr. A.G. Bus (U.L. Leiden)	Dr. G.J. Reezigt (R.U. Groningen)
Prof.dr. B.P.M. Creemers (GION Groningen)	Prof.dr. P. Reitsma (V.U. Amsterdam)
Prof.dr.J.R.M. Gerris (K.U. Nijmegen)	Prof.dr. A.M.M.J. Ruijsenaars (U.L. Leiden)
Prof.dr. P. Ghesquière (K.U. Leuven)	Prof.dr. P.Sleegers (K.U. Nijmegen)
Prof.dr. P. Janssen (K.U. Leuven)	Prof.dr. T.A.B. Snijders (R.U. Groningen)
Dr. P. de Jong (V.U. Amsterdam)	Dr. J.Vermunt (U.L. Leiden)
Dr. K.P.E. Gravemeijer, (U.U. Utrecht)	Dr. M. Voeten (K.U. Nijmegen)
Prof.dr. B. van Hout-Wolters (U.v.A. Amsterdam)	Prof.dr. T.H.A. van der Voort (U.L. Leiden)
Prof.dr. G.Kanselaar (U.U. Utrecht)	Dr. W. van Werkhoven (U.U. Utrecht)
Prof.dr. F.K. Kieviet (Em. hoogleraar Leiden)	Dr. B. Witziers (U.T. Enschede)

Van de redactie

Afscheid professor N. Verloop

Op 1 april 1999 heeft Prof. N. Verloop de functie van voorzitter van *Pedagogische Studiën* neergelegd. In naam van de hele redactie en – in ruimere zin – van alle auteurs en lezers van *Pedagogische Studiën*, bedank ik Prof. Verloop voor wat hij de voorbije 15 jaar voor het tijdschrift gedaan en betekend heeft.

In de 75-jarige geschiedenis van het tijdschrift zijn er weinigen die gedurende zo'n lange periode tijd en in zoveel verschillende hoedanigheden hun stempel hebben gedrukt op het tijdschrift. In 1985 begon Prof. Verloop bij *Pedagogische Studiën* als redactiesecretaris en hij bleef dit tot in 1991. Eerst met Prof. F. Kieviet, daarna met Prof. J. Lowyck als voorzitter. In 1991 ruilde hij de functie van secretaris in voor die van gewoon redactieraadslid, en begin 1995 volgde hij Prof. W. Nijhof op als voorzitter.

Gedurende de eerste periode bij *Pedagogische Studiën* ontpopte Prof. Verloop zich tot een efficiënt en begeesterd secretaris: hij stelde nieuwe beoordelingsformulieren voor reviewers op, schreef richtlijnen voor auteurs, bedacht systemen om de beoordelingstijd van manuscripten in te korten, ontwikkelde een duidelijke procedure voor de aanmaak van themanummers, enzovoort.

Ook in de periode als gewoon redactielid legde hij een meer dan modale redactionele activiteit aan de dag. Naast het indienen van talloze beoordelingsrapporten en het redigeren van menig themanummer of kroniek, porde hij ook vaak zelf potentiële auteurs aan tot het indienen van een manuscript, bijvoorbeeld naar aanleiding van zijn aanwezigheid op een internationaal of nationaal congres, een studiedag, een promotie, een inauguratie of een afscheidscollege.

Het wekte dan ook weinig verbazing dat Prof. Verloop in 1995 gekozen werd als opvolger van Prof. Nijhof, een voorzittersfunctie die hij de voorbije vier jaar op voorbeeldige wijze heeft vervuld. Tijdens die periode slaagde hij erin om de positie van *Pedagogische Studiën* als kwalitatief hoogstaand Nederlandstalig tijdschrift te handhaven in een klimaat dat – wat het onderzoek betreft – in toenemende mate onder druk komt te staan van de internationali-

sering en de verscherpte internationale publicatie-eisen, en – wat de uitgeverij betreft – steeds sterker te kampen krijgt met criteria van de financiële (over)leefbaarheid. Hoewel *Pedagogische Studiën* onder zijn voorzitterschap verder evolueerde in de richting van een empirisch georiënteerd onderwijsblad voor academici (Depaepe & Bakker, 1998), nam en steunde Prof. Verloop diverse initiatieven om het tijdschrift “een gemeenschappelijke studeerkamer” (Depaepe & Bakker, 1998, p. 9) te laten zijn voor onderwijskundigen en pedagogen, uit de empirisch-analytische en de hermeneutische onderzoekstraditie, meer fundamenteel georiënteerd en meer toepassingsgericht. Als voorzitter ving hij de problemen op van de vervanging van de onvervangbare mevrouw Hamburg, de opvolging van de vorige redactiesecretaris (Dr. J. Stijnen) en van verscheidene andere belangrijke redactieraadsliden. En “last but not least” nam hij het voortouw in de organisatie van het geslaagde 75-jarig jubileum van het oudste pedagogisch tijdschrift in Nederland, dat enkele maanden geleden in aanwezigheid van vele lezers van *Pedagogische Studiën* in het Utrechtse Congresgebouw werd gevierd, en in de redactie van de fraaie feestbundel *75 jaar onderwijs en opvoeding - 75 jaar Pedagogische Studiën*, dat naar aanleiding van dit jubileum als laatste aflevering van de vorige jaargang is verschenen (Verloop, 1998).

Met Prof. Verloop verdwijnt opnieuw een steunpilaar van *Pedagogische Studiën*. In naam van de hele redactieploeg en van al wie het tijdschrift een warm hart toedraagt, wens ik Prof. Verloop het allerbeste toe in zijn toekomstige professionele bezigheden.

Lieven Verschaffel

Depaepe, M., & Bakker, N. (1998). Een gemeenschappelijke studeerkamer: 75 jaar *Pedagogische Studiën*. In N. Verloop (Red.), *75 jaar onderwijs en opvoeding - 75 jaar Pedagogische Studiën* (pp. 9-44). Groningen: Wolters-Noordhoff.

Verloop, N. (Red.) (1998). *75 jaar onderwijs en opvoeding - 75 jaar Pedagogische Studiën*. Groningen: Wolters-Noordhoff.

De standaardmethode voorbij

Introductie op het thema

J.J. Hox

Eigenlijk zouden sociaal-wetenschappelijke onderzoekers, naast hun eigen vakgebied, ook nog moeten proberen de methodologische en statistische vernieuwingen bij te houden. In de praktijk is dat echter ondoenlijk. Ter illustratie: aan de Erasmus Universiteit te Rotterdam wordt door het expertisecentrum Social Research Methodology (SRM) de sociaal-wetenschappelijke methodologische literatuur geïnventariseerd en gecatalogiseerd. Het resultaat is een CD-ROM waaraan jaarlijks ongeveer 3000 publicaties worden toegevoegd. Zelfs wanneer maar tien procent hiervan de moeite waard is, dan nog moet men per jaar 300 publicaties lezen om methodologisch bij te blijven.

Ervaren onderzoekers passen daarom in de praktijk methoden en technieken toe die omschreven kunnen worden met de term 'accepted good practice' (Hox, 1997), dat wil zeggen, methoden en technieken die al jaren in gebruik zijn en waarvan methodologen en onderzoekers inmiddels de sterke en zwakke punten kennen. Dikwijls zijn het ook de methoden en technieken die de kern van het universitaire onderwijs vormen, en waarvoor goede software beschikbaar is. En, heel belangrijk: ze zijn ook bij tijdschriftredacties en reviewers bekend en geaccepteerd. Door je primair van deze technieken te bedienen, kun je je als auteur eigenlijk geen buil vallen.

Mijns inziens zou in kwalitatief hoogwaardig universitair onderzoek zoveel mogelijk gebruik moeten worden gemaakt van 'current best methods'. Current best methods zijn die methoden en technieken, die bij de methodologen goed bekend zijn en waarvan de methodologen zich inmiddels een redelijk beeld van de sterke en zwakke punten hebben gevormd. De omschrijving 'current best methods' is afkomstig uit de managementfilosofie, waarin gestreefd wordt naar een voortdurende bewaking en verbetering van de kwaliteit van het geleverde product. Dat gebeurt in het bijzonder

door het controleren en verbeteren van het productieproces (Deming, 1982). Current best methods spelen bij die kwaliteitsverbetering een belangrijke rol, omdat ze vastleggen en documenteren wat op dit moment de beste manier is om bepaalde doelen te bereiken. In het sociaal-wetenschappelijk onderzoek moeten we met current best methods denken aan methoden en technieken die in principe een verbetering zijn van de geaccepteerde standaardpraktijken, maar die door inhoudelijk georiënteerde onderzoekers nog weinig worden toegepast.

Een belangrijk aspect van current best methods is dat ze *current* zijn; ze omvatten de methoden die op het huidige moment de beste zijn. Daaruit spreekt de verwachting, dat ze op een gegeven moment vervangen zullen worden door nieuwe en betere methoden. Een voorbeeld van zo'n beweging is gemakkelijk te geven. Het gebruik van multiniveau-analyse was zo'n tien jaar geleden voorbehouden aan de experts; slechts enkele statistici en methodologen hielden zich hiermee bezig. Tegenwoordig vermoed ik dat een artikel dat een 'platte' analyse toepast op multiniveau-gegevens door gerenommeerde tijdschriften als het *Tijdschrift voor Onderwijsresearch* of *Pedagogische Studiën* niet meer geaccepteerd wordt. De 'huidige beste methode' is hard op weg om standaardmethode te worden.

Nadenken over de current best methods is daarbij niet alleen de taak van de inhoudelijke onderzoekers. Een belangrijk kenmerk van current best methods is dat ze *beschreven* zijn. Er moet gedocumenteerd zijn om welke methoden het gaat, en er moet uiteengezet worden waarom en hoe die toegepast moeten worden.

Dit nummer van *Pedagogische Studiën* bevat vier methodologische bijdragen. Elk van deze bijdragen beschrijft methoden of technieken die in principe zo toegepast kunnen worden in het onderwijskundig onderzoek. Het eerste

artikel, van De Leeuw, geeft een overzicht van het methodologisch onderzoek naar het verzamelen van gevoelige informatie. Methodologen en statistici hebben hier veel onderzoek naar gedaan en oplossingen voorgesteld, variërend van randomized response (Van der Heijden, 1994) tot computerondersteund interviewen. Veel van de technieken die De Leeuw bespreekt, zijn direct toepasbaar in onderwijskundig onderzoek. Van Hattum en Kef bespreken de toepassing van computerondersteunde interviewmethoden bij onderzoek in scholen en bij jongeren. Een niet te onderschatten voordeel daarvan blijkt te liggen in de datakwaliteit: er zijn goede redenen om aan te nemen dat computerondersteund interviewen leidt tot een betere kwaliteit van de verzamelde gegevens. Het artikel van Wester en Peters gaat over recente ontwikkelingen bij de analyse van kwalitatieve gegevens. Hoewel computerondersteunde kwalitatieve analyse bepaald niet nieuw meer is (vgl. Peters, Wester & Richardson 1989) komen we deze in de praktijk weinig tegen. Het gevolg zijn onderzoeksverslagen waarin alleen melding wordt gemaakt dat er 'een kwalitatieve analyse is uitgevoerd', zonder dat de verslagen technische details bevatten. Daarmee is het betreffende onderzoek in

feite onrepliceerbaar geworden. Het laatste artikel, van Hox, behandelt een aantal statistische simulatietechnieken die eveneens al een tijd beschikbaar zijn. Deze technieken zijn binnen het bereik gekomen van inhoudelijk georiënteerde onderzoekers dankzij de beschikbaarheid van snelle en goedkope computers en vooral dankzij het opnemen van dit soort technieken in standaard-computerprogrammatuur zoals SPSS.

Ik hoop, samen met de andere auteurs, dat lezing van de vier volgende artikelen voor u aanleiding zal zijn een aantal van deze technieken ook in uw eigen onderzoek toe te passen.

Literatuur

- Deming, W.E. (1982). *Quality, productivity, and competitive position*. Cambridge: MIT.
- Heijden, P.G.M. van der (1994). *Hoe te leven met statistiek*. Oratie, Universiteit Utrecht.
- Hox, J.J. (1997). *Er is nieuws onder de zon*. Oratie, Universiteit Utrecht.
- Peters, V., Wester, F. & Richardson, R. (1989). *Kwalitatieve analyse in de praktijk en Handleiding bij Kwalitan*. Nijmegen: ITS.

E. D. de Leeuw

Samenvatting

De validiteit van onderzoek naar een gevoelig onderwerp kan bedreigd worden door een grote mate van ontbrekende gegevens en sociaal-wenselijke antwoorden. Over het vragen naar sensitieve informatie is dan ook veel methodologisch onderzoek gedaan. Er zijn verschillende methoden om de vertekening in onderzoek bij gevoelige onderwerpen terug te dringen. Deze richten zich onder meer op het vergroten van de privacy bij het beantwoorden van de vragen, het aanpassen van de vraagvorm zodanig dat sociale wenselijkheidstendenties minder kans krijgen, het speciaal trainen en begeleiden van interviewers om optimaal te functioneren in moeilijke situaties en het verzamelen van additionele informatie uit andere bronnen. In dit artikel wordt een overzicht gegeven van onderzoeksmethoden voor gevoelige vragen, de achterliggende theorieën, en de voor- en nadelen van de verschillende methoden.

1 Introductie

Bij onderzoek naar gevoelige onderwerpen kunnen zich verschillende problemen voordoen: ondervraagden kunnen weigeren aan het gehele onderzoek mee te doen (unit-nonrespons), ondervraagden kunnen weigeren om aan bepaalde onderdelen van het onderzoek mee te doen (partiële nonrespons), ze kunnen weigeren op een bepaalde vraag antwoord te geven (item-nonrespons), of ze kunnen bij het beantwoorden van de vraag de werkelijkheid 'mooier' voorstellen (sociale wenselijkheid).

Hierdoor kan de validiteit van de onderzoeksconclusies bedreigd worden. Wanneer ondervraagden weigeren om alle vragen of een gedeelte van de vragen te beantwoorden, dan vormt dit een bedreiging voor de generaliseerbaarheid van de conclusies: de externe validiteit wordt bedreigd. Wanneer ondervraagden wel de vraag beantwoorden, maar deze niet geheel eerlijk beantwoorden, bijvoorbeeld door een sociaal wenselijk antwoord te geven, wordt met name de begripsvaliditeit bedreigd: de onderzoeker meet dan iets anders dan men wil meten.

Onderzoekers hebben verschillende methoden ontworpen om de mogelijke vertekening in onderzoek bij gevoelige onderwerpen terug te dringen en de validiteit van de conclusies te vergroten. Deze methoden richten zich onder meer op het vergroten van de privacy bij het beantwoorden van de vragen, het aanpassen van de vraagvorm zodanig dat sociale-wenselijkheidstendenties minder kans krijgen, het speciaal trainen en begeleiden van interviewers zodat deze optimaal functioneren in moeilijke situaties en het verzamelen van additionele informatie uit andere bronnen. In dit artikel zal ik op deze punten ingaan. Alvorens dit te doen geef ik eerst een overzicht van mogelijke gevoelige onderwerpen en van mechanismen die kunnen leiden tot vertekening van de conclusies in onderzoek naar deze onderwerpen.

2 Wat zijn gevoelige onderwerpen

Wanneer men aan een aantal willekeurige mensen vraagt wat nu een gevoelig onderwerp is bij enquêtes, dan noemt men vaak criminaliteit, druggebruik en seksueel gedrag. Andere voorbeelden zijn vragen naar gezondheid, risicogedrag en ziekte, alcoholgebruik, inkomen en financiële situatie. Dit zijn alle onderwerpen die voor de onderzochte potentieel bedreigend kunnen zijn. En daarmee raken we de kern van

* Met dank aan Joop Hox, Don Mellenbergh, Peter Swanborn en een anonieme reviewer voor hun commentaar en aan Jan-Huib Blans, Benoit Riandey en de leden van de AAPOR- en de ASA/SMRS-internet-discussielijsten voor literatuursuggesties.

het begrip 'gevoelig'. Een onderwerp is gevoelig wanneer het voor de onderzochte persoon potentieel bedreigend is of, in andere woorden, wanneer het beantwoorden sociale of emotionele 'kosten' met zich meebrengt. Een onderzochte zal alleen dan eerlijk antwoord geven wanneer hij of zij de verwachte 'kosten', bijvoorbeeld gevoelens van schaamte, op vindt wegen tegen de te verwachte 'baten', bijvoorbeeld het nut van het onderzoek voor de volksgezondheid (zie ook Dillman, 1978, over sociale ruiltheorie in de context van dataverzameling).

Onderzoeksvragen kunnen op twee manieren bedreigend voor respondenten zijn. Ze kunnen een intrinsieke of een extrinsieke dreiging vormen (Lee, 1993).

2.1 Intrinsieke dreiging

Er is sprake van een intrinsieke dreiging wanneer de vragen betrekking hebben op onderwerpen die zeer persoonlijk zijn, pijnlijk zijn, of stressvol. De dreiging komt dan vanuit de respondent zelf. Voorbeelden hiervan zijn vragen naar echtscheiding, of de dood van een ouder. Maar ook vragen naar positieve seksuele belevingen of eerste verliefdheid kunnen voor respondenten een dermate privé karakter hebben, dat men er niet over kan of wil praten. Ook vragen die respondenten dwingen om hun zelfbeeld negatief bij te stellen, zullen als zeer bedreigend worden ervaren.

Een goed voorbeeld van dit laatste type vragen wordt gegeven door Burt (1995). Zij vertelt dat daklozen vaak ontkennend antwoorden op de vraag of zij dakloos zijn. Bij doorvragen blijkt dan dat zij zichzelf niet als dakloze zien, maar als iemand die door omstandigheden tijdelijk verblijft in een opvanghuis of speciaal pension. Een ander duidelijk voorbeeld is toegeven dat men op school gepest wordt. Een kind zal dit niet gauw toegeven, noch tegenover anderen noch tegenover zichzelf.

Soms kan het voor respondenten pijnlijk of stressvol zijn om 'tegenover zichzelf' toe te geven dat zij zich niet gedragen zoals ze eigenlijk volgens zichzelf zouden moeten. Er is dan sprake van internalisatie van sociale normen. Bradburn en Sudman (1979) wijzen dan ook op het belang van de respondent om een positief zelfbeeld te handhaven. Dit kan op twee manieren, afhankelijk van het onderwerp. Wanneer

het onderwerp betrekking heeft op zeer gewenst of gewaardeerd gedrag, dan zal men niet graag toegeven dat men dit gedrag niet of haast niet vertoont. Zo zal men bijvoorbeeld de frequentie van tandenpoetsen vaak wat overdrijven; men handhaaft het positieve zelfbeeld van een reine, zichzelf goed verzorgende persoon. Wanneer het gedrag betrekking heeft op ongewenst gedrag dat tegen de normen ingaat, dan zal men dit gedrag ontkennen of minimaliseren. Zo zal alcoholconsumptie in het algemeen ondergerapporteerd worden. Men drinkt een à twee glaasjes en zal niet snel toegeven dat er meer gedronken wordt.

2.2 Extrinsieke bedreiging

Vragen naar ongewenst gedrag kunnen naast een intrinsieke dreiging ook een extrinsieke dreiging met zich meebrengen. Er is sprake van extrinsieke dreiging als de antwoorden sancties kunnen oproepen. Dit kunnen sociale sancties zijn, bijvoorbeeld groepsdruk, sociale controle, zoals in het geval van ongewenst gedrag, te veel drinken of het gebruiken van geweld tegen medeleerlingen. Wanneer er sprake is van strafbaar gedrag, dan kunnen de sancties bedreigend zijn omdat zij strafrechtelijke consequenties hebben. De kern van extrinsieke dreiging ligt in de angst van de respondent dat de antwoorden bekend worden of zelfs doorgegeven worden aan officiële instanties zoals de belastingen of de politie, of bij schoolonderzoek aan ouders of leerkrachten.

Een speciaal geval van extrinsieke dreiging gaat uit van vragen die een politieke of beleidsconsequentie kunnen hebben: Wat gebeurt er met de resultaten van het onderzoek, zouden deze bijvoorbeeld gebruikt kunnen worden om een (on)gewenste politieke beslissing te legitimeren? Zo moest een studie naar de gezondheid van tieners in de New Yorkse wijk Harlem halverwege afgebroken worden, omdat lokale actiegroepen uit angst voor ongewenste beleidsconsequenties opriepen tot een boycot (Lee, 1993, p. 7). Anticipatie op beleidseffecten is ook bij onderwijskundig onderzoek een bekend fenomeen.

Wat bedreigend is kan van cultuur tot cultuur en van tijdperk tot tijdperk verschillen. Vragen naar marihuana-gebruik kunnen in een kleine stad in het midden-westen van Amerika heel gevoelig zijn en in Amsterdam veel min-

der. Het toegeven van het hebben van pacifistische ideeën was in Londen tijdens de Eerste Wereldoorlog gevaarlijk, want men riskeerde niet alleen sociale uitsluiting maar zelfs geweld, terwijl het tijdens de grote vredesdemonstraties in de jaren tachtig in Den Haag zelfs bewonderenswaardig was dat men pacifistische ideeën uitdroeg. Er zijn ook minder extreme voorbeelden; zo is de vraag naar voltooide schoolopleiding in Zweden nogal gevoelig en is de vraag naar inkomen in Amerika één van de gevoeligste vragen (Lyberg & Dean, 1992).

In de nu volgende paragrafen worden systematisch methoden behandeld waarmee het mogelijk is om goed onderzoek te doen naar gevoelige onderwerpen. Welke methoden en technieken een onderzoeker dient te gebruiken, hangt af van de te onderzoeken groep en het onderwerp, en met name van het type dreiging dat voor die specifieke groep uitgaat van het te onderzoeken onderwerp.

3 Opvangen van extrinsieke en intrinsieke dreiging bij onderzoek naar gevoelige onderwerpen

3.1 Onderzoek naar gevoelige onderwerpen: extrinsieke dreiging

Wanneer ondervraagden bang zijn dat een eerlijk antwoord negatieve consequenties kan hebben, is er sprake van extrinsieke dreiging. Deze negatieve consequenties kunnen vrij mild zijn, zoals blijk van afkeuring door de directe omgeving. De gevreesde sancties kunnen echter ook ernstiger zijn. Voorbeelden zijn: sociale stigmatisering bij het bekend worden van ongewenst gedrag, kortingen op uitkeringen bij bekend worden van bijverdienden en strafrechtelijke vervolging bij erkennen van crimineel gedrag.

De onderzoeker dient in dit geval het onderzoek zodanig op te zetten en uit te voeren dat de extrinsieke dreiging minimaal is: de angst voor sancties en mogelijke negatieve consequenties dient weggenomen te worden. Met andere woorden, de respondent dient er van verzekerd te zijn dat de antwoorden volstrekt vertrouwelijk zijn en dat persoonlijk identifi-

ceerbare individuele gegevens niet openbaar worden gemaakt. Kernbegrip hierbij is vertrouwelijkheid van de gegevens. Twee verschillende processen spelen een rol bij het verzekeren van de vertrouwelijkheid.

Allereerst dienen, *voordat* de dataverzameling begint, alle mogelijke maatregelen genomen te worden die de vertrouwelijkheid van de gegevens tijdens en na het onderzoek zullen garanderen. Uiterst belangrijk zijn een onderzoeksprotocol (taken, verantwoordelijkheden, gedragscode), een uitgewerkte publicatiepolicy (*wat wordt hoe in welke vorm en waar gepubliceerd*) en een gedetailleerd draaiboek met maatregelen voor de dataopslag (anonymiseren, versleutelen steekproef, versleutelen van datafiles wanneer tijdens interviews gebruik wordt gemaakt van computergestuurd interviewen, link-filesystemen voor data-opslag, e.d.). Zie ook de lijst van aanbevelingen van de Commissie Kordes (1997), en de gedragscode van de VMO (1997).

Daarnaast is het belangrijk om bij het opzetten van het onderzoek voor iedere onderzoeks-fase expliciet na te gaan welke maatregelen de vertrouwelijkheid bevorderen en de extrinsieke dreiging wegnemen. Voorbeelden zijn: iedereen die bij de uitvoering van het onderzoek betrokken is, ook de interviewers, hoort een geheimhoudingsverklaring te tekenen; tijdens de interviewerinstructie wordt de interviewers verteld welke maatregelen er genomen zijn om de vertrouwelijkheid te beschermen. Het is niet de bedoeling dat de interviewers dit allemaal in detail aan de respondent vertellen, integendeel (zie hierna), maar deze kennis werkt wel (onbewust) door in de houding en het gedrag van de interviewers. De interviewers of dataverzamelaars weten zo dat ze aan een professioneel en zorgvuldig opgezet onderzoek meewerken, waarbij uitdrukkelijk aandacht aan de vertrouwelijkheid van de gegevens besteed is. Dit zullen ze ook op een positieve wijze uitstralen. Tevens kan men de dataverzameling zelf zodanig opzetten dat de extrinsieke dreiging geminimaliseerd wordt: bijvoorbeeld door de setting zo neutraal mogelijk te houden en iedere schijn van censuur of veroordeling te vermijden, of een gedeelte van de vragen op een speciale manier (schriftelijk met verzegelde enveloppen, randomized response) te laten beantwoorden (zie ook 3.2.).

Ten tweede, de onderzoeker dient de respondent te verzekeren dat de gegevens vertrouwelijk zijn. Hierboven is al beschreven hoe dat in het gedrag van de uitvoerders tot uiting kan komen, maar het moet respondenten ook expliciet, in woorden, gezegd worden. Men kan prachtige veiligheidsmaatregelen genomen hebben, maar wanneer de respondent niet weet dat de informatie vertrouwelijk is en nooit op een voor hem/haar bedreigende wijze gebruikt gaat worden, zullen deze veiligheidsmaatregelen geen effect hebben op het vraag-antwoordproces bij de respondent. Een expliciete verzekering van vertrouwelijkheid of anonimiteit is daarom noodzakelijk. Hierbij is het van belang dat dit in voor de respondent begrijpelijke taal gedaan wordt én dat het niet te veel nadruk krijgt. Een lange, uitgebreide verantwoording maakt respondenten eerder wantrouwig dan dat het geruststelt. Men gaat zich bezorgd afvragen wat er allemaal gevraagd gaat worden als zo expliciet de nadruk op de vertrouwelijkheid wordt gelegd. Vergelijkend onderzoek naar verschillende soorten introducties laat zien dat een korte verzekering van vertrouwelijkheid het beste werkt: in woord en gedrag wordt duidelijk gemaakt dat het vanzelfsprekend is dat de gegevens vertrouwelijk zijn (Hippler, Schwarz, & Singer, 1990).

3.2 Onderzoek naar gevoelige onderwerpen: intrinsieke dreiging

Er is sprake van intrinsieke dreiging wanneer gevraagd wordt naar onderwerpen die voor de respondent zeer persoonlijk, pijnlijk of stressvol zijn, of het positieve zelfbeeld van de respondent aantasten. De dreiging komt dan vanuit de respondent zelf. Voorbeelden zijn vragen over het gepest worden door medeleerlingen, verliefdheden, verkering of seks. De onderzoeker kan op twee manieren deze dreiging verminderen: door de methode van dataverzameling zo privé mogelijk te maken, en door de vraagverwoording aan te passen. Deze twee methoden sluiten elkaar natuurlijk niet uit, integendeel ze vullen elkaar aan en worden dan ook gelijktijdig gebruikt.

Allereerst gaan we in op de methode van *dataverzameling*. Uit vergelijkend onderzoek is gebleken dat men bij schriftelijke vragenlijsten opener en eerlijker antwoord geeft op gevoeli-

ge vragen. Voor een overzicht zie De Leeuw (1992). Wanneer het niet mogelijk is een schriftelijke vragenlijst of post-enquête te gebruiken, dient het interview zodanig te zijn ingericht dat dit voor de geïnterviewde zo min mogelijk stress geeft. Hierbij speelt de persoon van de interviewer een grote rol. Deze dient niet te schokken te zijn en in houding en gedrag vanzelfsprekend en zakelijk te reageren, zonder een koude indruk te maken. Kernbegrippen voor het interviewgedrag zijn 'niet te schokken', 'nuchter en matter of fact' en 'vriendelijk'. Ook dient een interviewer goed op de hoogte te zijn van het taalgebruik dat normaal is voor een specifieke doelgroep. Men denke bijvoorbeeld aan een onderzoek naar veilig vrijen onder homoseksuele jongeren. Op de rol van de interviewer wordt in 4.2. nader ingegaan.

Het spreekt haast vanzelf dat men bij een interview de aanwezigheid van derden dient te vermijden (vgl. b.v. Aquilino, 1993). Maar ook de interviewer zelf is een toehoorder, die de privacy kan aantasten. Daarom is het belangrijk om bij gevoelige vragen andere methoden binnen het interview toe te passen die de privacy van de respondent zo veel mogelijk garanderen. Men noemt dit wel een 'mixed-mode'-benadering. Een veel gebruikte en succesvolle methode is om tijdens het interview de respondent zelf een korte vragenlijst te laten invullen die deze daarna in een gesloten en dichtgeplakte envelop aan de interviewer meegeeft. Deze methode resulteert in meer accurate en opener antwoorden dan wanneer alle vragen door de interviewer gesteld worden (Makkai & McAlister, 1992). Bij zeer gevoelige vragen kan de interviewer de respondent verzoeken de vragenlijst in een bijgeleverde gefrankeerde envelop te stoppen en te posten. Wanneer er gebruik gemaakt wordt van computergestuurd interviewen, kan de interviewer de respondent een papieren vragenlijst laten invullen bij zeer gevoelige vragen, maar nog beter werkt het als de interviewer aan de respondent vraagt om zelf achter het toetsenbord plaats te nemen. Uit vergelijkend onderzoek blijkt dat een computer gestuurd zelfinterview zelfs nog iets meer openheid in de antwoorden oplevert dan een door de respondent zelf ingevulde papieren vragenlijst (De Leeuw, Hox, & Snijders, 1995; Weisband & Kiesler (1996); zie ook de bijdra-

ge van Van Hattum & Kef in dit nummer). In sommige gevallen kan een schriftelijke vragenlijst moeilijkheden opleveren, bijvoorbeeld bij heel jonge kinderen, ouderen, functioneel anal-fabeten, of visueel gehandicapten. Nieuwe ontwikkelingen om dit probleem op te lossen zijn het gebruik van een walkman in combinatie met een schriftelijke vragenlijst en het gebruik van Audio-CASI. De combinatie van walkman en schriftelijke vragenlijst is bij kinderen zeer succesvol (Scott, 1997). In deze toepassing horen de kinderen de vraag via de walkman en lezen tegelijk mee en kruisen hun antwoord aan in de vragenlijst. In Audio-CASI wordt de geluidskaart van de computer benut. De respondent hoort de vraag, ziet tegelijkertijd de tekst op het scherm en tikt het antwoord in. Het antwoord dient dan wel een simpel ja/nee te zijn of een getal dat eenvoudig is in te tikken. Voor een toepassing van Audio-CASI bij visueel gehandicapte jongeren zie De Leeuw, Hox, Kef & Van Hattum (1997) en Van Hattum en Kef in dit nummer. In de Verenigde Staten is Audio-CASI ondermeer succesvol toegepast bij onderzoek naar drugsgebruik en seksueel risico gedrag onder adolescenten (Turner, Ku, Rogers, Lindberg, Pleck & Sonenstein, 1998).

Bij zeer gevoelige onderwerpen wordt soms gebruik gemaakt van meer gecompliceerde indirecte technieken zoals de randomized response. Kenmerk van de randomized response methode is dat de respondent twee vragen voorgelegd krijgt. Op basis van een toevalsmechanisme (werpen van een munt, dobbelsteen) geeft de respondent antwoord op een van de twee vragen en slechts de respondent weet welke vraag beantwoord is. Bijvoorbeeld: wanneer een respondent met 'ja' antwoordt, noteert de interviewer het antwoord, maar weet niet of dit nu een 'ja' is op de vraag 'slaat u uw vrouw' of op de vraag 'bent u in maart geboren'. Via statistische technieken en het bekend zijn van de kansen op een antwoord kan men een goede schatting geven van het percentage dat 'ja' op de gevoelige vraag heeft geantwoord. Randomized response is in 1965 ontwikkeld door Warner. In de oorspronkelijke benadering leverde deze methode alleen univariate schattingen, het antwoord kon niet gerelateerd worden aan andere variabelen. In de tachtiger jaren zijn er echter methoden ontwikkeld waarbij de randomized-responsevariabele als

afhankelijke variabele in een logistische regressie gebruikt kan worden (zie b.v. Scheers & Dayton, 1988). Men kan de antwoorden op een randomized-responsevraag nu wel degelijk in verband brengen met antwoorden op andere vragen (Van der Heijden & Van Gils, 1996). Voor een algemeen overzicht over randomised response zie b.v. Umesh en Peterson (1991). In Nederland zijn randomized response technieken onder andere gebruikt door Van der Heijden, Van Gils, Bouts en Hox (1998). Voor het succesvol gebruiken van randomized response technieken is het van belang dat de methode helder en duidelijk aan de respondent wordt uitgelegd. De respondent moet de techniek goed kunnen toepassen, én er van overtuigd zijn dat deze ook inderdaad de vertrouwelijkheid bevordert.

Naast de methode van ondervragen kan de onderzoeker de vraagvorm zodanig aanpassen dat de intrinsieke dreiging van het gevoelige onderwerp wordt tegengegaan. In het algemeen dient de vraag zo verwoord te zijn dat elke schijn van afkeuring of censuur vermeden wordt. Door de verwoording van de vraag zelf of door het toevoegen van extreme antwoord-categorieën kan men daarenboven de suggestie geven dat het betreffende gedrag sociaal acceptabel is. Zo kan de onderzoeker in de vraagformulering de suggestie meegeven dat (bijna) iedereen het wel eens doet, bijvoorbeeld: 'Zelfs de meest kalme ouders ...'. Ook kan expliciet het gedrag verondersteld worden en naar de frequentie gevraagd worden ('Hoe vaak spijbel je?' in plaats van 'Spijbel je wel eens?'), of kan in de vraagvorm een beroep op een autoriteit gedaan worden om zo het sociaal onwenselijke karakter tegen te gaan ('Uit medisch onderzoek blijkt dat wijn drinken de kansen op hartproblemen vermindert, hoeveel wijn'). Men gebruikt dan *bewust* suggererende (leading) vragen, om sociale wenselijkheid tegen te gaan. In het algemeen dient men suggererende vragen natuurlijk te vermijden, en dient men vragen neutraal te stellen. Maar bij onderzoek naar sociaal erg onwenselijke zaken, kan het zeer welbewust 'laden' van een vraag een goed instrument zijn. (vgl. Sudman & Bradburn, 1982). Voor goede aanwijzingen voor het formuleren van vragen zie Fowler (1995). Een speciale techniek, die enige succes heeft

geboekt bij het terugdringen van missing data op inkomensvragen, is de zogenaamde 'door-in-the-face' techniek (Hippler, Schwarz & Singer, 1990). Men begint met een open vraag (b.v. 'Hoeveel verdient u?'). Aan de respondenten die hier geen antwoord op willen geven wordt dan een iets gematigder vervolgvraag gesteld; deze wordt verwoord als tegemoetkoming aan de respondent. Bijvoorbeeld: 'Ik begrijp dat het een lastige vraag is, mag ik u vragen verdient u meer of minder dan X', eventueel gevolgd door een gesloten vraag met enkele inkomenscategorieën. Opgemerkt dient te worden dat deze vraagwijze met name uitgetest is voor inkomensvragen en niet standaard toegepast kan worden bij andere gevoelige vragen.

Bij gevoelige onderwerpen is het aan te raden om de antwoordmogelijkheden van gesloten vragen zo ruim mogelijk te maken. Zeker wanneer de vraag ongewenst gedrag betreft. Daarnaast kan men bij een gesloten vraag er zorg voor te dragen dat er een aantal extreme antwoordcategorieën zijn. Deze extremen brengen dan aan de respondent de boodschap over dat de onderzoeker extreme uitschieters heel gewoon vindt en dat mensen dit blijkbaar antwoorden. Het verdient verder aanbeveling antwoordcategorieën *niet* te nummeren; cijfers kunnen ook een suggestie overbrengen, denk maar aan schoolcijfers.

Ter illustratie van de het gebruik van extreme antwoordcategorieën, vergelijk de volgende twee vragen naar alcoholconsumptie:

Hoeveel glazen alcohol drinkt U per dag?

- ☐ geen
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ meer

Hoeveel glazen alcohol drinkt U per dag?

- ☐ geen
- ☐ 1 à 2
- ☐ 3 à 4
- ☐ 5 à 6
- ☐ 7 à 8
- ☐ 8 à 10
- ☐ meer dan 10

illustratie. Hierbij was het doel het aantal valse negatieven tegen te gaan. Valse negatieven zijn mensen die het gedrag wel vertonen, maar het niet toe willen geven. Daarnaast is er ook het gevaar op valse positieven: mensen die pochen en een gedrag rapporteren dat ze zelf in werkelijkheid niet vertonen. Met name adolescenten die in een 'peer group' ondervraagd worden, zijn voor dit laatste gevoelig. Zowel bij het extreem formuleren van een vraag ('laden' van een vraag), als bij het toevoegen van extreme antwoordcategorieën, dient men voorzichtig te werk te gaan. Anders raakt men van de regen in de drup en verruilt men de ene fout (valse negatieven) voor de andere (valse positieven).

3.3 Samenvattend

Bij onderzoek naar gevoelige vragen is het een goede gewoonte om de respondenten vertrouwelijkheid van de gegevens of anonimiteit te garanderen. Daarnaast dient men een situatie te creëren waarin de respondent zich zo goed mogelijk voelt. Dit kan onder andere op de volgende wijze:

1. Leg in het kort aan de respondent uit waarom deze vragen gesteld worden (nut, doel).
2. Verwoord de vragen zodanig dat iedere schijn van afkeuring of censuur vermeden wordt (oordeel niet). Indien mogelijk, geef de suggestie dat het gedrag in kwestie sociaal acceptabel is (laden van vragen).
3. Maak antwoordcategorieën zo ruim mogelijk.
4. Breng de vraag zo zakelijk mogelijk (matter of fact).
5. Maak de dataverzamelmethode zo privé mogelijk (b.v. vragenlijst met gesloten envelop, CASI, randomized response)
6. Garandeer vertrouwelijkheid van de gegevens.
7. Breng de boodschap over dat de gegevens niet op een voor de respondent negatieve wijze gebruikt kunnen worden (N.B.: maak het niet te zwaar).
8. Wees respondentgericht.

4 Mogelijkheden om onderzoek naar gevoelige onderwerpen te optimaliseren

Hiervoor is aangegeven hoe een onderzoeker door een zorgvuldige keuze van dataverzamelmethode en vraagvorm de intrinsieke en extrinsieke dreiging kan verminderen. Daarnaast kunnen onderzoekers nog andere technieken toepassen om de mogelijke vertekening in onderzoek bij gevoelige onderwerpen terug te dringen en de validiteit van de conclusies te vergroten. Drie van deze technieken: het uittesten van de vragen, het trainen van de interviewer, en het verzamelen van aanvullend materiaal, worden hierna kort besproken.

4.1 Uittesten van de vragen

Wanneer men een nieuw geconstrueerde vragenlijst gebruikt of wanneer men een bestaande vragenlijst in een nieuwe situatie of bij een andere groep toepast, dan is het sterk aan te raden de vragen eerst aan een pretest te onderwerpen. Twee veel gebruikte methoden zijn 'beoordeling door deskundigen' en 'testen met behulp van kwalitatieve, cognitieve methoden'. Deze methoden worden vaak in combinatie toegepast.

Bij een *beoordeling door deskundigen* legt de onderzoeker het onderzoeksontwerp en de te gebruiken vragenlijst voor aan een aantal experts op het onderzoeksterrein. Dit kunnen collega-onderzoekers zijn, maar ook leden van belangengroeperingen (ervaringsdeskundigen), of sleutelfiguren in het veld (gezondheidswerkers, begeleiders, e.d.). Het is belangrijk dat de volledige onderzoeksopzet geschetst wordt, zodat de vragen in hun context beoordeeld kunnen worden. In een standaard vooronderzoek beoordelen de deskundigen de vragen op begrijpelijkheid, en geven commentaar op het taalgebruik. Ook wordt hen vaak gevraagd kritisch naar de vraagvolgorde te kijken om zo de mogelijke invloed van het antwoord op een voorgaande vraag op het daarna gegeven antwoord (contexteffect) op te kunnen sporen. Daarnaast wordt ook de lay-out van de vragenlijst en de onderzoekssituatie zelf van commentaar voorzien.

Wanneer men een gevoelig onderwerp onderzoekt, wordt bovenstaande procedure

aangevuld met een aantal vragen naar de mate van 'dreiging'. Zo kan de onderzoeker aan de deskundigen vragen om de vragen en antwoordmogelijkheden te scoren op mate van bedreigendheid voor de doelgroep. Men kan ook een oordeel vragen over de mate van privacy tijdens de bevragingssituatie en vragen naar suggesties of deze nog te verhogen is. Het voordeel van een beoordeling door deskundigen is dat deze snel en eenvoudig werkt, en nuttige, ter zake doende en toepasbare adviezen oplevert. Een nadeel is dat men de beoogde respondenten niet zelf ondervraagt. Daarom leggen onderzoekers vaak een tweede versie van de vragenlijst, waarin het commentaar van de deskundigen al is verwerkt, voor aan een aantal vertegenwoordigers van de te onderzoeken groep: men doet een pretest.

Bij een pretest legt de onderzoeker de vragen voor aan een kleine groep proefrespondenten (vijf à tien). Bij het samenstellen van de proefgroep wordt er zorg voor gedragen dat er van iedere voor het onderzoek belangrijke subgroep in ieder geval een vertegenwoordiger is; bij de keuze van stratificatiecriteria let men met name op de externe validiteit. Voorbeelden van stratificatiecriteria zijn leeftijd, leefsituatie en sekse. De proefrespondenten wordt uitgelegd dat het hier niet om een gewoon interview gaat, maar dat hun hulp gevraagd wordt bij het uittesten van de vragenlijst. In de instructie moet duidelijk naar voren komen dat de onderzoekers juist commentaar op de vragen en vragenlijst willen en dat het heel belangrijk is om te weten wat onduidelijk geformuleerd is en wat de respondent niet begrijpt. Vraagtechnieken die toegepast worden zijn onder andere hardop denken bij het beantwoorden van een vraag, het in eigen woorden laten parafraseren van de vraag en het direct vragen (direct probes) naar de betekenis van een woord in de vraag. Voor een overzicht van deze cognitieve interviewtechnieken zie Forsyth en Lessler (1991); voor een praktische beschrijving van verschillende methoden zie Fowler (1995). Ook bij kinderen is deze vorm van kwalitatieve pretest goed toe te passen. Wel dient men er zorg voor te dragen dat de groepjes homogeen zijn met betrekking tot leeftijd en dat er aparte groepen zijn voor jongens en voor meisjes. Voor een voorbeeld zie Scott (1997).

Bij een onderzoek naar gevoelige onderwer-

pen is het wenselijk daar in de pretest expliciet aandacht aan te schenken. Naast de standaardvragen naar het begrijpen van de vraag en de in de vraag gebruikte woorden, kan dan gevraagd worden wat de proefrespondenten van de vragen zelf vonden. Dit kan gebeuren aan de hand van open-interviewtechnieken, maar men kan ook meer gestructureerde technieken gebruiken. Zo kan men de proefrespondent de vragen op een ratingschaal laten beoordelen naar bedreigendheid, opdringerigheid en privé-karakter. Vaak wordt zo'n beoordeling twee keer gevraagd. De eerste keer wordt gevraagd om de eigen mening weer te geven: 'Hoe bedreigend vindt u deze vraag?' De tweede keer wordt gevraagd naar de opvatting van de meeste mensen in hun omgeving: 'Hoe bedreigend denkt u dat de meeste mensen in uw omgeving deze vraag vinden?' Deze laatste indirecte manier van vragen, is in feite een directe vraag naar de sociale norm zoals deze door de respondent ervaren wordt. Voor een beschrijving zie Bradburn, Sudman, Blair en Stocking (1989). Deze laatste indirecte benadering kan voor kinderen erg lastig zijn; men kan daarom kinderen beter een directe vraag stellen, eventueel in aanvulling met een vignetbenadering. Daarin kan de onderzoeker aan de proefrespondent vragen om de gehele uiteindelijke onderzoekssituatie (vragen + methode + setting) te beoordelen. De onderzoeker laat dan verschillende aspecten systematisch beoordelen door middel van vignetten (kleine verhaaltjes). Bij adolescenten en volwassenen kan men ook een globale benadering kiezen. In een open interview wordt eerst een schets gegeven van de opzet van het onderzoek en worden de vragen beschreven. Daarna wordt verder gesproken over hoe dit onderzoek op de respondent overkomt. Het is heel belangrijk niet alleen naar commentaar en meningen te vragen, maar op het einde van het testinterview ook te vragen naar suggesties voor verbetering van de onderzoekssituatie.

4.2 Persoon van de interviewer

Hoewel er veel onderzoek is gedaan naar interviewer-respndent interacties, is er weinig hard empirisch bewijs over welke interviewerkenmerken nu de kwaliteit van de antwoorden beïnvloeden. Er zijn wel duidelijke effecten gevonden van sekse en ras van de interviewer, maar alléén bij vragen die te maken hebben met

aan sekse- of rasgerelateerde onderwerpen. Het zal geen verbazing wekken dat vrouwen iets opener over gynaecologisch onderzoek praten met vrouwelijke interviewers. Ook bleek dat bij gezondheidssenquêtes 'gewone' standaard interviewers, die een korte aanvullende training over het onderwerp van de enquête hadden gekregen, betere, opener antwoorden verkregen dan gezondheidswerkers die een aanvullende interviewertraining hadden gekregen. Dit kan veroorzaakt worden door de houding die de interviewer uitstraalt: gezondheidswerkers kunnen te veel problemen verwachten bij moeilijke en gevoelige vragen en dit onbewust op de respondent overbrengen. Uit empirisch onderzoek is gebleken dat interviewers die zelf een vraag te persoonlijk vinden of moeilijkheden verwachten bij deze vraag, ook 'slechtere' antwoorden krijgen van respondenten, bijvoorbeeld onderrapportage of veel weet-niet-antwoorden. Voor een overzicht zie Bradburn (1983).

Uit het bovenstaande blijkt dat er weinig criteria zijn om interviewers op te selecteren bij onderzoek naar gevoelige onderwerpen. Het belangrijkste criterium is dat de interviewer zelf geen problemen met het onderwerp heeft en zich op zijn/haar gemak voelt bij de te ondervragen personen. Het is van groot belang dat de interviewers een goede training krijgen. Zoals bij elk onderzoek dienen interviewers een basistraining te ontvangen (het stellen van vragen, het leren doorvragen, het geven van feedback, het gebruik van verbale en nonverbale communicatie). Een degelijke basistraining blijkt al in staat te zijn om de respons op bedreigende vragen te verhogen (Billiet & Loosveldt, 1988). Tijdens deze basistraining is het aan te raden om de interviewers zelf de vragenlijst te laten invullen en in een nagesprek in te gaan op vragen die ze zelf gevoelig vonden. Niet alleen leren de interviewers zo de vragenlijst kennen, maar ze ervaren ook hoe het is om deze vragen te beantwoorden; door het bespreken van twijfels en problemen leren de interviewers ook beter met hun eigen gevoelens over de vragen omgaan. Bovendien verkrijgt de onderzoeker zo extra materiaal om naderhand statistisch mogelijke interviewereffecten te kunnen onderzoeken en te corrigeren (vgl. Hox, 1994).

Naast deze basistraining is het aan te beve-

len een aanvullende training te geven waarin de interviewers informatie krijgen over de achtergrond van het onderwerp. Bij onderzoek onder specifieke groepen is het belangrijk dat de interviewers op de hoogte worden gebracht van het taalgebruik en de cultuur van deze groepen. Men denke bijvoorbeeld aan onderzoek naar het effect van veilig-vrijencampagnes op het seksuele gedrag van homoseksuele jongeren of aan onderzoek naar verdovende-middelengebruik bij middelbare scholieren. Tijdens de training dient de neutrale interviewerrol benadrukt te worden. Zo is het belangrijk om voor ogen te houden dat een interviewer neutraal is, geen oordeel uitsprekt, en niet makkelijk te schokken is. Door middel van rollenspel en proefinterviews kan men de interviewers hiermee laten oefenen. Daarbij dient extra aandacht geschonken te worden aan de nonverbale communicatie: een interviewer kan in woorden wel de juiste neutrale reactie geven, maar door toonhoogte en lichaamstaal toch een veroordelende of andere emotionele boodschap overbrengen. Hiermee wordt niet bedoeld dat een interviewer zich als een robot zou moeten gedragen. Een interviewer dient vriendelijk en open te zijn en te reageren op de respondent; een interviewer mag alleen nooit 'oordelen' of 'vooroordelen'!

Ook na de training is een goede begeleiding van de interviewers van belang. Een veldgids met basisinformatie en telefoonnummers van relevante (hulp)organisaties kan een goede steun zijn. Daarnaast behoren de interviewers een telefoonnummer te hebben waarmee ze in geval van nood iemand van de onderzoeksleiding altijd kunnen bellen. Maar ook regelmatig contact met iemand van de onderzoeksleiding is aan te raden. Zo kunnen regelmatige korte telefoongesprekken op initiatief van de onderzoeksleiding niet alleen dienen om zakelijke regelingen te treffen, maar ook de interviewers de gelegenheid te geven even stoom af te blazen. Bij een langer lopend onderzoek is het verstandig een aantal 'terugkomdagen voor interviewers' in het schema op te nemen. Deze dagen bieden onderzoekers de gelegenheid om aanvullende instructies te geven en de neuzen weer dezelfde kant uit te laten wijzen; daarnaast geven ze interviewers de gelegenheid om elkaar te ontmoeten, en problemen te bespreken. Naast interviewtechnische begeleiding

kan ook emotionele begeleiding en opvang van interviewers belangrijk zijn. In welke mate en hoe dit dient te geschieden is natuurlijk zeer afhankelijk van het onderwerp van onderzoek en kan per onderwerp sterk variëren.

4.3 Verzamelen van aanvullende informatie

In een aantal gevallen is het mogelijk om aanvullende informatie te verzamelen om zo de kwaliteit van de verkregen data te onderzoeken en eventueel te corrigeren. Deze aanvullende informatie kan zowel tijdens het onderzoek zelf verzameld worden, bijvoorbeeld door het opnemen van een aantal extra methodologische vragen, als na het onderzoek verzameld worden, bijvoorbeeld statistische informatie uit bestaande bronnen.

Bij het opnemen van methodologische vragen kan men denken aan een sociaalwenselijkheidsschaal. Zo vond Gibson (1994) dat heroïneverslaafden die sociaalwenselijke antwoorden gaven volgens een sociaalwenselijkheidsschaal minder druggerelateerd risico gedrag toegaven dan heroïneverslaafden die laag scoorden op de sociaalwenselijkheidschaal. Hij verbond hieraan de conclusie dat de hoog-sociaalwenselijke groep waarschijnlijk de mate van risicogedrag heeft ondergerapporteerd. Behalve via een algemene sociaalwenselijkheidsschaal kan men ook direct respondenten vragen hoe ze bepaalde vragen ervaren, bijvoorbeeld door te vragen hoe ze zich voelden bij het beantwoorden van specifiek vragen (1= heel erg gespannen tot 5= absoluut niet gespannen). Ook kan men aan het einde van de vragenlijst de respondenten vragen specifieke vragen te beoordelen op 'bedreigendheid'.

Wanneer men gebruik maakt van interviewers bij de dataverzameling, dan kan men ook de interviewers als observator gebruiken en aan het einde van ieder vraaggesprek een aantal vragen laten beantwoorden over het interviewproces. Zo kan men de interviewer vragen hoe de sfeer was tijdens het interview, of de respondent gespannen was en zo ja bij welke vragen, en of er in het algemeen vragen waren die problemen opleverden.

Behalve via de interviewer en de respondent zelf kan men ook uit andere bronnen aanvullende gegevens halen. Dit kan zowel op individueel niveau, bijvoorbeeld gegevens uit het gebruikte steekproefkader, als op geaggregeerd

of groepsniveau, bijvoorbeeld gezondheidsstatistieken of schoolgegevens. Deze gegevens kunnen onder andere gebruikt worden om na te gaan of speciale groepen geweigerd hebben aan het onderzoek mee te doen (selectieve non-respons).

In een aantal speciale gevallen wordt ook wel een extra onderzoek opgezet om de uit interviews of vragenlijsten verkregen gegevens te kunnen controleren. Er worden dan niet-reactieve methoden gebruikt zoals observatiegegevens (vgl. Webb, Campbell, Schwartz & Sechrest, 1966). Voorbeelden zijn het tellen van lege drankflessen op vuilnisbakkendag bij een studie naar alcoholconsumptie en het verzamelen van verkoopcijfers van tandenborstels, dental floss en tandenstokers bij een lokaal onderzoek naar tandartsvoorlichting. Uit de voorbeelden blijkt al dat dit type onderzoek veel inventiviteit en tijd vraagt.

4.4 Statistische controle en correctie

Waarom nu al die moeite om extra gegevens te verzamelen? Met behulp van deze extra gegevens kan een statisticus nagaan of er ernstig gevaar is voor vertekening van de antwoorden. Bedreigingen voor de validiteit kunnen met deze extra informatie worden opgespoord en in een aantal gevallen statistisch worden gecorrigeerd.

Zo dienen bij ieder onderzoek de resultaten van een analyse naar de nonrespons gerapporteerd te worden. Hoe groot is de nonrespons, wie zijn de nonrespondenten, verschillen zij op belangrijke achtergrondvariabelen van diegene die wel meedoen. Als dit laatste het geval is dan kan men overwegen om de data met wegingsmethoden te heranalyseren, en zo voor de nonrespons te corrigeren (vgl. Bethlehem et al, 1996).

Wanneer men informatie heeft over welke vragen bedreigend gevonden werden, of wanneer men de score op een sociale wenselijkheidsschaal ter beschikking heeft, dan kan men voor verschillende subgroepen nagaan of zij ook verschillende antwoordpatronen op de gevoelige vragen vertonen. In het meest ideale geval kan men de gegevens 'ijken' met behulp van data uit andere bronnen.

Bij al deze statistische analyses is het belangrijk dat men in het rapport duidelijk vermeldt welke analyses er gebruikt zijn en welke verschillen er gevonden zijn.

5 Samenvatting

In dit artikel is ingegaan op de problemen die zich voor kunnen doen bij onderzoek naar gevoelige onderwerpen. In het eerste gedeelte is een overzicht gegeven van de mechanismen die kunnen leiden tot vertekening van de antwoorden. Hierbij stonden de begrippen intrinsieke en extrinsieke dreiging centraal. In het tweede gedeelte zijn methoden besproken om de vertekening in onderzoek bij gevoelige onderwerpen terug te dringen en zo de validiteit van de conclusies te vergroten. Aan bod kwamen: 1) speciale methoden van dataverzameling die zich richten op het vergroten van de privacy bij het beantwoorden van de vragen, 2) het aanpassen van de vraagvorm zodanig dat sociale-wenselijkheidstendenties minder kans krijgen, 3) het uittesten en optimaliseren van de vragenlijst, 4) het speciaal trainen en begeleiden van interviewers zodat deze optimaal functioneren in moeilijke situaties, en 5) het verzamelen van additionele informatie uit andere bronnen.

Deze methoden moeten gezien worden als gereedschap dat behoort tot het instrumentarium van de onderzoeker. Dit gereedschap dient met zorg uitgekozen en gehanteerd te worden, en zal zeker niet altijd bij ieder onderzoek worden toegepast. Welke mix van methoden gebruikt wordt, hangt in grote mate af van de onderzoeksvraag en daarmee samenhangende factoren (b.v. hoe gevoelig is het onderwerp, waar wordt het onderzoek gehouden, en wordt een speciale (rand)groep onderzocht), én van de beschikbare middelen (tijd, personeel, geld).

Voor onderzoekers is het belangrijk om zich te realiseren dat met enige inventiviteit meer kan dan men aanvankelijk verwacht had. Het is ook belangrijk om open te staan voor ervaringen van anderen, én om de eigen ervaringen met anderen te delen: maak praktische en methodologische kennis openbaar!

Literatuur

- Aquilino, W.S. (1993). Effects of spouse presence during the interview on survey responses. *Public Opinion Quarterly*, 57, 358-376.
- Bethlehem, J. et al. (1996). *Bascula Reference Manual*. Centraal Bureau voor de Statistiek, Voorburg/Heerlen.

- Billiet, J. en Loosveldt G. (1988). Improvement of the quality of the responses to factual survey questions by interviewer training. *Public Opinion Quarterly*, 1988, 52, 190-21.
- Bradburn, N.M. (1983). Response effects. In: P.H. Rossi, J.D. Wright en A.B. Anderson (Eds), *Handbook of survey research*. New York: Academic Press.
- Bradburn, N.M. en Sudman, S. (1979). *Improving interview method and questionnaire design*. San Francisco: Jossey Bass.
- Bradburn, N.M., Sudman S., Blair, E., en Stocking C. (1989). Question threat and response bias. In: E. Singer en S. Presser (Eds) *Survey research methods: A reader*. Chicago: University of Chicago Press. (Het betreffende artikel verscheen in eerste instantie in *Public Opinion Quarterly*, 1978, 42, 2).
- Burt, M. (1995). Handling sensitive information on US homeless surveys. Paper gepresenteerd op de European Conference on Demography, Statistics and Privacy, Parijs, oktober 1995.
- Commissie Kordes (1997). *Privacy-wetgeving en het gebruik van persoonsgegevens voor wetenschappelijk en statistische doeleinden*. Advies aan de minister van Onderwijs, Cultuur en Wetenschappen, Den Haag/Zoetermeer, 1997, hoofdstuk 6.
- Dillman, D.A. (1978). *Mail and telephone surveys*. New York: Wiley.
- Forsyth, B.H. en Lessler, J.T. (1991). Cognitive laboratory methods: A taxonomy. In: P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz en S. Sudman (Eds), *Measurement errors in surveys*. New York: Wiley.
- Fowler, F.J. jr. (1995). *Improving survey questions: Design and evaluation*. Thousand Oaks, CA: Sage.
- Gibson, D. (1994). Drug injectors may underreport drug risk but not sexual risk behavior. *Abstract book tenth international conference on Aids*. Yokohama, Japan, 7-12 Aug. 1994.
- Heijden, P.G.M. van der, Gils, G. van, Bouts, J., & Hox, J. (1998). A comparison of randomized response, CASAQ, and direct questioning: Eliciting sensitive information in the context of social security fraud. *Kwantitatieve Methoden*, 59, 15-34.
- Heijden, P.G.M. van der, en Gils, G. van (1996). Some logistic regression models for randomized response data. In: A. Forcina, G.M. Marchetti, R. Hatzinger, and G. Galmatti (eds). *Statistical modeling, Proceedings of the 11th international workshop on statistical modeling*, Orvieto, Italy.
- Hippler, H.-J., Schwarz, N. en Singer, E. (1990). Der Einfluss von Datenschutzzusagen auf die Teilnamebereitschaft an Umfragen. *ZUMA-Nachrichten*, 27, 54-67.
- Hox, J.J. (1994). Hierarchical regression models for interviewer and respondent effects. *Sociological Methods and Research*, 22, 3, 300-318.
- Lee, R.M. (1993). *Doing research on sensitive topics*. London: Sage.
- Leeuw, E.D. de, Hox, J., Kef, S. & Van Hattum, M. (1997). *Overcoming the problems of special interviews on sensitive topics: computer assisted self-interviewing tailored for young children and adolescents*. MLS-publicatin, nr. 60. Universiteit van Amsterdam.
- Leeuw, E.D. de (1992). *Data quality in mail, telephone, and face to face surveys*. Amsterdam: TT-publikaties.
- Leeuw, E.D. de, Hox, J.J. en Snijders, G.J.M. (1995). The effect of computer assisted interviewing on data quality; A review. *Journal of the Market Research Society*, 37, 325-344.
- Lyberg, L.E. en Dean, P. (1992). *Methods for reducing nonresponse rates: A review*. Paper gepresenteerd op de jaarlijkse conferentie van AAPOR, St Petersburg, Florida, mei, 1992.
- Makkai, T., en McAllister, I. (1992). Measuring social indicators in opinion research; A method to improve accuracy on sensitive questions. *Social Indicators Research*, 27, 2, 169-186.
- Scott, J. (1997). Children as respondents: Methods for improving data quality. In: L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, D. Trewin, C. Dippo, en N. Schwarz (Eds). *Survey measurement and process quality*. New York: Wiley.
- Scheers, N.J. en Dayton, C.M. (1988). Covariate randomized response models. *Journal of the American Statistical Association*, 83, 969-974.
- Sudman, S. en Bradburn, N.M. (1982). *Asking questions. A practical guide to questionnaire design*. San Francisco: Jossey-Bass.
- Turner, C.F., Ku, L., Rogers, S.M., Lindberg, L.D., Pleck, J.H., Sonenstein, F.L. (1998). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science*, 280, 867-873.
- Umesh, U.N. en Peterson, R.A. (1991). A critical evaluation of the randomized response method. Applications, validation, and research agenda. *Sociological Methods and Research*, 20, 1, 104-138.

VMO (Vereniging van marktonderzoeksbureaus)
Bureau-informatie (1997), *Gedragcode m.b.t marktonderzoek, persoonsgegevens en privacy-bescherming*.

Webb, E.J., Campbell, D.T., Schwartz, R.D. en Sechrest, L. (1966). *Unobtrusive measures: nonre-active research in the social sciences*. Chicago: Rand McNally.

Weisband, S. & Kiesler, S (1996). Self-disclosure on computer forms: Meta analysis and implications. CHI '96 (<http://www.al.arizona.edu/~weisband/chi/chi96.html>).

Auteur

Edith D. de Leeuw doceert methoden en statistiek aan het University College Utrecht; daarnaast is zij werkzaam als zelfstandig consultant 'methoden en statistiek'. Zij is secretaris van de sociaal-wetenschappelijke sectie van de Vereniging voor Statistiek en president van de internationale onderzoeksgroep on Logic and Methodology van de ISA.

Adres: Plantage Doklaan 40, NL-1018 CN Amsterdam

E-mail: edithL@educ.uva.nl

Abstract

Asking sensitive information: A review of methods

E.D. de Leeuw, *Pedagogische Studiën* 1999, 76, 92-103.

Missing data and socially desirable answers threaten the validity of studies on sensitive topics. To find ways to counteract these effects and to improve data quality, methodologists have studied the question-answer process with sensitive questions and proposed optimized methods for asking sensitive information. This article gives an overview of methods to improve the data quality when asking sensitive questions. Both theory and practical applications are discussed.

Data verzamelen bij basisschoolleerlingen en adolescenten met een visuele beperking met behulp van de computer

M.J.C. van Hattum en S. Kef*

Samenvatting

In onderzoek wordt veelvuldig gebruik gemaakt van vragenlijsten. Deze methode biedt voordelen wanneer informatie verzameld wordt over gevoelige onderwerpen. Een nadeel van vragenlijstonderzoek is dat alleen relatief eenvoudige vragenlijsten gebruikt kunnen worden. Een methode om dit probleem te verhelpen is het gebruik van de computer bij dataverzameling: Computer Assisted Data Collection (CADAC).

Na een literatuuroverzicht wordt ingegaan op mogelijke vormen van CADAC. Vervolgens worden twee studies beschreven. In het eerste onderzoek is een Disk by Mail survey over pesten geïmplementeerd in het regulier basisonderwijs. De tweede studie bestaat uit een survey over persoonlijke netwerken, sociale steun en het welbevinden van adolescenten met visuele beperkingen. Hierbij is gebruik gemaakt van een combinatie van Computer Assisted Personal Interviewing en Computer Assisted Self Interviewing. Bij de beschrijvingen van beide studies wordt ingegaan op logistieke aspecten en data kwaliteit. Het artikel wordt afgesloten met suggesties voor het gebruik van CADAC bij kinderen en adolescenten.

Inleiding

In sociaal-wetenschappelijk onderzoek wordt veelvuldig gebruik gemaakt van zelfrapportages. Met name bij onderzoek waarbij gevoelige onderwerpen aan bod komen, blijken zelfrapportages, beter dan interviews, een goede manier om het gevoel van anonimiteit van de respondent te vergroten en sociale wenselijkheid te reduceren (De Leeuw, 1992; Tourangeau & Smith, 1996). Tevens heeft deze methode als voordeel dat de respondent in

eigen tempo de gestelde vragen beantwoordt. Dit komt ten goede aan de kwaliteit van de antwoorden, met name wanneer de onderzoeker te maken heeft met speciale populaties zoals kinderen, adolescenten of bejaarden (De Leeuw & Collins, 1997). Een bijkomend voordeel van zelfrapportages die via de post verstuurd worden naar respondenten, zijn de lage kosten en de geringe hulpmiddelen die noodzakelijk zijn.

Een nadeel is dat alleen relatief eenvoudige vragenlijsten gebruikt kunnen worden. Gecompliceerde sprongen en lastige vertakkingen in de vragenlijst vormen een bedreiging voor zowel de datakwaliteit als de motivatie van de respondent om de vragenlijsten geheel af te ronden. Een manier om dit probleem te verhelpen is het gebruik van Computer Assisted Data Collection (CADAC). Hierbij wordt de vragenlijst via de computer aangeboden aan de respondent, waarbij de meest gecompliceerde vragenlijsten in een eenvoudige vorm op het beeldscherm verschijnen. CADAC maakt het zelfs mogelijk complexe vragenlijsten voor te leggen zonder dat de hulp van een interviewer noodzakelijk is. Het interviewprogramma regelt de voortgang van het interview en de volgorde van de vragen. De vragen worden vanaf het scherm gelezen, waarna een antwoord ingetypt kan worden. Nadat het antwoord gegeven is, verdwijnt de vraag van het scherm en wordt de volgende vraag automatisch aangeboden. Alleen die vragen die op de betreffende respondent van toepassing zijn, verschijnen op het beeldscherm, waardoor er geen mogelijkheid bestaat dat vragen per ongeluk overgeslagen worden. De kans op missing values wordt dan ook minimaal, aangezien de computer de route van de respondent door de vragenlijst controleert en checkt of de respondent een antwoord heeft ingetypt alvorens er kan worden doorgegaan naar de volgende vraag.

Nicholls, Baker en Martin (1997) geven aan dat de non-respons, veroorzaakt door fouten

* De auteurs staan op alfabetische volgorde.

van de interviewers of de respondenten duidelijk zichtbaar vermindert, maar dat er weinig verschil is tussen een computergestuurde vragenlijst en een papieren versie wat betreft het aantal 'weet niet' en 'geen mening' antwoorden.

Bij een computergestuurde afname kunnen de items aan iedere respondent in een specifieke willekeurige volgorde aangeboden worden via randomisering. Op deze wijze worden de volgorde-effecten gerandomiseerd, waardoor men mag hopen op meer valide uitkomsten.

In het algemeen wordt in de literatuur aangegeven dat respondenten een computergestuurde afname van een vragenlijst als prettig ervaren; ze vinden het interessant, leuk en prima uit te voeren (Witt & Bernstein, 1992). Respondenten blijken een onderschatting te maken van de tijd die ze nodig hebben om de computergestuurde vragenlijst in te vullen (Higgins, Dimnik & Greenwood, 1987). Verder wordt naar aanleiding van een onderzoek van Beckenbach (1995) geconcludeerd dat meer dan 80% van de respondenten geen enkel probleem ondervond om de computer en het interviewprogramma te gebruiken. Een klein aantal respondenten klaagde over fysieke problemen, zoals oogproblemen door de inspanning van het lezen via een beeldscherm. Wel moet worden benadrukt dat respondenten beter niet onder tijdsdruk gezet kunnen worden, aangezien er aanwijzingen bestaan dat tijdsdruk negatief interacteert met perceptuele en motorische vaardigheden die nodig zijn voor het lezen van vragen vanaf een beeldscherm en het correct intypen van een antwoord.

Zoals eerder genoemd, blijken respondenten bij computergestuurde dataverzameling geneigd te zijn zich meer bloot te geven over gevoelige onderwerpen. In een meta-analyse van Weisband en Kiesler (1996) wordt een duidelijk effect gevonden ten voordele van de computergestuurde dataverzameling. Dit effect was sterker aanwezig bij de vergelijking tussen CADAC en een face-to-face-interview, maar ook bij de vergelijking tussen CADAC en de zelfrapportage bleken respondenten zich significant meer open te stellen bij de computergestuurde methode. Eveneens is aangetoond dat dit effect groter wordt naarmate de informatie gevoeliger ligt. Over het algemeen worden

geen grote verschillen gevonden wat betreft betrouwbaarheid en validiteit tussen een computergestuurde vragenlijst en een papieren versie van de vragenlijst (Harrel & Lombardo, 1984).

Samenvattend kan geconcludeerd worden dat vergelijkingsonderzoek tussen papier-en-pen-vragenlijsten en computergestuurde vragenlijsten aangetoond heeft dat bij de computergestuurde versie minder sprake is van non-respons en de respondenten meer geneigd zijn zich open te stellen ten aanzien van gevoelige onderwerpen. Dit betekent dat de kwaliteit van de data van computergestuurde vragenlijsten hoger is. Daarnaast is de computergestuurde methode een extra stimulans voor de respondenten door het leuke en nieuwe karakter ervan.

In de volgende paragraaf worden diverse vormen van CADAC beschreven. Daarna wordt een tweetal voorbeelden van computergebruik bij dataverzameling uitgewerkt. Hierbij gaat het om twee specifieke doelgroepen, namelijk basisschoolleerlingen in de leeftijd van ongeveer 8 tot 12 jaar en adolescenten met een visuele beperking. Ingegaan wordt op de organisatorische aspecten, data kwaliteit en speciale aanpassingen die gemaakt moesten worden.

1 Mogelijke vormen van computergestuurde dataverzameling

In de literatuur worden naast de in dit artikel gebruikte term Computer Assisted Data Collection (CADAC) diverse namen gebruikt voor computergestuurde dataverzameling, zoals Computer Assisted Interviewing (CAI) of Computer Assisted Survey Information Collection (CASIC). Kenmerkend voor de computergestuurde methode is dat de vragen gelezen worden van het beeldscherm en dat de antwoorden direct in de computer ingevoerd worden door een interviewer dan wel door de respondent zelf. In deze paragraaf worden de diverse vormen op een rijtje gezet en vergeleken met de vormen die gewoonlijk in onderzoek gebruikt worden. Hierbij wordt gebruik

Tabel 1
Overzicht survey vormen: traditioneel en computergestuurd

<i>Traditionele vragenlijst</i>	<i>Computergestuurde vragenlijst</i>
Telefonische enquête of interview	Computer Assisted Telephone Interviewing (CATI)
Persoonlijk (ofwel: face-to-face-)interview	Computer Assisted Personal Interviewing (CAPI)
Schriftelijke vragenlijst met aanwezigheid van een interviewer	Computer Assisted Self Interviewing (CASI)
Schriftelijke vragenlijst (zelfrapportage)	Computer Assisted Self Administering Questionnaire (CSAQ)
<i>Vormen:</i> Post-enquête Elektronische enquête	<i>Vormen:</i> Disk By Mail (DBM) Electronic Mail Survey (EMS)

gemaakt van een overzicht van De Leeuw, Snijkers en Hox (1994) en De Leeuw, Hox en Snijkers (1995).

Een computergestuurd survey kan op vier manieren plaatsvinden. De eerste is een telefonische enquête: Computer Assisted Telephone Interviewing (CATI). De interviewer zit achter een computer en stelt de vraag die op het beeldscherm verschijnt aan de respondent met wie telefonisch contact gelegd is. Het antwoord wordt door de interviewer direct ingetypt. Ten tweede noemen de auteurs het persoonlijke interview: de Computer Assisted Personal Interviewing (CAPI). De interviewer brengt een bezoek aan de respondent met een draagbare computer en neemt een face-to-face-interview af, waarbij de antwoorden rechtstreeks ingetypt worden. De derde manier is Computer Assisted Self Interviewing (CASI), waarbij (een gedeelte van) de vragenlijst in aanwezigheid van een interviewer door de respondent zelf ingevuld wordt. De respondent leest de vragen van het beeldscherm en typt een antwoord in. Deze methode kan gecombineerd worden met de CAPI-methode. Dan geeft een interviewer na een aantal vragen de (draagbare) computer over aan de respondent die vervolgens bijvoorbeeld een aantal gevoeligliggende vragen (over bijvoorbeeld de hoogte van een inkomen of eenzaamheidsgevoelens) zelf beantwoordt. De laatste manier, die op twee manieren kan plaatsvinden, wordt Computer Self-Administered Questionnaires (CSAQ) genoemd. Men spreekt van Disk By Mail (DBM), indien de vragenlijst via een diskette per post verstuurd wordt naar de respondent die hem vervolgens, na beantwoording, retourneert. Een voorwaarde voor DBM is dat de

betreffende respondent de beschikking heeft over een eigen computer of laptop. Een andere manier waarop de CSAQ methode gebruikt wordt, is door de vragenlijst elektronisch te versturen via een computernetwerk. Deze vorm wordt ook wel Electronic Mail Survey (EMS) genoemd. Voor een gedetailleerd beschrijving wordt verwezen naar De Leeuw, Snijkers en Hox (1994).

Tabel 1 geeft een schematische weergave van de diverse survey-vormen, met daarnaast de gebruikte terminologie voor de betreffende computergestuurde vorm.

2 Onderzoek bij basisschoolleerlingen en adolescenten

Zoals we in de vorige paragraaf gezien hebben, kan het computergebruik bij dataverzameling op verschillende manieren plaatsvinden. In deze paragraaf worden twee toepassingen beschreven van computergebruik bij dataverzameling onder speciale populaties, namelijk basisschoolleerlingen en blinde of slechtziende adolescenten.

In de eerste studie die beschreven wordt, is een Disk By Mail uitgevoerd over pesten in het regulier basisonderwijs. De respondenten bestaan uit leerlingen van groep 6, 7 en 8 van het basisonderwijs (leeftijd ongeveer 8 tot 12 jaar).

De tweede studie is een onderzoek naar persoonlijke netwerken van adolescenten met een visuele beperking. Centraal in deze studie staat de ervaren sociale steun en het welbevinden van deze jongeren. De leeftijd van de respondenten varieert van 14 tot 24 jaar. Bij de dataverzameling is gebruik gemaakt van een vra-

genlijst die gedeeltelijk door de interviewer en gedeeltelijk door de respondent ingevuld wordt, oftewel een gemengde CAPI/CASI methode. Bij beide onderzoeken is bij de constructie van de gecomputeriseerde vragenlijst gebruik gemaakt van het software programma Ci3 (Sawtooth, 1994).

2.1 Studie 1: Een Disk By Mail onderzoek bij kinderen in het regulier basisonderwijs

In 1995 is een grootschalig vragenlijstonderzoek van start gegaan over pesten in het regulier basisonderwijs. In dit onderzoek zijn bij leerkrachten van groep 6, 7 en 8 en bij hun leerlingen data verzameld over hun beleving, visie en handelen omtrent pesten (Van Hattum, 1997). In dit artikel wordt enkel de dataverzameling bij de leerlingen beschreven.

Bij dit onderzoek is gebruik gemaakt van CSAQ. Bekend is dat kinderen terughoudend zijn ten opzichte van volwassenen, zoals leerkrachten en ouders, wat betreft het toegeven dat ze gepest worden dan wel zelf pesten (Besag, 1989; Heenan, 1992; Olweus, 1992; 1993). Zoals in de inleiding aangegeven, is het gebruik van de computer bij het beantwoorden van een schriftelijke vragenlijst over een sensitief onderwerp, zoals pesten, een methode om dit probleem te verminderen. Daar komt bij dat, aangezien de respondenten in dit geval kinderen zijn, de vragenlijst er aantrekkelijk en eenvoudig moet uitzien, hetgeen met CSAQ goed te realiseren is.

In de vorige paragraaf zijn diverse vormen van CSAQ genoemd. In het onderzoek is gekozen voor Disk By Mail, aangezien de respondenten hiervoor de schoolcomputer, waarmee ze vertrouwd zijn, konden gebruiken. Dankzij het door de regering gefinancierde Comenius project, dat in 1989 van start gegaan is, heeft het primair basisonderwijs de toegang tot één of meer personal computers per school (Van der Ree, 1994), waardoor implementatie van DBM mogelijk is.

Logistieke aspecten

De volgende logistieke aspecten komen aan bod: het gebruikte instrumentarium, de subjecten en de fase van dataverzameling.

De vragenlijst voor de leerlingen richt zich op de houding ten aanzien van pesten, de manier

waarop volwassenen met pesten omgaan en het feitelijke pesten in de school, waarbij gevraagd wordt of het betreffende kind als actieve pester opereert en/of zich slachtoffer van pesten voelt. Deze vragenlijst bestaat uit 99 items. Voor dit artikel wordt gebruik gemaakt van de actorschaal, die bestaat uit 21 items over het zelf pesten van anderen (b.v. 'Ik heb soms zin om andere kinderen uit mijn groep te schoppen of slaan', 'Ik heb soms zin om gemene en vervelende dingen tegen andere kinderen uit mijn groep te zeggen'), en van de slachtofferschaal, die bestaat uit 20 items (b.v. 'Ik word vaak geschopt en geslagen door andere kinderen', 'Andere kinderen doen stiekem rot tegen mij zonder dat de meester/juf het ziet'). De items worden op een vijf puntsschaal gescoord (1= 'helemaal niet waar' tot 5= 'helemaal waar') (Van Hattum, 1997). Eveneens wordt een schaal voor het meten van sociale wenselijkheid in dit onderzoek meegenomen (Van der Wolf, 1990). Deze schaal bestaat uit 9 items, die eveneens op een vijf puntsschaal gescoord worden (1= 'helemaal niet waar' tot 5= 'helemaal waar').

Bij de constructie van de computergestuurde vragenlijst zijn de vragen binnen een subonderdeel gerandomiseerd. Hierdoor wordt de variatie in de volgorde-effecten verhoogd. Voor iedere vraag is een range van mogelijke antwoorden gegeven. Voor de zogenaamde 'weet niet' antwoorden is de code '9' gedefinieerd. Deze mogelijkheid voor het gebruik van de code '9' verscheen niet op het scherm, maar werd in een speciale instructie toegelicht.

De mogelijkheid is gecreëerd om tussendoor tijdelijk een pauze te nemen of de vragenlijst in etappes in te vullen. Dit is gedaan met het oog op kinderen die mogelijk sneller moe zijn of met het oog op de leerkracht die het betreffende kind weer in de klas wil hebben voor bijvoorbeeld een klassikale instructie. Bij de programmering is eenvormigheid in de layout van de vragenlijst bewerkstelligd. Ook is een instructie ingebouwd waarmee de naam van het kind diverse malen tussendoor ter aandachtiging op het scherm terugkomt.

Eveneens was een papieren versie van de vragenlijst beschikbaar als back-up.

Naast deze vragenlijst over pesten is de Syracuse Amsterdam/Groningen Sociometrische Schaal (SAGS) voorgelegd aan de kin-

deren (Defares, Kema, Van Praag, & Van der Werff, 1977). De SAGS is een meer objectieve methode om de sociale positie van een kind in zijn groep vast te stellen. Met behulp van dit instrument kunnen geïsoleerde en populaire kinderen in de groep geïdentificeerd worden. Deze gegevens kunnen vergeleken worden met de subjectieve ervaring van kinderen, verkregen uit de vragenlijst. Bij de afname van de SAGS is gebruik gemaakt van een papieren versie. Deze is klassikaal afgenomen, aangezien de instructie voor dit instrument complex is.

Uit de populatie reguliere basisscholen in Nederland is een aselechte steekproef getrokken van 217 scholen. Deze scholen zijn in november 1994 benaderd voor deelname aan het onderzoek; in eerste instantie reageerden 119 scholen hierop positief. Als redenen voor een weigering werden met name genoemd; 'geen tijd', 'pesten is geen probleem op onze school', 'we doen op onze school al wat op het gebied van pesten'. De gebruikte methode (DBM) werd niet als reden van weigering genoemd.

Het onderzoek is uitgevoerd in de groepen 6, 7 en 8. Op een klein aantal scholen is groep 5 meegenomen, aangezien deze groep een combinatiegroep vormde met groep 6 en de betreffende leerkracht een totaaloverzicht wilde hebben ten aanzien van pesten. Uiteindelijk hebben 106 scholen (263 groepen) de vragenlijst daadwerkelijk ingevuld en geretourneerd, hetgeen resulteert in een responspercentage van 49%. Op 8 scholen (18 groepen) is een papieren versie van de vragenlijst afgenomen. Dit betekent dat 556 kinderen een papieren versie van de vragenlijst ingevuld hebben en 5872 kinderen de computergestuurde versie.

In maart 1995 is aan de deelnemende scholen een pakket toegestuurd met twee of meer diskettes (afhankelijk van het aantal computers), een korte instructie op papier voor zowel de leerkracht als de leerlingen en een begeleidende brief. Op de diskettes bevonden zich zogenaamde 'bat-files' om de vragenlijst op te starten, tussendoor te onderbreken en wederom op te starten vanaf de positie waar de betreffende respondent gebleven was, de data te bewaren en back-ups te maken.

De instructie voor de leerkracht was gericht

op het opstarten van de vragenlijst voor de kinderen, en voor de vragenlijst die door de leerkracht zelf ingevuld diende te worden. De instructie voor de leerlingen bestond uit een vel met bondig weergegeven punten, zoals het gebruik van de <Enter> en de <Backspace> en met uitleg over de pieptoon die het kind te horen zou krijgen wanneer het een antwoord zou intypen dat buiten de beschikbare range zou vallen. Op deze instructie werd eveneens aangegeven dat de code '9' ingetypt kon worden wanneer ze echt geen enkel antwoord konden bedenken.

In deze periode is een helpdesk ingesteld voor telefonische advisering en was het mogelijk in geval van nood personen van deze helpdesk in te schakelen. Daarnaast hebben elf scholen met relatief weinig computers tijdelijk de beschikking gekregen over een aantal laptops van de universiteit en over persoonlijke assistentie. Op één school is om algemene assistentie gevraagd; deze school maakte zich ernstige zorgen over hun bekwaamheid deze 'computerdingen' ten uitvoer te brengen.

Op zes scholen is een papieren versie van de vragenlijst afgenomen. Deze scholen waren relatief groot, waardoor het onderzoek erg veel tijd zou gaan vergen, aangezien de leerlingen één voor één de vragenlijst invullen. Ook kwam het voor dat de directie aangaf wegens onervarenheid de computergestuurde dataverzameling niet aan te durven. Uiteindelijk is bij nog twee scholen een papieren versie afgenomen aangezien de computer op school hardwareproblemen vertoonde tijdens de afname.

Na beëindiging van de dataverzameling (juli 1995) hebben de participerende leerkrachten allen een rapportage ontvangen waarin de resultaten van hun groep beschreven zijn. Bij de leerkrachten die meegedaan hebben aan de computergestuurde dataverzameling (211 groepen) werd eveneens een papieren vragenlijst meegestuurd waarin gevraagd werd naar bevindingen van het computergebruik bij het onderzoek. In deze vragenlijst werd gevraagd naar de voorkeur van de leerkracht na afloop en hun ervaring met dit specifieke onderzoek. Ook werd navraag gedaan naar de reacties van hun leerlingen. Van de 211 leerkrachten hebben er 195 de vragenlijst compleet ingevuld teruggestuurd.

Data kwaliteit

We zullen nu Disk By Mail, uitgevoerd in het basisonderwijs, op een aantal punten kritisch beschouwen. In de eerste plaats gaan we, naar aanleiding van een korte evaluatie onder leerkrachten die DBM hebben uitgevoerd, dieper in op de bevindingen van het gebruik van DBM in het onderwijs. Daarna volgt een vergelijking tussen de papieren versie en de computergestuurde versie, waarbij we de kwaliteit van de data nader beschouwen. Uiteindelijk geven we een kostenvergelijking tussen beide methoden.

Na de afronding van de dataverzameling is de deelnemende leerkrachten gevraagd de computergestuurde afname te evalueren. Zoals aangegeven, is hierbij een papieren versie van de vragenlijst gebruikt om selectieve uitval te vermijden van de groep leerkrachten die negatieve ervaringen hadden met de computergestuurde methode.

De resultaten van deze evaluatie waren bemoedigend. Leerkrachten waren positief over het gebruik van DBM bij onderzoek. Volgens de leerkrachten ervoer gemiddeld genomen 68% van de leerlingen het invullen van de vragenlijst per computer als erg leuk. Daarnaast had 45% van de leerkrachten het idee dat het gebruik van DBM bij onderzoek naar een sensitief onderwerp als pesten minder belastend en stressvol is voor de kinderen, terwijl 52% aangaf dat de methode niet verschilt wat betreft emotionele stress met een papieren versie. Het gebruik van de computer leverde geen extra problemen op, zoals technische problemen of problemen met het lezen van vragen vanaf het scherm en dergelijke. Indien er problemen ontstonden, waren dit problemen met de inhoud van de vragen zelf: dezelfde problemen die bij een papieren versie gewoonlijk naar voren komen.

Correlaties zijn berekend tussen achtergrondvariabelen op leerkrachtniveau en de antwoorden op de evaluatie. Tegen de verwachting in kan worden geconstateerd dat de oudere leerkrachten met weinig computerervaring ietwat minder problemen aangaven en wat positiever waren ten aanzien van DBM dan jongere leerkrachten. Ook in de lagere groepen met jongere kinderen werd meer enthousiasme bij het gebruik van de computer aangetroffen. Overige achtergrondvariabelen, zoals denomi-

natie, sekse van de leerkracht, opleiding van de leerkracht, school- of klassegrootte correleerden niet met positieve of negatieve ervaringen met DBM. Deze resultaten zijn stimulerend voor onderzoekers die een DBM-survey in het onderwijs zouden willen uitvoeren.

Een vergelijking is gemaakt tussen de resultaten, gevonden met de DBM-vragenlijst en de resultaten van de papieren versie. De DBM vragenlijst is in 245 groepen (=5872 kinderen) afgenomen. De papieren versie is in 18 groepen (=556 kinderen) afgenomen, die naar verhouding vrij groot van omvang bleken. De groepen waren vergelijkbaar wat betreft achtergrondvariabelen op leerkrachtniveau (bijvoorbeeld aantal jaren ervaring, opleiding, groepsniveau).

Bij vergelijking werd een veel groter percentage missing values op itemniveau gevonden bij de papieren versie ($p < .001$). In de DBM-groep werd een gemiddeld percentage van 5.7 gevonden, terwijl in de papieren conditie sprake was van gemiddeld 14.1% missing values. Opvallend hierbij was dat de corresponderende standaarddeviaties zeer varieerden. Zo werd bij de DBM-conditie een standaarddeviatie van 3.4 gevonden, terwijl deze bij de papieren conditie 25.0 was. Niet alleen het gemiddelde aantal missing values op itemniveau is lager bij de DBM-conditie, maar eveneens de individuele variabiliteit. Dit kan veroorzaakt zijn doordat kinderen die niet geheel zorgvuldig en geconcentreerd zijn bij de papieren versie makkelijker een vraag of soms een hele pagina kunnen overslaan, terwijl deze kinderen bij een computergestuurde versie gedwongen worden de vragen één voor één van het beeldscherm te lezen.

In de vragenlijst is een schaal, bestaande uit 9 items, voor het meten van sociale onwenselijkheid opgenomen (Van der Wolf, 1990). Een lage score op deze items betekent dat het betreffende kind geneigd is sociaal wenselijk te antwoorden. Een significant verschil is gevonden tussen beide condities ($p < .001$); de kinderen van de DBM-conditie waren ietwat minder geneigd sociaal wenselijk te antwoorden ($m = 30.6$, $s.d. = 5.1$) in vergelijking met de kinderen die de papieren versie ingevuld hebben ($m = 29.9$, $s.d. = 5.5$). Een aantal kinderen uit de DBM-conditie gaf na de afname aan het gevoel te hebben de antwoorden 'ergens in te stoppen,

waar niemand zo in kon bladeren en kijken'. Dit zou er op kunnen duiden dat computergebruik bij dataverzameling ten aanzien van gevoelige onderwerpen ook door kinderen als minder bedreigend wordt beschouwd.

Dit komt wederom naar voren bij een vergelijking van de beide condities op de geneigdheid van leerlingen open en eerlijk te antwoorden op de actor- en slachtofferitems. Een significant verschil werd gevonden tussen de beide condities. De kinderen die de computergestuurde vragenlijst ingevuld hebben, geven aan vaker te pesten dan de kinderen die de papieren versie ingevuld hebben ($p < .001$). De gemiddelden waren 30.5 (s.d. 9.6) voor de DBM-conditie en 27.7 (s.d. = 9.2) voor de papieren conditie. Daarnaast gaven de kinderen bij de DBM-conditie aan vaker slachtoffer te zijn van pesten dan de kinderen die een papieren versie ingevuld hebben ($p < .001$). Hun gemiddelde score op de slachtofferitems was 26.4 (s.d. = 9.8), terwijl dit bij de papieren conditie 23.1 (s.d. = 8.6) was.

Beargumenteed zou kunnen dat dit verschil op de actor- en slachtofferschaal tussen beide groepen daadwerkelijk bestaat. De mogelijkheid voor het toetsen van deze hypothese was voorhanden door de sociale positie van het kind in de klas, zoals gemeten met de SAGS, te vergelijken met de gevonden uitkomsten op de actor- en slachtofferschaal (Van Hattum, 1997). Aangezien deze sociometrische data bij alle kinderen verzameld is via de traditionele papieren versie kunnen de resultaten niet toegeschreven worden aan een methode-effect.

In de literatuur wordt aangegeven dat slachtoffers van pesten of agressie kinderen zijn die niet populair zijn en een verworpen positie in de groep hebben. De daders zijn eveneens niet erg populair, maar bevinden zich ook niet in de verworpen groep. Zij hebben een controversiële positie in de groep (Perry, Kusel & Perry, 1988; Hox & Van Hattum, 1996).

Wanneer we nu naar de verschillen tussen beide condities kijken, constateren we dat de DBM-leerlingen ietwat populairder zijn volgens de SAGS dan de kinderen die de papieren versie ingevuld hebben (DBM-leerlingen: $m = 2.3$, s.d. = .55; papieren versie: $m = 2.2$, s.d. .41; $p < .001$). Dus ondanks dat de kinderen die de DBM-versie ingevuld hebben een fractie

populairder zijn, geven deze kinderen eveneens aan meer gepest te worden. Dit duidt eveneens op een grotere mate van openheid en eerlijkheid bij het beantwoorden van een computergestuurde vragenlijst. Hiermee worden de gevonden resultaten ten aanzien van de geneigdheid van kinderen open en eerlijk te antwoorden nogmaals bevestigd.

Naast de kwaliteit van de data, vormen de kosten een belangrijk aspect bij de keuze voor een methode. Het vergelijken van kosten is een lastig probleem. Om toch te proberen een zo goed mogelijk beeld te schetsen van de kosten, hebben we getracht deze te berekenen en vervolgens te vergelijken met de kosten die we gemaakt zouden hebben indien het onderzoek met behulp van papieren versies van de vragenlijst uitgevoerd zou zijn. De kosten van de steekproeftrekking en van het operationaliseren van de concepten in items zijn niet meegeteld, aangezien deze voor beide condities gelijk zijn. Bij de DBM-conditie zijn de kosten meegerekend voor de aanschaf van het Ci3-programma, van de gebruikte diskettes, het programmeren, de helpdeskbeanning en de verzendkosten. Bij het berekenen van de kosten van de papieren conditie zijn de kopieerkosten, de (goedkoopste) verzendkosten en de kosten voor het invoeren en coderen van data meegerekend. Dit resulteerde in een bedrag van f 1.62 per ingevulde vragenlijst voor de DBM-procedure. Dit bedrag zou f 5.15 per ingevulde papieren vragenlijst zijn geweest (zie voor gedetailleerde informatie hieromtrent Van Hattum & De Leeuw, 1996)

Samenvatting

Resumerend kan geconstateerd worden dat het zeer goed mogelijk is een Disk By Mail onderzoek uit te voeren in Nederlandse basisscholen. Kinderen vanaf ongeveer 8 jaar blijken prima in staat een computergestuurde vragenlijst in te vullen en blijken dit een leuke methode te vinden. Vervolgens kunnen we constateren dat de kwaliteit van de data, verzameld met de computergestuurde methode, beter is dan die van de papieren versie. Als laatste valt een kostenvergelijking van beide methoden uit in het voordeel van de Disk By Mail procedure.

2.2 Studie 2: Een gemengd CAPI-CASI interview bij jongeren met een visuele beperking

In september 1994 is een onderzoek gestart naar de betekenis van het persoonlijk netwerk en sociale steun voor blinde en slechtziende jongeren. Het verzoek om dit onderzoek uit te voeren was afkomstig van de FOVIG, de Federatie van Ouders van Visueel Gehandicapten. De FOVIG stelt zich ten doel de ouderorganisaties te helpen hun doelstellingen te bereiken door het bundelen van hun krachten. De belangrijkste taken van de federatie zijn belangenbehartiging, onderlinge hulp van ouders en informatievoorziening. Aanleidingen tot dit onderzoek waren onder meer de resultaten van een eerdere studie naar de leefsituatie van visueel gehandicapten van 19 tot en met 32 jaar (Habekothé en Peters, 1993), waaruit bleek dat veel respondenten knelpunten in hun dagelijks leven ondervonden, zoals acceptatie van de visuele beperkingen, afhankelijkheid en problemen rondom sociale contacten.

Het belang van een vervolgonderzoek naar het psychosociaal functioneren en de betekenis van sociale contacten hierbij, werd bevestigd door aanwijzingen in de literatuur en door ervaringen van hulpverleners en begeleiders van blinde en slechtziende jongeren. Doel van het onderzoek is dat met behulp van de resultaten ouders en opvoeders van kinderen met een visuele beperking hun kinderen beter kunnen ondersteunen bij het verwerven van een zelfstandige maatschappelijke positie. Tevens kunnen aanbevelingen gedaan worden aan instellingen ten behoeve van opvoeding van en onderwijs aan blinde en slechtziende kinderen en jongeren. In dit artikel wordt slechts de dataverzameling beschreven; voor onderzoeksresultaten wordt verwezen naar Kef (1997a), Kef, Hox & Habekothé (1997) en Kef, Habekothé en Hox (1998).

Logistieke aspecten

Alle jongeren met een visuele beperking van 14 tot 24 jaar in Nederland zijn via het speciaal onderwijs en diverse instellingen benaderd om mee te doen aan dit onderzoek. Voor deze leeftijd is gekozen omdat in deze periode jongeren een identiteitsontwikkeling doormaken en hierbij een beroep gedaan wordt op hun relaties. Tevens wordt in deze periode de wens om bij een groep te horen sterker en zal het gevoel

'anders' te zijn door de visuele beperkingen mogelijk (opnieuw) een rol gaan spelen. Voor de vergelijking van de onderzochte groep met niet-gehandicapte jongeren werd geen aparte dataverzameling ondernomen. In plaats daarvan werd gebruik gemaakt van de resultaten van onderzoeksprojecten aan de Universiteit Utrecht (Rispen, Hermanns & Meeus, 1996) en de Rijksuniversiteit Leiden (Buysse, 1997).

Vanwege het persoonlijke en gevoelige karakter van de vragen, bijvoorbeeld over ervaren sociale steun en eenzaamheid, is gekozen voor een face-to-face-interview. Mede gezien de visuele beperkingen, de lengte en complexiteit van de vragenlijsten, hadden een schriftelijke vragenlijst of een telefonisch interview duidelijke nadelen. Bovendien is de datakwaliteit bij sensitieve complexe vragen beter bij een face-to-face-interview (De Leeuw, 1992). Bij het persoonlijk interview werd een draagbare computer gebruikt; er was dus sprake van een Computer Assisted Personal Interview (CAPI). In het onderhavige onderzoek naar onder meer persoonlijke netwerken, zijn de vragen dusdanig complex en gebaseerd op voorgaande antwoorden, dat computergestuurd interviewen duidelijk voordelen biedt.

Voor de vergelijking van de resultaten van de jongeren met een visuele stoornis met die van ziende jongeren, is een deel van de instrumenten overgenomen uit de onderzoeksprojecten van Rispen, Hermanns en Meeus (1996) en Buysse (1997). Van het eerste project zijn dat: de Rosenberg Self-Esteem Questionnaire, de Cantrill-ladder, de Utrechtse Coping Lijst verkorte versie en de Persoonlijke Netwerk Lijst. Uit het tweede project zijn overgenomen: de Social Network Map en de Social Network Grid. Deze vragenlijsten zijn omgezet in een computergestuurde vragenlijst. Het betreft hier geen inhoudelijke wijzigingen, alleen programmeertaal is toegevoegd¹. Geheugensteuntjes voor de interviewers, bijvoorbeeld over het gebruik van de helpfunctie, werden tevens geprogrammeerd zodat deze op het beeldscherm verschenen. Vervolgens is de computergestuurde versie geïmplementeerd op alle laptops en uitgebreid gecheckt. Iedere interviewer bezat tevens een disketteversie van de vragenlijst, die fungeerde als back-up. Mocht de vragenlijst op de laptop niet functioneren, dan kon de disketteversie gestart worden. Als de

laptop zelf problemen gaf, kon op de personal computer van de respondenten via de diskette-versie het interview toch afgenomen worden. Vrijwel alle blinde en slechtziende jongeren hebben een computer en gebruiken deze veelvuldig. Een belangrijk punt voor de vergelijkingsmogelijkheden was dat de overgenomen instrumenten uit het project van Rispens et al. (1996) in hun onderzoek schriftelijk werden afgenomen. Het was van essentieel belang dat verschillen in afnamevorm en mogelijkheden tot het geven van sociaal wenselijke antwoorden voorkomen werden. In ons onderzoek hebben de blinde of slechtziende jongeren daarom bij deze onderdelen, die samen 20% van de vragenlijst vormen, de antwoorden zelf ingetypt (het betreft in alle gevallen het intoetsen van één cijfer). Hiervoor waren specifieke aanpassingen noodzakelijk in verband met de visuele beperkingen van de respondenten. Zij konden de tekst op het beeldscherm niet of moeilijk lezen, zodat de interviewer de betreffende items en antwoordmogelijkheden van deze instrumenten uit een veldgids moest voorlezen. De respondent toetste zelf het antwoord in. Deze methode kan gekenmerkt worden als CASI, Computer Assisted Self Interviewing. Om het intypen door blinde en slechtziende jongeren mogelijk te maken, is een kartonnen aanpassing met braille en vergrote cijfers voor op het toetsenbord ontwikkeld. Tevens werd een geluidssignaal geprogrammeerd dat hoorbaar werd bij het indrukken van een toets. Dit geluidssignaal

was noodzakelijk voor de interviewer. Deze wist op die wijze wanneer de volgende vraag gesteld kon worden. Ter ondersteuning van het geheugen van de respondenten, voornamelijk bij de gestandaardiseerde vragenlijsten die vele antwoordcategorieën bezitten, zijn antwoordkaarten in braille en vergrotingen gebruikt. De uiteindelijk ontwikkelde methode in dit onderzoek kan dus een gemengd CAPI-CASI model genoemd worden.

Zestien interviewers, allen (ex)orthopedagogiekstudenten, hebben een driedaagse interviewtraining gevolgd. In deze training kwamen, naast het trainen van algemene interviewtechnieken, ook rollenspelen aan bod voor het oefenen van specifieke vaardigheden met betrekking tot het computergestuurd interviewen. Alle blinde en slechtziende jongeren van 14 tot 24 jaar in Nederland werden, via scholen en andere instellingen voor blinden en slechtzienden, gevraagd om hun medewerking. Van deze groep heeft 37% aan het onderzoek meegedaan. De jongeren met een visuele beperking (N = 316) zijn éénmaal in hun woonsituatie geïnterviewd, zonder aanwezigheid van anderen. Gedurende de dataverzameling waren er zes terugkomdagen voor de interviewers, waarop vragen, onduidelijkheden en ervaringen gedeeld werden. Tevens was het mogelijk telefonisch een fieldmanager te consulteren, twaalf uur per dag en zeven dagen per week. De dataverzamelingsfase heeft in totaal vijf maanden (maart-juli 1996) in beslag genomen.

Tabel 2
Betrouwbaarheid van de schalen (N=316) (Kef, 1997b)

Schaal	Alpha (N = 316)	Alpha ander onderzoek
Zelfwaardering	0.83	0.85 ¹
Locus of control-1	0.35	0.72 ²
Locus of control-2	0.60	0.69 ³
Coping strategieën, problem focussed	0.76	0.77 ⁴
Coping strategieën, emotion focussed	0.67	0.72
Coping strategieën, vermijden	0.66	0.53
Handicap-acceptatie	0.83	0.91 ⁵
Eenzaamheid	0.78	0.81 ⁶
Sociale wenselijkheid	0.59	0.60 ⁷

1 Rispens, Hermanns & Meeus (1996)

2 Dodds, Bailey, Pearson & Yates (1991)

3 Peetsma (1992)

4 Rispens, Hermanns & Meeus (1996)

5 Dodds, Bailey, Pearson & Yates (1991)

6 De Jong-Gierveld & Kamphuis (1985)

7 Lagerweij (1995)

Datakwaliteit

Aangezien bij alle respondenten een computer-gestuurde afname heeft plaatsgevonden, is een vergelijking van de resultaten met die van een eventuele papieren versie bij deze respondenten niet mogelijk. Echter, andere methoden om de datakwaliteit te kunnen beoordelen zijn wel mogelijk. Om de kwaliteit van de data te onderzoeken, is gekeken naar het voorkomen van missing data, de betrouwbaarheid van schalen en de interviewervariantie. De computer-gestuurde interviewmethode werd geëvalueerd bij zowel respondenten als interviewers.

Wat betreft het eerste aspect met betrekking tot datakwaliteit kan worden geconcludeerd dat in dit onderzoek geen missing data zijn. Dit is veroorzaakt door de geprogrammeerde route bij het computergestuurd interviewen. Het gevolg hiervan is dat bij de statistische analyses minder problemen zijn voorgekomen.

De betrouwbaarheid van de schalen is het tweede aspect voor datakwaliteit. De antwoorden op de gestandaardiseerde schalen, bijvoorbeeld de Rosenberg Self-Esteem Scale, werden geanalyseerd op hun betrouwbaarheid. Voor elke schaal werd Cronbachs alpha berekend, voor zowel de gehele groep adolescenten als voor subgroepen (bijvoorbeeld blinde versus slechtziende adolescenten). Verwacht werd dat het intikken van de antwoorden bij het CASI-onderdeel voor de blinden moeilijker was, wat zou kunnen leiden tot minder consistente antwoorden.

Voor de gehele groep adolescenten en voor subgroepen was de betrouwbaarheid van de meeste schalen voldoende tot goed. Zoals blijkt uit tabel 2 is de betrouwbaarheid van de locus of control schaal-1, afkomstig van de Nottingham Adjustment Scale (Dodds, Bailey, Pearson & Yates, 1991) niet voldoende, ook niet na itemanalyse. De resultaten op deze schaal zijn dan ook niet gebruikt. De betrouwbaarheid van de schalen in ons onderzoek was enigszins lager dan de vermelde waarde van de oorspronkelijke instrumenten. Dit kan veroorzaakt zijn door de methode, een face-to-face-interview, waarin de betrouwbaarheid vaak lager is vergeleken met die in schriftelijke vragenlijsten (De Leeuw, 1992). De verklaring hiervoor wordt gezocht in het feit dat bij een face-to-face-interview het overzicht bij de respondenten ontbreekt van eerder door hen gegeven antwoorden.

Zijn er interviewereffecten opgetreden op de variabele netwerk-grootte? In ander netwerk-onderzoek werden grote interviewereffecten aangetoond met betrekking tot het afbakenen van persoonlijke netwerken (Van Tilburg, 1998). Ook dit werd geanalyseerd voor de gehele groep respondenten en tevens voor subgroepen, omdat de blinde jongeren wellicht meer assistentie nodig zouden kunnen hebben. Voor de gehele groep en voor subgroepen werden geen interviewereffecten gevonden.

Uit de evaluaties, gehouden direct na afloop van het interview, kwam naar voren dat de ervaringen van de meeste jongeren met het gemengde model CAPI/CASI, zeer positief zijn. Ze vonden het een goed interview en leuk om aan mee te doen. De antwoordkaarten die gebruikt werden ter ondersteuning van het geheugen, bevielen prima. De mogelijkheid om zelf antwoorden in te tikken bij met name sensitieve onderwerpen, viel bij iedereen bijzonder goed in de smaak. Het was een leuke afwisseling, maar ook voor het gevoel van privacy vonden de respondenten het bij die onderdelen belangrijk. De computer was niet beangstigend, het tegendeel is waar. Blinde en slechtziende jongeren gebruiken dagelijks een computer voor de informatieverwerking, bijvoorbeeld huiswerk en kranten, en zijn dus zeer ervaren op dit gebied. Eén blinde respondent wist zelfs aan het bromgeluid van de laptop het merk te herkennen.

De interviewers waren tevens overwegend positief. Omdat er geen fouten gemaakt kunnen worden in de volgorde van de vragen, kan men zich goed concentreren op de inhoud van de vragen en de communicatie met de respondent. Het omgaan met het interviewprogramma en de laptop gaf geen problemen. Als nadeel werd door de interviewers het gemis aan overzicht genoemd, bijvoorbeeld hoe ver gevorderd het interview was op een zeker moment. In een vervolgstudie kan dit echter opgelost worden door dit te programmeren. Zo kan duidelijk worden gemaakt op het beeldscherm dat men bezig is met onderdeel drie uit in totaal tien onderdelen. Tevens is aan interviewers gevraagd: 'Wat is er speciaal aan het computergestuurd interviewen van blinde en slechtziende jongeren?' Uit de resultaten kwamen elf clusters of concepten naar voren. De drie belangrijkste concepten of aspecten in het

(computergestuurd) interviewen van blinde en slechtziende adolescenten zijn:

- de noodzaak om elke actie of gebeurtenis verbaal uit te leggen aan de respondenten ('Ik moet even een paar commando's intikken voordat we kunnen starten');
- het belang van paralinguïstische en verbale communicatiekanalen tijdens het interview (knikken wordt niet gezien, hummen wel gehoord);
- de houding van de interviewer tegenover een persoon met een beperking (gelijkwaardig tegemoet treden en niet medelijdend).

Om data te verkrijgen met een hoge kwaliteit, is het noodzakelijk om vooraf aan deze aspecten veel aandacht te besteden, zoals in dit onderzoek naar tevredenheid van de interviewers ook gedaan is. Geconcludeerd kan worden dat CAPI en CASI door zowel de respondenten als de interviewers geaccepteerd en leuk gevonden worden.

Samenvatting

Een gemengd CAPI-CASI model kan met succes gebruikt worden in onderzoek bij jongeren met een visuele beperking. Op basis van de ervaringen van respondenten, interviewers en de resultaten wat betreft datakwaliteit, heeft deze methode vele voordelen getoond. Ook andere vormen van CADAC zouden succesvol gebruikt kunnen worden bij deze doelgroep, bijvoorbeeld Disk By Mail of Computer Assisted Telephone Interviewing. De laatst genoemde methode is overigens in de afgelopen jaren al tweemaal met succes toegepast bij blinde en slechtziende adolescenten (Wesselink, 1996; Reek & Sloep, in voorbereiding). Vanwege de visuele beperkingen van de respondenten zijn soms specifieke aanpassingen nodig, zoals antwoordkaarten in braille, en moeten bepaalde interviewvaardigheden in de training voldoende aandacht krijgen.

3 Conclusie en aanbevelingen

Uit twee landelijke grootschalige onderzoeken, een onder kinderen van 8 tot 12 jaar en een onder jongeren met een visuele beperking, blijkt dat de resultaten met computergestuurd dataverzamelen zeer positief zijn. Zowel de kinderen als de adolescenten zijn positief over

de gebruikte methode. Ook interviewers kunnen er goed mee omgaan en ervaren de vele voordelen ervan. De methode is effectief, snel, makkelijk en leuk. Bovendien zijn er ook belangrijke voordelen voor de onderzoekers: het invoeren en verwerken van de data verloopt vele malen sneller en effectiever, de methode is goedkoper, de datakwaliteit is hoog en is vaak hoger vergeleken met andere methoden.

Natuurlijk is er een aantal belangrijke aandachtspunten bij het gebruik van CADAC. Zo is het van essentieel belang dat de opzet van de computergestuurde dataverzameling wordt aangepast aan de doelgroep. Aspecten zoals de leesvaardigheid en oog-handcoördinatie moeten aandacht krijgen, maar ook specifieke aanpassingen in het geval van een zintuiglijke of andersoortige beperking van de respondent. Tevens vraagt de voorbereiding van de computergestuurde dataverzameling meer tijd dan bij een schriftelijke vragenlijst of een gewoon persoonlijk interview. De geprogrammeerde vragenlijst moet uitgebreid getest worden in pilot-studies en de organisatie van de dataverzameling vraagt om een zeer systematische aanpak (De Leeuw, Hox, Kef & Van Hattum, 1997). Bij de voorbereiding hoort ook een intensieve training van interviewers en andere personen die bij de computergestuurde dataverzameling betrokken zijn. Naast de gebruikelijke interviewtraining, gericht op gesprekstechnieken en administratieve vaardigheden, is het noodzakelijk om aandacht te besteden aan het omgaan met de laptop of computer en het computerprogramma. Verder is de ondersteuning tijdens het verzamelen van de data belangrijk. Een gedetailleerde handleiding, verschillende back-upsystemen, terugkomdagen en telefonische ondersteuning (helpdesk) zijn hiervan enkele voorbeelden.

Gezien de maatschappelijke ontwikkelingen op het gebied van computers zou deze methode veel meer, bij kinderen en jongeren, gebruikt kunnen worden. Vele toepassingen in wetenschappelijk onderzoek, praktijkonderzoek, onderwijs en begeleiding van kinderen en jongeren, met of zonder beperkingen, zijn denkbaar. In plaats van schriftelijke vragenlijsten sturen we diskettes, in plaats van een persoonlijk interview met behulp van formulieren met hun bijbehorende nadelen, gebruiken we een laptop. Met de nieuwste ontwikkelingen in de

multi-mediasystemen, zoals geluid (Audio-CASI) en video, zal de waarde van CADAC alleen maar toenemen, zeker bij gebruik met speciale doelgroepen.

Noot

- 1 Een technisch rapport met daarin de vragenlijst en een disketteversie van de vragenlijst, kan opgevraagd worden bij de Faculteit der Maatschappij- en Gedragswetenschappen van de Universiteit van Amsterdam. Tel. 020 - 525 15 66.

Literatuur

- Beckenbach, A. (1995). Computer assisted questioning: the new survey methods in the perception of the respondent. *BMS*, 82-100.
- Besag, V.E. (1989). *Bullies and victims in schools*. Milton Keynes: Oxford University Press.
- Buysse, W.H. (1997) *Personal social networks and behavior problems in adolescence*. Proefschrift. Leiden: Grafisch Bedrijf UFB.
- Defares, P.B., Kema, G.N., Praag, E. van & Werff, J.J. van der (1977). *SAGS. Syracuse Amsterdam/Groningen Sociometrische Schaal*. Lisse: Swets & Zeitlinger.
- Dodds, A.G., Bailey, P, Pearson, A. & Yates, L. (1991) Psychological factors in acquired visual impairment: the development of a scale for adjustment. *Journal of Visual Impairment and Blindness*. 85, 306 - 310.
- Habekothé, H.T. & Peters, M.J.G.F. (1993). *Leefsituatie van visueel gehandicapten in de leeftijd van 19 tot en met 32 jaar*. Amsterdam: Universiteit van Amsterdam.
- Harrel, T.H. & Lombardo, T.A. (1984). Validation of an automated 16PF administration procedure. *Journal of Personality Assessment*, 48, 216-227.
- Hattum, M.J.C. van (1997). *Pesten. Een onderzoek naar beleving, visie en handelen van leraren en leerlingen*. Dissertatie Universiteit van Amsterdam.
- Hattum, M.J.C. van & Leeuw, E.D. de (1996). *A Disk By Mail survey of teachers and pupils in dutch primary schools; logistics and data quality*. MLS-publicatie nr. 57. Universiteit van Amsterdam.
- Heenan, S.A. (1992). *The problems of carrying out small scale research into bullying in schools*. Unpublished Master thesis, University of East Anglia, School of Education.
- Higgins, C.A., Dimnik, T.P., & Greenwood, H.P. (1987). The DISQ survey method. *Journal of the Market Research Society*, 29, 437-445.
- Hox, J.J. & Hattem, M.J.C. van (1996). *De invloed van sociometrische positie op pesten in de klas. Modellen en analyseprocedures*. Amsterdam: CSO-Kohnstamm Instituut.
- Hutchinson, H. & Metegrano, M. (Eds.) (1991) *Ci3 System*. Evanston: Sawtooth Software, Inc.
- Jong-Gierveld, J. de & Kamphuis, F.H. (1985) The development of a Rasch-type Loneliness scale. *Applied Psychological Measurement*. 9, 3, 289-299.
- Kef, S. (1997a) The personal networks and social supports of blind and visually impaired adolescents. *Journal of Visual Impairment & Blindness*. volume 91, nr. 3, 236 - 244.
- Kef, S. (1997b) *Technisch rapport: de betekenis van het persoonlijk netwerk voor visueel gehandicapte jongeren*. MLS-publication, nr. 59, Universiteit van Amsterdam.
- Kef, S., Hox, J.J. & Habekothé, H.T. (1997) *(On)Zichtbare steun. Onderzoek naar visueel gehandicapte jongeren en hun netwerk*. Amsterdam: Thesis Publishers.
- Kef, S., Habekothé, H.T., & Hox, J.J. (1998) Jongeren met een visuele beperking: persoonlijk netwerk en welbevinden. *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs*, Vol. 14, nr. 4, 200 - 219.
- Lagerweij, N.W. (1995) *Milieuegedrag bij kinderen. Ontwikkeling van een instrument voor het meten van milieuegedrag*. Amsterdam: Proefschrift Universiteit van Amsterdam.
- Leeuw, E.D. de (1992) *Data quality in mail, telephone and face to face surveys*. Amsterdam: TT-Publikaties.
- Leeuw, E.D. de (1995) *Computer assisted data collection data quality and costs; A taxonomy and annotated bibliography*. MLS-publication, nr. 55, Universiteit van Amsterdam.
- Leeuw, E.D. de & Collins, M. (1997). Data collection method and survey quality: An overview. In: L. Lyberg et al. (eds). *Survey measurement and process quality*. New York: Wiley.
- Leeuw, E.D. de, Hox, J.J. & Snijders, G. (1995) The effect of computer-assisted interviewing on data quality. A review. *Journal of the Market Research Society*, 37, 4, 325-344.
- Leeuw, E.D. de, Hox, J., Kef, S. & Van Hattum, M. (1997). *Overcoming the problems of special interviews on sensitive topics: computer assisted*

- self-interviewing tailored for young children and adolescents. MLS-publication, nr. 60. Universiteit van Amsterdam.
- Leeuw, E.D. de, Snijders, G. & Hox, J.J. (1994). *The effect of computer assisted interviewing on data quality. A review.* MLS-publication, nr. 54. Universiteit van Amsterdam.
- Nicholls, W.L., Baker, R.P. & Martin, J. (1997). The effect of new data collection technologies on survey data quality. In: L. Lyberg et al. (eds). *Survey measurement and process quality.* New York: Wiley.
- Olweus, D. (1992). *Treiteren op school.* Amersfoort: College Uitgevers.
- Olweus, D. (1993). *Bullying at school. What we know and what we can do.* Oxford: Blackwell.
- Peetsma, T.T.D. (1992) *Toekomst als motor? Toekomstperspectieven van leerlingen in het voortgezet onderwijs en hun inzet voor school.* Amsterdam: SCO, Stichting Kohnstamm Fonds voor Onderwijsresearch.
- Perry, D.G., Kusel, S.J. & Perry, L.C. (1988). Victims of peer aggression. *Developmental Psychology*, 24, 807-814.
- Ree, B. van der (1994). *The Comenius project for primary and special education.* Zoetermeer: Ministerie van OCenW.
- Reek, S. & Sloep, A. (in voorbereiding) *Liefde is blind. Onderzoek naar dating en seksualiteit van blinde jongeren.* Afstudeerscriptie. Amsterdam: Universiteit van Amsterdam.
- Rispens, J., Hermanns, J.M.A. & Meeus, W.H.J. (1996) *Opvoeden in Nederland.* Assen: Van Gorcum.
- Sawtooth, (1994). *Ci3 User Manual* (edited by Harla Hutchinson & Margo Metegrano). Evanston: Sawtooth Software.
- Tilburg, T. van (1998) Interviewer effects in the measurement of personal network size. A non-experimental study. *Sociological methods & research*, Vol. 26, No. 3, p.300-328.
- Tourangeau, R. & Smith, T.W. (1996). Asking sensitive questions; the impact of data collection, question format, and question context. *Public Opinion Quarterly*, 60, 275-304.
- Weisband, S. & Kiesler, S. (1996). Self disclosure on computer forms: Meta analysis and implications. *CHI '96* (<http://www.al.arizona.edu/~weisband/chi/chi96.html>).
- Wesselink, P. (1996) *Opvoedingsbeleving van visueel gehandicapte jongeren.* Afstudeerscriptie. Amsterdam: Universiteit van Amsterdam.
- Witt, K.J. & Bernstein, S. (1992). Best practices in Disk-by-Mail surveys. *Proceedings of the 1992 Sawtooth conference*, Sawtooth software: Evanston, Illinois.
- Wolf, J.C. van der (1990). *Effectief onderwijs op kleurrijke scholen.* Delft: Eburon.

Auteurs

M.J.C. van Hattum is momenteel als docent en onderzoeker verbonden aan de Universiteit Utrecht, Vakgroep Pedagogiek.

Adres: Universiteit Utrecht, Faculteit Sociale Wetenschappen, Heidelberglaan 1, 3584 CS Utrecht.

Email: M.vanhattum@fsw.ruu.nl

S. Kef is als Assistent in Opleiding (AiO) en docent werkzaam bij de Faculteit der Maatschappij- en Gedragswetenschappen van de Universiteit van Amsterdam, afdeling Pedagogiek.

Adres: Faculteit FMG, Wibautstraat 4, 1091 GM Amsterdam. Email: sabina@educ.uva.nl

Abstract

Computer Assisted Data Collection in case of pupils in primary education and visually impaired adolescents

M.J.C. van Hattum & S. Kef, Pedagogische Studiën, 1999, 76, 104-116.

In social sciences questionnaires like interviews and selfreports are often used. These methods have many advantages, especially when sensitive questions are asked. However, questionnaires have a serious draw-back: only relatively simple questionnaires can be used. Computer Assisted Data Collection (CADAC) can overcome these problems. After a resumé of literature on computer assisted data collection, several forms of CADAC are presented. Two studies are described in which the CADAC-method is used. In the first study, a Disk-by-Mail survey about bullying in primary education has been implemented. The second study consists of a survey about personal networks, social support and well-being of visually impaired adolescents. A combination of Computer Assisted Personal Interviewing and Computer Assisted Self Interviewing is used. The description of both studies consists of general logistics and data quality. The article is finished with recommendations for the adaptation of CADAC for children and adolescents.

Kwalitatieve analyse van interviews

Fasen, stappen en computergebruik in intensief onderzoek

F. Wester en V. Peters

Samenvatting

In dit artikel wordt een beschrijving gegeven van de opzet en uitvoering van een kwalitatieve analyse van interviewmateriaal. Allereerst wordt de kenmerkende cyclus: waarnemen – analyseren – reflecteren, die meer malen wordt doorlopen, toegelicht. Daarna worden de fasen van de kwalitatieve analyse vanuit het perspectief van de gefundeerde theoriebenadering nader uitgewerkt, waarbij steeds de belangrijkste stappen en de knelpunten en valkuilen die tijdens die fase dreigen, worden toegelicht. Daarbij wordt aan de hand van het computerprogramma Kwalitan geïllustreerd op welke wijze de computer kan worden ingeschakeld ter ondersteuning van deze analyse.

Inleiding

De afgelopen tien jaar is er een sterke ontwikkeling geweest in de beschikbare programma's voor het manipuleren van tekstachtig materiaal. Daarmee is de grootste barrière weggenomen om de computer in kwalitatief onderzoek als hulpmiddel in te schakelen. Dat is zeker het geval voor het kwalitatief onderzoek uit de interpretatieve onderzoekstraditie. Kenmerkend voor dit soort kwalitatief onderzoek is het arbeidsintensieve karakter van de iteratieve, interpreterende analyse van het tekstmateriaal. In deze bijdrage willen wij de systematiek van de kwalitatieve analyse beschrijven en de rol van de computer daarbij toelichten. Eerst zullen wij ingaan op de aard van de kwalitatieve analyse. Vervolgens geven wij een schets van de algemene functies die de computer daarbij kan vervullen. Daarna gaan we specifiek in op de systematiek van waarneming en analyse en de rol van de computer daarbij in onze uitwerking van de *gefundeerde theoriebenadering* (vgl. Wester, 1995). Het gaat hier veelal om vergelijkend onderzoek op grond

van een relatief flink aantal interviews, ook wel aangeduid als kwalitatief survey (Hijmans, 1995, p. 519; Kuyper, 1993; Hijmans, 1994). Ten slotte zullen enkele conclusies getrokken worden over het gebruik van de computer in de kwalitatieve analyse.

1 De analyse in het kwalitatieve onderzoek

De analyse¹ is lange tijd het stiefkindje van het kwalitatief onderzoek geweest. Juist op dit gebied zijn er de laatste vijftien jaar veel publicaties verschenen en zijn er recentelijk nieuwe ontwikkelingen, die samenhangen met de rol van de computer in de analyse (Fielding & Lee, 1991; Miles & Huberman, 1993; Weizman & Miles, 1994; Peters & Wester, 1995).

1.1 Kenmerken van de analyse

Voor het kwalitatieve onderzoek zijn twee betekenissen van het begrip analyse relevant. Ten eerste gaat het bij analyse om uiteenleggen, dat wil zeggen dat wordt geëxpliciteerd welke aspecten van het onderzochte verschijnsel relevant zijn. Die aspecten komen naar voren in het analytisch kader van de onderzoeker. In kwalitatief onderzoek moet dat kader veelal aan de hand van het onderzoek worden uitgewerkt. Ten tweede gaat het bij analyse om het zoeken naar patronen in het materiaal, dat wil zeggen de samenhang tussen de onderscheiden aspecten rond een verschijnsel. Via analyse wil men een overzichtsbeld krijgen van een verschijnsel en nagaan welke factoren met de aspecten van het verschijnsel samenhangen.

Beide betekenissen van analyse verwijzen naar de rol van het analytisch kader van de onderzoeker. Omdat in kwalitatief onderzoek dit kader veelal nog moet worden uitgewerkt en/of toegespitst op het onderzoeksveld, is analyse sterk met het ontwikkelen van het analy-

tisch kader verbonden. De analyse is dan ook niet één aparte fase in het onderzoek: zij is een steeds terugkerend moment van de spiraalvormige cyclus: reflectie – waarneming – analyse – reflectie, die de onderzoeker tijdens het onderzoek voortdurend doorloopt en waarbij het analytisch kader steeds verder uitgewerkt wordt. In reflectiemomenten wordt nagedacht over onderzoeks- en analysevragen en over het begrippenkader om die te beantwoorden. In de waarneming wordt het materiaal geproduceerd waarin die antwoorden naar voren moeten komen.

De aard van deze handelingen en de ontwikkeling daarin in de loop van het onderzoek brengen met zich mee dat de onderzoeker behoefte heeft aan hulpmiddelen waarin een en ander overzichtelijk kan worden uitgevoerd. Een van de belangrijkste hulpmiddelen voor het reflectieproces zijn de *memo's* waarin de onderzoeker ontwikkelingen vastlegt in begrippen en onderzoeksvragen. Daarnaast zal de onderzoeker het materiaal ordenen en bewerken door te coderen², samen te vatten en overzichten te produceren, waarbij de computer een rol kan spelen.

1.2 Analyse en de onderzoeksopzet

Men kan een aantal handelingen onderscheiden, waarvan sommige meer malen in het analyseproces terugkeren.

Een eerste handeling is het *formuleren van de probleemstelling*. In de probleemstelling wordt aangegeven welke aspecten van onderzoeksthema relevant worden geacht en eventueel welke context daarbij een rol speelt. Voor kwalitatief onderzoek betekent dit dat de onderzoeker een beredeneerde vraagstelling heeft uitgewerkt waarin het onderzoekersperspectief veelal in de vorm van *sensitizing concepts* naar voren komt. Dit is dus analysewerk, dat gebaseerd is op een theoretische doordienking van het onderzoeksthema, meestal aan de hand van de onderzoeksliteratuur. In de loop van het onderzoek zal de probleemstelling steeds weer aangepast worden, omdat er ontwikkelingen zijn in het analytisch kader waardoor de onderzoeksvragen explicieter kunnen worden geformuleerd. De probleemstelling is dan ook gedurende het gehele onderzoek een aandachtspunt voor de reflectieprocessen van de onderzoeker. De uiteindelijke probleemstel-

ling kan worden geformuleerd als het analytisch kader is uitgewerkt.

Een tweede handeling is het uitwerken van instrumenten voor de *gegevensverzameling*. Ook hier gaat het om aspecten van het analytisch kader: men wil immers gegevens verzamelen die relevant zijn voor de onderzoeksvragen. Het gaat hier om een aparte stap, omdat via de verkenning van het onderzoeksveld en de onderscheidingen die daarbij relevant zijn, *veldbetrokken begrippen* in het analytisch kader worden opgenomen. Bovendien houdt de gegevensverzameling altijd keuzen in wat betreft de te onderzoeken cases en de kenmerken die daarbij moeten worden vastgesteld. Ook hiervoor geldt dat de onderzoeker in de loop van het onderzoek steeds beter kan formuleren om welke aspecten het gaat en welke gegevens daarvoor verzameld moeten worden.

Een derde activiteit is *transcriptie* van het ruwe gegevensmateriaal. In kwalitatief onderzoek levert de waarneming meestal ruw en voor de onderzoeker ongestructureerd materiaal op, zoals op tape vastgelegde interviews, veldaantekeningen en archiefstukken. Op een of andere manier moet daarvan materiaal gemaakt worden dat geschikt is voor analyse. Bij transcriptie spelen weer keuzen een rol die consequenties hebben voor de mogelijkheden in de analyse (bijvoorbeeld nemen we naast de gesproken woorden ook aarzelingen en houdingen mee). Vooral bij observatiemateriaal zijn expliciete regels nodig omtrent het wat en hoe van de registratie van de geobserveerde verschijnselen.

Aan de transcriptie is tevens een activiteit verbonden die meer technisch van aard is: een *eerste ordening* van het materiaal, zodat het geschikt is voor de analyse-in-engere-zin. Tegenwoordig gaat het hier meestal om een zodanige bewerking van het ruwe materiaal, dat we de computer als hulpmiddel kunnen inschakelen. Zo vragen sommige programma's voor kwalitatieve analyse dat men het materiaal bijvoorbeeld indeelt in segmenten en zonder tekstverwerkingscodes bewaart.

Een vierde analysehandeling is dan de *analyse-in-engere-zin*. Globaal genomen gaat het hier om het lezen van een sample van het onderzoeksmateriaal. Hier zijn verschillende betekenissen van *lezen* van belang. Ten eerste gaat het natuurlijk om lezen in de zin van het

waarnemen van verbaal weergegeven betekenissen (zoals we een krant of boek lezen). Daarnaast gaat het om selectief lezen vanuit de onderzoeksvragen (c.f. aren lezen, uitgelezen waar). Bovendien gaat het om interpreterend lezen, dat wil zeggen dat we interpreteren vanuit het analytisch kader (c.f. kaartlezen, sterren lezen). Deze lezingen van het materiaal worden meestal uitgedrukt in codes of trefwoorden die bij het materiaal worden opgenomen. Via reflectie op vragen, materiaal en codes worden codes aangepast en nieuwe analysevragen geformuleerd. Deze worden dan weer op een volgende selectie uit het materiaal uitgeprobeerd, waarbij een nieuwe cyclus van materiaalselectie, lezen, interpreteren en reflecteren is ingezet. Dit cyclische proces gaat door tot de onderdelen van het analytisch kader zijn uitgewerkt.

Een vijfde analysehandeling zijn de *nadere analyses* die meestal gericht zijn op het vergelijken van cases via overzichten, tabellen en matrices om patronen in het materiaal te ontdekken.

De laatste analysehandeling die wij onderscheiden is de *rapportage*, een aandachtspunt in elke onderzoeksfase. Via tussentijdse verslagen kan de voortgang van het onderzoek worden bewaakt. Het schrijven van het eindverslag, waarin de beantwoording van de definitieve onderzoeksvraag wordt beargumenteerd, vormt de afronding van de analyse.

Door analyse op te vatten als een steeds terugkerende moment in een cyclisch proces wordt een aantal kenmerken van de kwalitatieve analyse duidelijk. Ten eerste wordt de analyse altijd gestuurd door de *vragen* die in het reflectieproces naar voren zijn gekomen en de observaties hebben gestuurd. In het verzamelde materiaal worden antwoorden op die vragen geformuleerd door middel van het voorlopige *analytisch kader*, dat daartoe eventueel wordt aangepast, en waarmee het materiaal wordt geordend en bewerkt. Ten tweede is de analyse in de loop van het onderzoek niet steeds op dezelfde vragen gericht. De spiraalvormige cyclus duidt een ontwikkeling in het onderzoek aan, waarbij de opgedane inzichten een nieuwe cyclus aansturen, waarbij vragen veranderen. Bovendien kan men de loop van het onderzoek uiteenleggen in een aantal fasen waarin achter-

eenvolgens een aantal subdoelen worden bereikt. Bij elk subdoel staat een ander soort vragen centraal.

Zo kan men in een onderzoek volgens de gefundeerde-theoriebenadering (Glaser & Strauss, 1967; Strauss & Corbin, 1990) een aantal fasen in het theorieontwikkelingsproces onderscheiden. In onze uitwerking van de gefundeerde-theoriebenadering worden de volgende fasen met doelstellingen onderscheiden:

- exploratiefase: de onderzoeker formuleert een probleemstelling en probeert aan de hand van een verkenning van het veld zo veel mogelijk relevante begrippen te formuleren;
- specificatiefase: de meest centrale begrippen worden uitgewerkt in termen van dimensies en variabelen waarmee het veld kan worden beschreven;
- reductiefase: de kern van de theorie en de relatie ervan met de centrale begrippen wordt geformuleerd;
- integratiefase: de theorie, als antwoord op de definitieve probleemstelling, wordt systematisch onderzocht.

In deze vorm van kwalitatief onderzoek staat de uitwerking van het analytisch kader centraal. De fasen kunnen als stappen in dat proces worden opgevat. Dit leidt ertoe dat de analyse in elke fase op de beantwoording van andersoortige vragen is gericht.

Deze in globale termen beschreven analysehandelingen komen min of meer terug in analyseprocedures die door verschillende auteurs in specifiekere stappen zijn beschreven (bijv. Spradley, 1979, 1980; Miles & Huberman, 1984, 1993; Maso, 1987; Wester, 1987, 1995). Recent zijn daar weer nadere uitwerkingen bij gekomen, mede onder impuls van de ontwikkeling van een groot aantal computerprogramma's ten behoeve van kwalitatief onderzoek. Men kan hierbij denken aan programma's als Ask Sam, The Ethnograph, Nudist, Atlas-ti en

Kwalitan (voor een consumentenbondachtige beschrijving zie Weizman & Miles, 1994; zie ook Muhr, 1992; Richards & Richards, 1990 en Seidel, 1988).

In elke fase van het analyseproces kan de computer worden ingeschakeld om de onderzoeker te voorzien van relevant materiaal, om analysebewerkingen te verwerken en de bevindingen vast te leggen. De analysestappen kunnen daardoor verder worden gespecificeerd (voor een uitwerking binnen de gefundeerde theoriebenadering zie Wester, 1995). Daarbij is de computer een zeer belangrijk hulpmiddel, die regelmatig wordt ingeschakeld bij allerlei analyseactiviteiten. Maar het is belangrijk te beseffen dat de computer een hulpmiddel blijft: de feitelijke analyse wordt door de onderzoeker zelf uitgevoerd.

2 Nadere uitwerking van de analyse volgens de gefundeerde theoriebenadering

Na de beschrijving van de terugkerende handelingen en de fasering van de analyse gaan wij hier specifiek in op de verschillende fasen van de gefundeerde theoriebenadering. In feite gaan we concreter in op hetgeen in de vorige paragraaf werd aangeduid als de analyse-inengere-zin. Waar het gaat om de ondersteuning van de computer refereren wij daarbij aan het programma Kwalitan (Peters, 1995), dat speciaal voor dit soort analyse is ontwikkeld. Men zou dezelfde ondersteuning in principe ook via andere programma's kunnen realiseren, hoewel er tussen de diverse programma's grote verschillen bestaan in de manier waarop het materiaal geprepareerd moet worden, de mogelijkheden die het programma biedt en de wijze waarop men met het programma communiceert.

In de volgende uitwerking volgen we de eerder omschreven fasen in het onderzoeksproces, waarbij we beginnen met bewerkingen op het materiaal nadat een eerste hoeveelheid materiaal is verzameld. We gaan er daarbij vanuit dat het gaat om een kwalitatief survey waarbij de gegevens verzameld worden via een flink aantal interviews. We zullen dan ook aangeven hoe de interviewstrategie er in de verschillende fasen uitziet.

2.1 Voorbereiding

Het geschikt maken van het materiaal voor een kwalitatieve analyse houdt allereerst in dat gezorgd moet worden voor de invoer van het materiaal in een computerbestand. Dat betekent meer dan alleen maar het uittypen of scannen van een tekst. Het gaat hier om wat eerder de transcriptie van het ruwe materiaal is genoemd. De onderzoeker moet bepalen hoe het materiaal in een bestand wordt vastgelegd. Dat lijkt voor interviews nogal simpel, maar in feite moet de onderzoeker allerlei beslissingen nemen, zoals het wel of niet opnemen van aarzelingen, stiltes en dergelijke in het transcript, het corrigeren van versprekingen, het opnemen van stemverhogingen, of juist het globaal uitschrijven wat zoal bedoeld is te zeggen. Als het gaat om video-opnames zijn de beslissingen die de onderzoeker moet nemen al veel complexer, omdat visuele informatie in een tekst moet worden omgezet. Maar zelfs als het ruwe materiaal teksten betreft, zoals de notulen van vergaderingen of een beleidsrapport, is er de vraag of men de gehele tekst wil opnemen in de transcriptie of dat men zich beperkt tot de relevante gedeelten daaruit.

Daarnaast delen wij het ruwe materiaal op in segmenten. We zien dat bij vrijwel alle computerprogramma's in de pre-analytische fase een of andere vorm van voorstructurering noodzakelijk is. In het geval van Kwalitan betekent dat, dat het materiaal vooraf wordt ingedeeld in segmenten; deze segmenten kan het programma vervolgens als eenheden onderscheiden. Om te voorkomen dat de onderzoeker van het onderscheiden van segmenten als het ware een aparte analysefase gaat maken, adviseren wij de tekst in segmenten uiteen te leggen aan de hand van 'logische posities' in de tekst, bijvoorbeeld als een interviewer een nieuwe vraag stelt, of in een document als het gaat om een nieuwe alinea of paragraaf. Dit onderscheiden van tekstsegmenten maakt het mogelijk enige structurering aan het materiaal te geven, die de start van de interpretatieve analyse bevordert, bijvoorbeeld omdat het mogelijk maakt dat segmenten over hetzelfde onderwerp met elkaar worden vergeleken. In programma's als Kwalitan is het altijd mogelijk de splitsing in segmenten aan te passen indien de analyse dat nodig maakt (bijvoorbeeld omdat je later meer over het onderwerp weet).

Indien de onderzoeker in dit stadium van de analyse behoefte heeft aan een nadere verkenning van het materiaal, kan hij daarvoor ook weer een beroep doen op de computer. Zo heeft Kwalitan een optie, om een overzicht te maken van de woorden die voorkomen in de tekst en de frequenties waarin ze voorkomen. Hoewel dit een analyseprocedure is die meer past bij vormen van inhoudsanalyse en als zodanig weinig verbindingen heeft met de gefundeerde theoriebenadering, kan een dergelijk overzicht toch nuttig zijn. Op deze manier kan de onderzoeker bekend raken met de soorten woorden en uitdrukkingen die respondenten gebruiken of die in de documenten voorkomen. Inspectie van deze lijst zet de onderzoeker op het spoor van relevante veldbetrokken begrippen.

2.2 De exploratiefase

De belangrijkste analyseactiviteit van de onderzoeker in deze fase wordt *open coderen* genoemd. Hierbij worden de teksten gelezen en geïnterpreteerd, worden er (nieuwe) codes geformuleerd en aan de segmenten toegevoegd.

Er kunnen hier verschillende technieken worden toegepast bij het open coderen. De belangrijkste is *gericht lezen*. Daartoe recapituleert de onderzoeker welke begrippen en categorieën bij de uitwerking van de probleem-

stelling, de selectie van respondenten en de uitwerking van de topiclijst een rol hebben gespeeld. Met dit voorlopig analytisch kader wordt het materiaal gelezen, vanuit de vraag: wat wordt er over het betreffende thema (begrip, categorie) gezegd. Deze begrippen en categorieën worden aan de interviewtekst verbonden waar dat relevant is en eventueel worden specifiekere codes geformuleerd die aangeven hoe erover het betreffende thema wordt gesproken en wat erover wordt gezegd. Deze manier van lezen sluit dus sterk aan bij het perspectief van de onderzoeker.

Een tweede techniek is *samenfattend lezen*. Daartoe leest de onderzoeker elk fragment en probeert met een of meer codes de strekking van de tekst en de belangrijkste termen die erin voorkomen aan te geven. Deze codes worden aan de fragmenten toegevoegd. Deze manier van lezen sluit meer aan bij het perspectief van de respondent.

Een derde techniek is *materiaalscanning*. Hier maken we expliciet gebruik van de mogelijkheden van de computer als tekstverwerker om een overzicht te krijgen van het woordgebruik van de respondent. De meeste woorden zijn uiteraard niet specifiek relevant voor het centrale onderzoeksthema, maar van bepaalde woorden is dat wel het geval. Deze woorden kunnen in de overzichtslijst worden geselecteerd.

Werkbestand	Segment	Codes	Output	Memo	Inhoud	Bestand	Overig
			Toon segmenten				
			----- tekst van het segment -----				
+---- huidig segment ----+			I:Leer jij voor goede cijfers?				
documentnr. : 2			Hi:Ja.				
documentnaam : lint119			I:Kan je zeggen waarom je dat doet?				
segmentnummer: 9			Ja, ik vind het niet zo leuk om naar school te gaan eigenlijk, dan denk ik van ja, toch heb je die ondergrond eigenlijk nodig voor je latere opleiding.				
+---- bewerk de codes ----+			Des te hogere cijfers je hebt des te eerder word je aangenomen.				
1. lesstof			Dus waarom ga je niet leren om die cijfers te halen en dan een beetje mooi rapport te krijgen?				
2. cijferleren			Ja en ook om niet te blijven zitten en nog langer op die school te blijven.				
3. critische acceptatie			Maar ik ga echt niet leren van leuk om te weten of zo, en ik hoef het eigenlijk niet te leren.				
4.			Ik denk dat de meeste mensen eigenlijk leren om de repetities en om een goed cijfer te				
5.							
6.							
7.							
8.							
9.							
10.							
11.							
12.							
+-----+							

teerd en als codes automatisch aan de fragmenten waarin ze voorkomen worden gekoppeld. Deze fragmenten kunnen vervolgens worden geselecteerd voor nadere lezing in gerichte en samenvattende zin.

De codes die tijdens het open coderen worden geformuleerd (zie de drie manieren van lezen hiervoor), worden gekoppeld aan de segmenten waarin het thema dat met de code wordt aangeduid, voorkomt. Figuur 1 laat zien hoe een en ander er bij Kwalitan uitziet. Op het scherm is de informatie met betrekking tot één segment weergegeven. Linksboven staat aangegeven welk segment momenteel wordt getoond. Rechts op het scherm staat de tekst geprojecteerd, terwijl aan de linkerkant ruimte is om de codes in te voeren. Aan elk segment kan een groot aantal (25) codes worden verbonden.

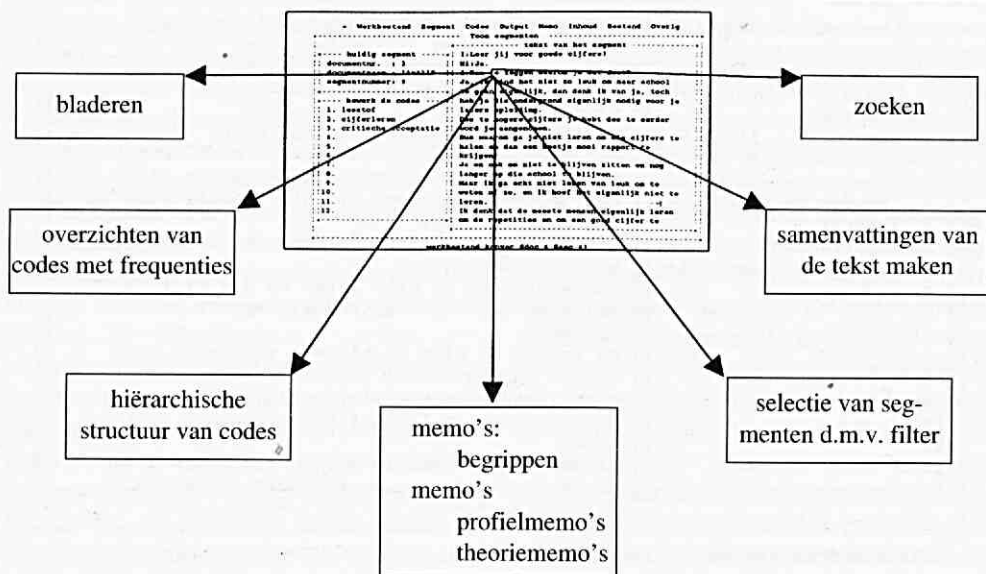
In Figuur 2 staan de verschillende mogelijkheden aangegeven waarmee vanuit het basis-scherm voor het coderen van het materiaal naar andere mogelijkheden van de computer kan worden overgegaan. Zo zijn er enkele opties om door de tekst te bladeren, om samenvattingen te maken van de tekst en de segmenten binnen het werkbestand achtereenvolgens door te lopen (bijvoorbeeld om codes te controleren die men eerder heeft gebruikt).

Eerder is al gewezen op de mogelijkheden

om activiteiten en beslissingen bij te houden in memo's³. In de exploratiefase zullen vooral begrippenmemo's worden geproduceerd, waarin de betekenis van de (belangrijkste) codes worden besproken. Daarnaast moeten in theoretische memo's de relaties tussen de codes met het analytisch kader worden bijgehouden. Kwalitan biedt de mogelijkheid om memo's in te voeren, aan te passen en tijdens de analyse op te roepen.

Nadat de codes aan de segmenten zijn toegevoegd, zal de onderzoeker graag een overzicht van de toegekende codes willen hebben. Via de computer vragen we een alfabetisch geordende lijst op van codes die aan de segmenten zijn toegevoegd (zie Figuur 3). Deze lijst geeft de onderzoeker de mogelijkheid te overzien welke codes hij gebruikt om het materiaal te beschrijven, met andere woorden de lijst laat de bouwstenen zien voor het analytisch kader. Daarnaast stelt deze lijst de onderzoeker in staat om typefouten in de codes te corrigeren en na te gaan of er mogelijk synoniemen voor hetzelfde thema zijn gebruikt.

De verkennende fase levert aldus een groot aantal codes op, die voor een deel met het voorlopige analysekader van de onderzoeker zijn verbonden. Daardoor krijgt de onderzoeker een beeld van de centrale begrippen of hoofdthe-



Figuur 2. Code-invoerscherm met segment uit een voorbeeld interview

Werkbestand	Segment	Codes	Output	Memo	Inhoud	Bestand	Overig
Overzicht				codes			
acceptatie			1				1
actuele relevantie			9				
beh. vaardigh. leren			1	bereik : alle documenten			
behoefte act. relev.			2	volgorde : alfabet			
cijferleren			9				
critische acceptatie			2	aantal verschillende codes:		25	
groeps vrijheid			4	totaal aantal codes		97	
groepsafhankelijkh.			2	aantal segmenten		41	
groepsefficiëntie			2				
groepssamenwerken			3	gemiddelde frequentie		3.88	
huiswerk			1	codes per segment		2.37	
individueel ideaal			1				
ineff.tijdsstructuur			1				
inter. vrijw. leren			1				
interesse			3				
irrelevantie			1				
klassikaal saai			2				
leerkr. ontsp. orde			4				
leerkracht gelijke			1				

werkbestand konver #doc 4 #seg 41

Figuur 3. Overzicht van codes in het werkbestand

ma's op het veld van onderzoek. De verkennende fase wordt afgesloten wanneer nieuw materiaal of een nieuwe lezing van het materiaal geen nieuwe hoofdthema's voor analyse meer opleveren.

Bij deze globale beschrijving van de werkwijze in de verkennende fase kan een aantal aantekeningen worden gemaakt, gelet op de problemen die onderzoekers zoal tegenkomen.

a. Het 'tijdverlies' door het transcriptie werk

Het transcriberen van interviews vraagt meer tijd dan onderzoekers vaak gepland hebben en levert een grote papierberg op, waardoor ook de analyse wordt bemoeilijkt. Dat roept de vraag op of daar niet op bezuinigd kan worden. Dat is zeker het geval voor de integratiefase van het onderzoek, maar niet in de beginfasen. Een eerste transcriptie van tape naar flop kan eventueel door hulp personeel worden uitgevoerd, maar de onderzoeker moet zelf de tape afluisteren om het basistranscript te corrigeren. Dat is geen verloren tijd, maar een eerste stap om thuis te raken in het materiaal. Daarbij komt dat de tekst eventueel kan worden aangevuld met toelichtingen uit de context van het interview om onbegrijpelijke of verwarrende teksten te verduidelijken. Na dit correctiewerk is de onderzoeker weer helemaal in het vraaggesprek en is dus in staat een profielmemo te

maken waarin in globale lijnen de plaats van de respondent bij belangrijke topics staat weergegeven.

b. Het tijdverslindende exploratieve lezen

De analyse volgens eerder genoemde technieken neemt zeer veel tijd in beslag. Wanneer in het onderzoek zo'n vijftig vraaggesprekken zijn gepland die alle op die manier gelezen moeten worden, zou het tijdbudget van het onderzoek makkelijk overschreden kunnen worden. Daarbij komt dat dit zoveel gecodeerd materiaal oplevert, dat ook de specificatiefase erg veel tijd vraagt.

De oplossing is dat men de veldverkenning op een beperkt aantal strategisch gekozen interviews uitvoert. Dit veronderstelt dat men in de veldverkenning fase weliswaar veel gesprekken heeft gevoerd (zeg 15), maar dat men aan de hand van de transcripties en de geproduceerde profielmemo's een zodanige selectie maakt (zeg 5) dat er een grote variatie is in wat naar voren is gebracht wat betreft de centrale onderzoekstopics. Immers in de veldverkenning fase gaat het erom zoveel mogelijk categorieën te ontwikkelen.

c. Het gevaar van fragmentatie

De line-by-line analyse en vooral het toekennen van codes aan fragmenten en de mogelijk-

heden die de computer biedt om fragmenten te vergelijken dwars door het materiaal heen, hebben het gevaar in zich dat de teksten worden gelezen los van hun context: het interview met een bepaalde respondent. De oplossing is dat men voortdurend de *profielmemo* van de respondent raadpleegt en dat deze profielmemo in iedere analyseronde wordt aangevuld wat betreft de score van de respondent op de centrale topics.

d. De sturing van de analyse

Het belangrijkste probleem is dat de onderzoeker leest en leest en steeds nieuw materiaal verzamelt en daardoor ook steeds nieuwe codes kan formuleren en dat oude codes weer steeds verder gespecificeerd kunnen worden, zodat het proces eindeloos voort kan gaan. Het verzadigingsprincipe (er komen geen nieuwe codes bij) werkt niet vanzelf. De onderzoeker moet het analyseproces sturen door naast interviewen en analyseren regelmatig tijd te reserveren om te reflecteren op de bevindingen. Hier kan de onderzoeker gebruik maken van een overzicht van de gebruikte codes en memo's schrijven over de relatie met het analytisch kader. De belangrijkste aanwijzing is Strauss' adagium (Strauss, 1987) 'think theoretical', waardoor structuur in het analytisch kader kan worden gebracht. De veldverkenkende fase kan worden afgesloten als de onderzoeker een voorlopig globaal antwoord op de probleemstelling kan formuleren en dus een idee heeft welke onderwerpen, begrippen en categorieën centraal staan in het onderzoek.

2.3 De specificatiefase

De specificatiefase is erop gericht de begrippen en codes die geformuleerd zijn rond de centrale onderwerpen in het onderzoek, precies uit te werken. De analyse is dus gericht op het ordenen van het analytisch kader om een beschrijving te kunnen geven van de aspecten die rond de centrale onderwerpen in het onderzoek van belang zijn.

De selectie van respondenten is hier gericht op aanvullingen van de onderzoeksgroep die in de veldverkenkende fase is ondervraagd. Het kan hier dus gaan om cases die juist verschillen van bepaalde onderzochte gevallen, maar ook om meer van dezelfde, die dus op andere cases lijken. De zo samengestelde onderzoeksgroep

maakt dan vergelijkingen mogelijk binnen een subgroep van gelijken én tussen subgroepen die verschillen.

Net als in de veldverkenkende fase zal de interviewer de topics bespreken door de respondent te laten uitleggen en toelichten, maar nu heeft de interviewer het voordeel dat hij/zij al wat thuis is in (de variaties in) het idioom waarmee over de onderwerpen wordt gepraat. Daarbij komt dat de analyse tot dan toe een globaal beeld heeft opgeleverd van centrale onderwerpen en bij sommige onderwerpen misschien ook al welke variaties daarbij een rol kunnen spelen. De analyse in de veldverkenkende fase *bevordert zo de mogelijkheden om door te vragen*. De interviewer is er nu op uit bij elk gespreksonderwerp een samenvatting te maken van het perspectief van de respondent, waarbij zoveel mogelijk al termen worden gebruik uit het analytisch kader, dat nu immers ook veldbetrokken begrippen bevat. Daarmee wordt tijdens de veldwerkfase al getoetst of de ontwikkelde begrippen passen op het veld van onderzoek.

Ook van deze vraaggesprekken wordt weer een volledig transcript gemaakt, dat in de computer kan worden ingelezen. De fragmenten worden nu onderscheiden aan de hand van de onderwerpen die in de veldverkenkende fase centraal zijn komen te staan in het onderzoek. Elk fragment krijgt aldus een of meer hoofdcodes, die op de betreffende onderwerpen slaan.

De lezing/codering van het materiaal geschiedt nu vanuit de overzichtslijst van codes, die in de veldverkenkende fase zijn gebruikt. Nieuwe codes worden alleen dan geformuleerd, wanneer aan de overzichtslijst (zie Figuur 3) geen passende code kan worden ontleend.

Is elk interview aldus gelezen en gecodeerd, dan gaat de onderzoeker over tot *vergelijkende analyses*. Hierbij worden alle interviews gebruikt die tot dan toe zijn gecodeerd.

Het belangrijkste doel van de specificatiefase is het uitwerken van centrale begrippen waarmee het verschijnsel dat men onderzoekt kan worden beschreven en de vraagstelling kan worden beantwoord. De analysehandelingen van de onderzoeker concentreren zich op voortdurende vergelijkingen binnen het materiaal. Op basis van de resultaten van de exploratiefase en van de reflecties op die resultaten

De belangrijkste ingang van het filter is de lijst van een of meer codes. De computer geeft alleen die segmenten weer op beeldscherm (of geprint) die aan de specificaties in het filter voldoen. Als je twee of meer codes in een filter opgeeft kunnen alle combinaties gemaakt wor-

Een handig hulpmiddel hierbij is een overzicht van de hiërarchische structuur van hoofd-

125
PEDAGOGISCHE
STUDIEN

Figuur 4. Het maken van een filter

Werkbestand	Segment	Codes	Output	Memo	Inhoud	Bestand	Overig
Boomstructuur codes							
boomstructuur	c:\11_persp				alle niveaus getoond	2 / 23	

> lesstof							
* relevantie							
* korte termijn							
+ actuele relevantie							
+ behoefte act. relev.							
+ interesse							
* middellange termijn							
+ cijferleren							
+ diplomaleren							
* lange termijn							
+ toekomstig relevant							
+ nuttig voor beroep							

> lesvorm							
* groepen							
* groepsefficiëntie							
* groepssamenwerken							
* leren>gezelligheid							

werkbestand konver #doc 4 #seg 41							

Figuur 5. De hiërarchische structuur van codes

en nevenscodes (zie Figuur 5). In zo'n overzicht zou eventueel tussen hoofd- en nevenscodes de dimensie als tussenlaag kunnen worden gedefinieerd. In het voorbeeld wordt deze tussenlaag gevormd door codes als 'korte termijn', 'middellange termijn' en 'lange termijn'. Deze codes kunnen vervolgens weer gebruikt worden om alle segmenten te selecteren die onder de codes van de 'tussenlaag' vallen.

Deze dimensies worden gebruikt om de *variabelen* te formuleren waarmee verschillen en overeenkomsten tussen respondenten met betrekking tot betreffende centrale onderwerp worden uitgedrukt. Daartoe moet de analyse vanuit deze dimensie per respondent worden uitgevoerd. Voor elke respondent wordt een profielmemo over de dimensie geschreven; vergelijking tussen de respondenten moet een variabele opleveren.

Zoals altijd moeten de reflecties die bij het uitwerken van begrippen en dimensies een rol hebben gespeeld in memo's worden vastgelegd.

Zijn voor alle centrale onderwerpen rond de probleemstelling dergelijke variabelen uitgewerkt, dan is de specificatiefase afgerond. Het analytisch kader is nu voldoende uitgewerkt om een beschrijvend overzicht te geven van de meest centrale onderwerpen in het onderzoek.

Bij deze globale beschrijving van de werk-

wijze in de specificatiefase kunnen de volgende aantekeningen worden gemaakt gelet op de problemen die de onderzoeker kan tegenkomen.

a. Onvolledige interviews

Na de specificatiefase is de onderzoeker goed in staat aan te geven welke onderwerpen in het interview aan de orde moeten komen en welke begrippen en categorieën daarbij een rol spelen. Dit heeft als voordeel dat het hierdoor mogelijk wordt om interviewers precies te instrueren zodat men (een deel van) de verdere vraaggesprekken kan uitbesteden. Een tweede consequentie van het uitgewerkte analytisch kader is dat de onderzoeker geconfronteerd wordt met lacunes in de vraaggesprekken die tot nu toe zijn gevoerd. Dat hangt samen met de start van het onderzoek: hoe opener het onderzoek, hoe meer lacunes in het analytisch kader, hoe groter de kans dat onderwerpen niet of niet volledig of niet diepgaand genoeg in de interviews aan bod zijn gekomen, gelet op de *uit-eindelijke analyse*. Met sommige respondenten zullen dus aanvullende gesprekken moeten worden gevoerd in de integratiefase.

b. Loskomen van de codes

De specificatiefase is erop gericht het perspectief van de onderzoeker verder uit te werken.

Dit betekent dat niet alle codes uit de verkennende fasen even belangrijk zijn. Sterker nog, het gaat niet om de codes op zich, ze zijn slechts *een hulpmiddel* om het begrippenkader van de onderzoeker uit te werken. Het analytisch kader bevat na de specificatiefase termen (variabelen, categorieën) die maar voor een klein deel met eerder gebruikte codes samen vallen. Daarbij komt dat in de verkennende fase codes zijn gebruikt, die later niet meer terugkeren omdat het niet om centrale onderwerpen in het onderzoek gaat. De onderzoeker moet dus loskomen van de codes en moet voorkomen dat de analyse uitloopt in een bureaucratisch afwerken van alle losse draden die met de codes zijn gegeven.

c. *Veldbetrokken literatuur*

Bij het uitwerken van het analytisch kader heeft de onderzoeker behoefte aan een bredere oriëntering dan het materiaal alleen. Bij de start van het onderzoek heeft de onderzoeker de probleemstelling gebaseerd op een globaal overzicht van het onderzoeksveld op grond van overzichtsartikelen. In de specificatiefase heeft de onderzoeker echter behoefte aan *specifieke literatuur met betrekking tot het onderzoeksveld*, waardoor de relevantie van de veldbetrokken begrippen ook buiten het eigen onderzoeksmateriaal kan worden onderzocht of eventuele onduidelijkheden in begripsformuleringen kunnen worden opgelost. Ook kunnen aan die literatuur begrippen ontleend worden om relevante onderscheidingen in het materiaal te benoemen.

d. *Aandacht voor reflectie*

In de specificatiefase heeft de onderzoeker de meeste voordelen van het werken met de computer. In de vergelijkende analyse werkt de onderzoeker met selecties uit het materiaal, die vanuit de analytische categorieën worden vergeleken (cf. Strauss & Corbin, 1990: "axial coding"). Bij die lezing van de fragmenten gaat het minder om het samenvatten en meer om het interpreteren van de tekst. Dit betekent dat het handwerk wordt afgewisseld door denkwerk waarbij de inzichten en de creativiteit van de onderzoeker een grote rol spelen. Reflectie op het analytisch kader en vastleggen van de zich ontwikkelende inzichten in memo's vormen een belangrijk deel van de analyseactiviteiten in de specificatiefase.

2.4 De reductiefase

In de reductiefase wordt de kern van de te ontwikkelen theorie bepaald om structuur te kunnen aanbrengen in de ontwikkelde variabelen en hun samenhangen. In de specificatiefase is het analytisch kader 'verticaal' ontwikkeld dat wil zeggen dat voor elk onderwerp in het onderzoek de structuur van begrippen, dimensies, variabelen en categorieën die daarop betrekking hebben, is uitgewerkt. In de reductiefase gaat het om een 'horizontale' uitwerking, dat wil zeggen de samenhang tussen de onderwerpen met elkaar en dus de manier waarop de begrippen en variabelen met elkaar zijn verbonden.

Wat betreft de kern van de theorie kan het gaan om een centraal begrip, bijvoorbeeld een typologie die de onderzoeker heeft ontwikkeld, die met vele andere variabelen samenhang vertoont. Ook kan het gaan om een proces, dat een ontwikkeling en samenhang beschrijft waarmee de verschillende variabelen kunnen worden geordend. Hier wordt weliswaar gesproken van een aparte fase om het kernbegrip te formuleren, maar vaak is de kern van de theorie in globale zin al tijdens de voorgaande fasen duidelijk geworden. De analyse in deze fase is dan alleen nog maar een toets of het betreffende kernbegrip inderdaad de samenhang in het materiaal brengt.

Ook in de reductiefase kunnen eventueel aanvullende interviews gedaan worden. De selectie van respondenten is hier gericht op het zoeken naar gevallen waarin de samenhang tussen de onderzochte verschijnselen naar voren moet komen. Het zal duidelijk zijn dat het uitgewerkte kader betekent dat de interviewer al veel meer systematisch te werk kan gaan en bij elk topic uit de vragenlijst door zal vragen vanuit de ontwikkelde begrippen en categorieën.

De transcriptie van de vraaggesprekken is hier ook selectief. De meeste topics kunnen aan de hand van het analytisch kader in samenvattende vorm worden uitgeschreven, terwijl eventueel alleen nog voor specifieke topics en fragmenten waarin samenhangen naar voren komen volledige transcripties worden gemaakt.

De analyse is hier voornamelijk reflectie aan de hand van specifieke selecties uit het materiaal, overzichten en tabellen. Er zijn verschillende strategieën als hulpmiddel om de

kern op het spoor te komen. Naast memo-onderzoek en literatuurstudie (los van het interviewmateriaal) kan men de profielmemo's van de respondenten een voor een bestuderen, om na te gaan wat per respondent de kern lijkt te zijn die het een en ander met elkaar in verband brengt. Daarnaast kunnen uit alle interviews die fragmenten geselecteerd worden, die samenhang tussen verschillende onderwerpen thematiseren. Ook kan gebruik gemaakt worden van allerlei reeds geproduceerde overzichten met scores van respondenten bij verschillende variabelen, zodat naar samenhang gezocht kan worden.

De reductiefase is vooral gericht op de uitwerking van de kern van de theorie. Veelal gaat het hier om een kernbegrip, waarmee de overige centrale begrippen zijn verbonden. In veel kwalitatief onderzoek heeft het kernbegrip de vorm van een typologie. In de reductiefase is de analyse gericht op de uitwerking van dit kernbegrip en op het onderzoeken van de relaties van dit begrip met andere begrippen. De belangrijkste analysehandeling is dan ook *selectief coderen*, waarbij de bovengenoemde opties worden toegepast. Het invoeren van codes, het maken van filters, het selecteren van segmenten en het maken van prints van selecties uit het materiaal zijn ook hier administratieve handelingen die de computer kan overnemen.

Het doel van deze analytische activiteiten is om naar de kern van de nieuwe theorie te zoeken door de relaties tussen de belangrijkste variabelen na te gaan. Een handig instrument in deze fase is een optie om grafisch te werken, zodat je een schema van de relaties tussen de variabelen kunt uitwerken en dat in het werkbestand kunt opnemen. Als zo'n grafische omgeving niet in het programma is opgenomen, kunnen dit soort grafische memo's in een speciaal tekenprogramma worden gemaakt, maar uiteraard voldoen papier en potlood ook goed!

Het zoeken naar belangrijke variabelen en hun relaties betekent dat men ook allerlei soorten van matrices en tabellen wil kunnen maken. Als het gaat om matrices hoeft men niet meteen te denken aan de bekende datamatrix van de kwantitatieve analyse; er wordt gewerkt met kwalitatieve matrices waarin de cellen met

kwalitatieve beschrijvingen zijn gevuld. Miles en Huberman (1993) geven een groot aantal voorbeelden van de soort matrices die er zijn en hoe je er mee om kunt gaan.

Een andere belangrijke analysehandeling in deze fase bestaat uit de systematische en grondige bestudering van de memo's die in de loop van het onderzoek met betrekking tot de analyse zijn geschreven. De daarin opgenomen ideeën moeten gerecapituleerd en samengevat worden om het kernbegrip en de relaties met andere begrippen te onderbouwen.

De reductiefase is afgerond wanneer het kernbegrip is geformuleerd en de samenhang met de centrale onderwerpen van het onderzoek duidelijk is.

Bij deze globale beschrijving van de analyse in de reductiefase kunnen de volgende aantekeningen worden gemaakt.

a. Verandering van analyseniveau

In de voorgaande fasen is de onderzoeker in de analyse voornamelijk bezig geweest op het niveau van het onderzoeksmateriaal, waarbij fragmenten werden gecodeerd en werden vergeleken. Alleen in (profiel)memo's heeft de onderzoeker op het niveau van de respondent gewerkt. In de reductiefase is de situatie juist andersom: nu vindt de analyse voornamelijk op het niveau van de respondenten plaats. De relevantie van het uitgewerkte analysekader kan getoetst worden door aan de hand daarvan de interviews samen te vatten en na te gaan of daarmee alle relevante informatie uit de interviews is opgenomen.

b. Aandacht voor theoretische literatuur

De onderzoeker heeft in de reductiefase behoefte aan een overzicht van theoretische invalshoeken met betrekking tot het onderzoeksveld om eventueel aan gangbare theorieën het kernbegrip te kunnen ontleenen. Daarnaast is de onderzoeker gericht op de formulering van het kernbegrip en ook hierbij kan de theoretische literatuur goede suggesties aandragen.

2.5 De integratiefase

De integratiefase is erop gericht de theorie precies uit te werken, de relaties tussen het kernbegrip en de andere concepten in de theorie op

het materiaal te toetsen en een antwoord te formuleren op de probleemstelling en dit met onderzoeksmateriaal te illustreren. De analyse heeft dus het karakter van een systematisch onderzoek, waarin een vaststaand analytisch kader wordt gehanteerd.

De selectie van respondenten is gericht op een zodanige aanvulling op de reeds onderzochte gevallen, dat aan de doelstellingen van het onderzoek kan worden voldaan. Soms betekent dit dat naast de specifieke, extreme of bijzondere gevallen die in de eerste fasen op theoretische gronden zijn geselecteerd, juist nu aanvullend de aandacht is gericht op doorsnee gevallen, zodat het hele spectrum aan voorkomende gevallen in de onderzoeksgroep aanwezig is. Ook kan het onderzoeksdoel vereisen, dat aanvullend op aselechte wijze een flink aantal gevallen wordt geselecteerd om een representatief beeld van de populatie te kunnen schetsen.

De interviewer gaat in deze fase zeer gericht te werk. Weliswaar kunnen de gesprekstopics vereisen dat de respondent veel ruimte krijgt om het eigen perspectief naar voren te brengen, maar de interviewer is er met name op gericht bij elk gespreksonderwerp de respondent volgens het analytisch kader te positioneren. Het doorvragen is er dus op gericht duidelijk te krijgen waar de respondent thuis hoort.

Nog steeds kan de onderzoeker ervoor kiezen om ook deze gesprekken op tape op te nemen, maar de transcriptie van het materiaal is beperkt tot samenvattingen van wat gezegd wordt in termen van het analytisch kader.

De uiteindelijke analyse gebeurt op de profielmemo's van alle respondenten, die samenvattingen bevatten in termen van het analytisch kader. Hierbij kan voor elk geval worden getoetst of het kernbegrip in staat is de kern van het perspectief van de respondent weer te geven en hoe over de hele onderzoeksgroep de samenhang van het kernbegrip met de andere centrale begrippen eruitziet.

Deze fase behelst de uitwerking van de theorie in de zin dat de observatie-eenheden (respondenten) worden beschreven en gekarakteriseerd met behulp van de variabelen en relaties die tezamen de nieuwe theorie vormen. Matrices, tabellen, samenvattingen en allerlei soorten overzichten zijn de hulpmiddelen om pro-

fielmemo's en tabellen te produceren op het niveau van de nieuwe theorie. Een andere mogelijkheid is om via de computer quasi-kwantitatieve matrices te maken, waarin bijvoorbeeld staat aangegeven of een bepaalde code in het materiaal van een respondent of een document voorkomt, of in hoeveel segmenten van een document of interview een bepaalde code voorkomt. Deze matrices hebben een zodanig algemeen format, dat zij ingelezen kunnen worden in speciale programma's voor kwantitatieve analyse, zoals bijvoorbeeld een homogeniteitsanalyse (Homals) of correspondentieanalyse. Deze analyses zijn bedoeld om een beter beeld te krijgen van relaties tussen variabelen die zinvol gekwantificeerd kunnen worden, en zo bijvoorbeeld de eenheden (respondenten/documenten) in categorieën onder te brengen.

De laatste stap bestaat uit het uitschrijven van de theorie in een onderzoeksrapport. De basis hiervoor kan gevonden worden in de vele methodische en theoretische memo's die de onderzoeker in de loop van het onderzoek heeft geschreven.

Bij de analyse in deze fase zijn de volgende aantekeningen van belang.

a. Herinterviews met 'oude' respondenten

Omdat op dit moment in het onderzoek het analytisch kader van de onderzoeker pas definitief is geworden, kan nu ook duidelijk worden welke de beperkingen en lacunes zijn in de vraaggesprekken die in de eerste fase zijn gevoerd. Dit betekent dat deze respondenten aanvullend ondervraagd moeten worden. Dit kan belangrijke inzichten opleveren. Enerzijds zal blijken dat de respondenten over de gesprekstopics niet in precies dezelfde woorden zullen spreken, wat duidelijk kan maken of het analytisch kader deze 'alledaagse variatie' overstijgt. Anderzijds kan duidelijk worden dat sommige respondenten met betrekking tot de onderzochte onderwerpen van idee zijn veranderd. De vraag is dan of vanuit de theorie kan worden begrepen waarom die verandering heeft plaats gevonden.

b. Analyseren is schrijven

De meeste stappen in de analyse zijn op een of andere manier in schrijffproducten vastgelegd.

Daar kan de onderzoeker bij het uitschrijven van het onderzoeksrapport gebruik van maken, zoals de vele methodische en theoretische memo's, maar ook de profielfragmenten van respondenten of gecodeerde fragmenten uit het onderzoeksmateriaal, die als illustratie kunnen dienen. Het schrijven van het eindrapport vraagt om nauwkeurige formuleringen, onderbouwingen van onderzoeksstappen en gefundeerde interpretaties van het onderzoeksmateriaal, die met elkaar consistent moeten zijn. In die zin is het schrijven van het eindrapport als laatste stap in de analyse te zien.

3 Slot

In het voorgaande hebben we de systematiek van de analyse in de gefundeerde theoriebenadering en de ondersteuning van de computer daarbij geschetst. De stappen in de analyse zijn daarbij alleen globaal aan de orde gekomen; de specifieke richtlijnen per stap zijn niet nader uitgewerkt. Ook zijn niet alle opties die het computerprogramma geeft, geïllustreerd (zoals het opstellen van matrices en het analyseren op woordniveau). Desondanks mag duidelijk zijn geworden dat de computer in alle fasen van de analyse belangrijke hulpfuncties kan vervullen.

Men zou een andere werkwijze dan de gefundeerde theoriebenadering kunnen toepassen, maar dat levert hoogstens andere paden op door het materiaal. Men kan een ander computerprogramma gebruiken, maar dat is zoiets als andere schoenen aantrekken. De basishandelingen met de computer zijn altijd min of meer hetzelfde. De computer kan de systematiek in de analyse bevorderen. In dit artikel hebben we laten zien, dat de analyse van kwalitatief materiaal gekenmerkt kan worden als een zoektocht langs analytisch kader en vraagstelling enerzijds en het empirisch materiaal anderzijds. De ondersteuning van de computer kan ertoe bijdragen, dat deze zoektocht meer systematisch en meer gestructureerd plaats vindt. Daarvoor is het wel noodzakelijk, dat het computerprogramma goed aansluit op de stappen die vanuit methodologisch perspectief noodzakelijk zijn.

Overigens moet men de mogelijkheden van de computer niet met bruikbaarheid voor de analyse verwarren. Hoewel de ontwikkelingen in de programmatuur steeds meer mogelijkhe-

den binnen bereik brengen zoals hypertext-achtige werkwijzen en de manipulatie van foto- en filmmateriaal, betekent dat tegelijkertijd dat de programma's zo algemeen worden (Ask Sam, Atlas-ti) dat de onderzoeker zelf zijn weg moet zoeken en eigen routines binnen het programma moet ontwikkelen. Wanneer ontwerpers zich bij het ontwikkelen van computerprogramma's voor de kwalitatieve analyse meer laten leiden door de technische mogelijkheden die de computer biedt dan door de context van een stapsgewijze analyseprocedure, kan het gebruik van de computer makkelijk tot valkuil worden, waarbij toeters en bellen de systematiek in de analyse naar de achtergrond doen verdwijnen.

Noten

- 1 Wij wijzen er op, dat het begrip analyse in het interpretatieve onderzoek veel breder wordt opgevat dan in andere vormen van onderzoek gangbaar is. Dit wordt in deze paragraaf toegelicht.
- 2 Coderen moet in deze context niet verward worden met het coderen dat plaatsvindt in kwantificerend onderzoek, waarbij vooraf bepaalde categorieën worden toegekend aan de antwoorden van de respondent. Binnen de context van interpreterend onderzoek wordt onder coderen verstaan de onderzoekshandeling waarbij de onderzoeker trefwoorden bedenkt die de inhoud van tekstfragmenten typeren; vervolgens worden deze trefwoorden of codes toegekend aan het betreffende fragment.
- 3 Binnen Kwalitan worden vier soorten memo's onderscheiden, elk met een eigen functie. In begrippenmemo's wordt de omschrijving van begrippen (codes, variabelen, dimensies) bijgehouden. Profielfragmenten bevatten een beschrijving (profiel) van een respondent in termen van de begrippen die in de analyse ontwikkeld zijn. Theoriememo's bevatten theoretische overwegingen, zoals hypothesen en het analysekader in wording. Methodememo's, tot slot, bevatten beslissingen en verantwoordingen met betrekking tot het analyseverloop.

Literatuur

- Fielding, N.G., & Lee, R.M. (1991). *Using computers in qualitative research*. London: Sage Publications.
- Glaser, B.G., & Strauss, A.L. (1967). *The Discovery of grounded theory*. Chicago: Aldine.
- Hijmans, E.J.S. (1994). *Je moet er het beste van maken*. Nijmegen: ITS.
- Hijmans, E.J.S. (1995). Kwalitatief survey. In: Hüttner, H., Renckstorf, K. & wester, F. *Onderzoekstypen in de communicatiewetenschap*. Houten: Bohn Stafleu Van Loghum.
- Kuyper, M. (1993). *Op de achtergrond. Een onderzoek naar de problemen van partners van patiënten met een chronische ziekte*. Amsterdam: Thesis publications.
- Maso, I. (1987). *Kwalitatief Onderzoek*. Meppel: Boom.
- Miles, M.B. and Huberman, A.M. (1984, 1993). *Qualitative Data-analysis. A sourcebook of new methods*. Sage: London.
- Muhr, Th. (1992). *Atlas / ti* (computer program). Berlin.
- Peters, V. (1995). *Kwalitan. A support program for qualitative analysis*. Version 4.0. (computer program). Nijmegen: University of Nijmegen.
- Peters, V. & Wester, F. (1995). An introduction to the principles and practice of our method of analysis. In: Pieterse, H.J.C. (ed.) *Desmond Tutu's Message. A qualitative Analysis*. Kampen (NL): Kok.
- Richards, T., & Richards, L. (1990). *Nudist 2.0: User's manual*. Melbourne: Rplee.
- Spradley, J.P. (1979). *The Ethnographic Interview*. New York: Holt Rinehart Winston.
- Spradley, J.P. (1980). *Participant Observation*. New York: Holt Rinehart Winston.
- Strauss, A.L. (1987). *Qualitative Analysis for Social Scientists*. New York: Cambridge Univ. press.
- Strauss, A.L., & J. Corbin (1990). *Basics of Qualitative Research; Grounded Theory, Procedures and Tactics*. Sage; London.
- Seidel, J.V. (1988). *The Ethnograph version 3.0*. (computer program). Littleton: Qualis Research Associates.
- Weizman, E.A., & Miles, M.B. (1994). *Computer programs for qualitative analysis*. Thousand Oaks, CA: Sage.
- Wester, F. (1995). *Strategieën voor Kwalitatief Onderzoek*. Bussum: Coutinho (1e ed.1987).

Auteurs

Fred Wester is universitair hoofddocent bij de vakgroep Methoden van de Faculteit Sociale Wetenschappen van de Katholieke Universiteit Nijmegen. Postbus 9104, HE Nijmegen.
E-mail: f.wester@ped.kun.nl.

Vincent Peters is universitair docent bij de vakgroep Methoden van de faculteit Beleidswetenschappen van de Katholieke Universiteit Nijmegen. Postbus 9104, HE Nijmegen.
E-mail: v.peters@ped.kun.nl

Abstract

Qualitative analysis of interview data. Phases, steps and computer support in intensive research

F. Wester & V. Peters. Pedagogische Studiën, 1999, 76, 117-131.

In this article a description is given of the design and the execution of the qualitative analysis of research material. At first, the characteristic cycle of observation – analysis – reflection, that is run through several times, is clarified. After that, the phases of the qualitative analysis are elaborated from the perspective of the grounded theory approach; the most important steps of analysis and the bottlenecks and pitfalls are pointed out. In addition, the way the computer can support this kind of analysis is illustrated by means of the program Kwalitan.

Computerintensieve data-analyse: Bootstrap- en randomisatiemethoden in de statistiek*

J.J. Hox

Samenvatting

Dit artikel beschrijft twee computerintensieve methoden voor inductieve statistiek: bootstrappen en randomisatietoetsen. Bootstrappen houdt in dat uit de voorliggende gegevens een groot aantal steekproeven wordt getrokken, waarna in iedere bootstrapsteekproef de gebruikte statistische grootheid wordt berekend. De variabiliteit van de uitkomsten van de bootstrapsteekproeven wordt beschouwd als een indicatie van de steekproefvariatie in de echte steekproef. De bootstrapmethode wordt gebruikt om de standaardfout te schatten en daarmee een toets uit te voeren of betrouwbaarheidsintervallen te construeren. De randomisatietoets dient om verschillen tussen twee of meer groepen te toetsen. De gegevens van de twee groepen worden gecombineerd en vervolgens vele malen op een aselechte wijze in twee groepen verdeeld, waarna het verschil van de betreffende statistische grootheid in die twee groepen wordt berekend. De verdeling van de verschillen tussen twee toevallig samengestelde groepen wordt gebruikt als nulverdeling voor het daadwerkelijk geobserveerde verschil tussen de twee groepen. Bootstrap- en randomisatieprocedures zijn vooral bruikbaar in situaties waarin een klassieke statistische toets niet beschikbaar is, of wanneer ernstig aan de geldigheid van de aannamen van de klassieke toets getwijfeld wordt. Daarmee kunnen computerintensieve statistische methoden gezien worden als een alternatief voor de klassieke non-parametrische toetsen.

Inleiding

Op het terrein van de statistische analyse heeft zich de laatste jaren, op een voor inhoudelijke onderzoekers vrijwel ongemerkte, bijna stiekeme manier, een nieuwe tak van statistiek ontwikkeld. Deze 'nieuwe statistiek' bestaat uit een verzameling van methodes, die ik hier samenvat met *computerintensieve statistiek*. Het kenmerkende van al deze methoden is dat er bij de bepaling van de significantie of het betrouwbaarheidsinterval niet gebruik wordt gemaakt van de klassieke statistische modellen voor de verdeling van de toevallige fouten, zoals een model met een normale verdeling. In plaats daarvan wordt het toevalsproces met de computer gesimuleerd. Deze simulatie wordt enige honderden tot duizenden malen herhaald, en de variaties in de uitkomsten van deze simulaties vormen een gesimuleerde empirische verdeling van de toevallige fouten.

Het statistisch jargon spreekt bij deze nieuwe toepassingen van resamplingmethoden, bootstrappen, randomisatie- en permutatietoetsen (zie o.a. Efron & Tibshirani, 1993; Mooney & Duval, 1993; Noreen, 1989). De wijze van werken laat zich het beste toelichten aan de hand van een voorbeeld. Bij wijze van illustratie volgt om te beginnen een voorbeeld van de zogenaamde *bootstrap*-methode.

1 Bootstrappen

Stel dat onderzoekers een vragenlijst willen ontwikkelen, bijvoorbeeld een korte attitude-schaal bestaande uit zeven vragen. Het is een goed gebruik om in een vooronderzoek voor zo'n schaal te bepalen, hoe betrouwbaar deze meet. Tabel A in de Bijlage bevat de (gefinancierde) gegevens van een eerste vooronderzoek met vijftien personen. Bij deze proefafname blijkt de coëfficiënt alfa gelijk te zijn aan 0.74.

* Met dank aan G.L.H. van den Wittenboer en AD. Nijdam voor hun commentaar bij een eerdere versie.

Doorgaans wordt een alfa van 0.70 of hoger beschouwd als een indicatie dat een schaal voldoende betrouwbaar meet. Een betrouwbaarheid van 0.74 lijkt dus voldoende. Het probleem is echter dat de proefgroep erg klein is, zodat we moeten aannemen dat er flinke steekproeffluctuaties in het spel kunnen zijn. Wat we nodig hebben is een betrouwbaarheidsinterval voor de coëfficiënt alfa. Hiervoor is een statistisch model ontwikkeld door Feldt (Feldt, 1965; Feldt, Woodruff & Salih, 1987; zie ook Van den Brink & Mellenbergh, 1998). De procedure van Feldt levert voor onze schaal een 95% betrouwbaarheidsinterval op van 0.47-0.90. Een nadeel van de procedure van Feldt is dat deze aanneemt dat de items normaal verdeeld zijn met gelijke variantie. Bovendien geldt het betreffende model pas bij grote aantallen vragen en bij grote steekproeven. Zeven vragen en vijftien respondenten lijken daarvoor beide aan de lage kant.

Een 95% betrouwbaarheidsinterval voor de coëfficiënt alfa kan echter ook worden vastgesteld met de *bootstrap*-methode. De bootstrapmethode is gebaseerd op de logica waarop statistische toetsen zijn gebaseerd. Die logica gaat ervan uit dat als we verscheidene malen een steekproef trekken uit een populatie, we vanzelfsprekend niet steeds precies dezelfde groep respondenten zullen krijgen, en daarom ook niet steeds dezelfde uitkomsten zullen vinden. De uitkomsten zullen van steekproef tot steekproef variëren en we gebruiken een mathematisch statistisch model om te schatten hoe groot de variatie is die we mogen verwachten. Op basis van aannamen voor de verdeling kunnen vervolgens betrouwbaarheidsintervallen bepaald worden en statistische toetsen worden uitgevoerd. De procedure van Feldt is gebaseerd op zo'n statistisch model.

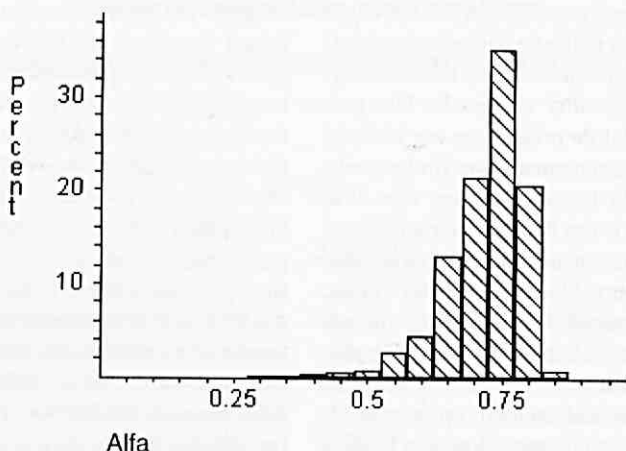
In plaats van een statistisch model te gebruiken, kunnen we natuurlijk ook duizendmaal met teruglegging een steekproef trekken, de gewenste uitkomsten berekenen, en bekijken hoe groot de variatie in uitkomsten daadwerkelijk is. In de praktijk is dat veel te bewerkelijk en bovendien te duur, maar in principe zou het kunnen. De bootstrapmethode bestaat uit een procedure die hier op lijkt. In plaats van duizendmaal een nieuwe steekproef te trekken uit de gehele populatie, trekt de computer duizendmaal een steekproef uit de reeds verzamel-

de gegevens. In het voorbeeld van de attitude-schaal wordt uit de voorliggende steekproef van vijftien respondenten duizendmaal met teruglegging een nieuwe steekproef van vijftien personen getrokken, en op die duizend nieuwe steekproeven wordt vervolgens duizendmaal de coëfficiënt alfa berekend. De variabiliteit van die duizend uitkomsten wordt gezien als een indicatie voor de grootte van de steekproeffluctuaties, zodat op basis hiervan een 95% betrouwbaarheidsinterval kan worden berekend en statistische toetsen kunnen worden uitgevoerd. Net als indertijd de Baron van Münchhausen trekken we ons aan onze laarzen het statistische moeras uit, vandaar de Engelse naam: *bootstrappen*.

Zoals hierboven is gesteld, trekken we de bootstrapsteekproeven met teruglegging. Telkens wanneer een respondent in de bootstrapsteekproef getrokken is, wordt deze weer teruggelegd bij de oorspronkelijke gegevens. De bootstrapsteekproeven verschillen dus van elkaar doordat sommige respondenten in een steekproef ontbreken, terwijl andere weer een aantal keren voorkomen. Omdat bij de bootstrap steeds opnieuw steekproeven uit de oorspronkelijke steekproef worden getrokken, heet deze strategie ook wel *resampling*.

Een resampling benadering van de betrouwbaarheid van coëfficiënt alfa houdt in dat we uit onze gegevens duizend bootstrapsteekproeven trekken. Vervolgens berekenen we duizendmaal de coëfficiënt alfa, en kijken bij de gevonden duizend alfa's welke de bovengrens en de ondergrens van de middelste 95% aanduiden. Het resultaat is dan een 95% bootstrapbetrouwbaarheidsinterval.

Figuur 1 geeft een histogram van de verdeling van coëfficiënt alfa in duizend bootstrapsteekproeven. Het 95% betrouwbaarheidsinterval volgens de bootstrapmethode loopt van 0.54 tot 0.82. Het percentage alfa's dat onder de kritische grens van 0.70 valt is 30%. De kans dat de betrouwbaarheidscoëfficiënt alfa in de populatie kleiner is dan 0.70, is aanzienlijk. Het betrouwbaarheidsinterval volgens de bootstrapmethode is kleiner dan het betrouwbaarheidsinterval volgens het model van Feldt (Feldt, 1965). In dit geval levert de klassieke statistische toets vermoedelijk een te grote standaardfout op. Aan het histogram van de



Figuur 1. Histogram van de waarden voor coëfficiënt alfa in 1000 bootstrapsteekproeven

bootstrap alfa's is goed te zien dat de steekproevenverdeling van alfa niet symmetrisch is. Vooral de scheve verdeling zorgt ervoor dat de kans op een fors lagere alfa bij vervolgonderzoek niet uitgesloten moet worden.

2 Randomisatie- en permutatietoetsen

Andere computerintensieve methoden zijn de zogenaamde randomisatie- of permutatietoetsen. Deze sluiten nauw aan bij de gedachten-gang in experimenteel onderzoek. Stel dat de gegevens van Tabel A verzameld zijn met behulp van een schriftelijke vragenlijst en we vragen ons af of dataverzameling door computerondersteund enquêteren meer betrouwbare gegevens zou geven (vgl. het artikel van Kef & Van Hattum in dit nummer). Volgens de beste principes van het experimentele onderzoek verzamelen we 50 respondenten, waarvan we er via loting 25 door de computer laten onder-vragen en de andere 25 een schriftelijke vragenlijst laten invullen. De computergestuurde interviewmethode levert een alfa op van 0.91, de vragenlijstmethode van 0.76. Mogen we nu concluderen dat de computergestuurde onder-vraging tot betrouwbaarder resultaten leidt?

Ook voor deze vraag is een statistische toets beschikbaar (Feldt, 1969; Hakstian & Whalen, 1976). Wanneer we deze toets toepassen, blijkt het verschil tussen de alfa's van beide groepen significant te zijn. De overschrijdingskans is

0.03, hetgeen wil zeggen dat de kans klein is dat dit soort verschillen op basis van toeval zijn ontstaan. We concluderen dat de wijze van vragen stellen invloed heeft op de betrouwbaarheid van de schaal.

De toets voor het verschil tussen twee alfa's berust op dezelfde aannamen als de hiervoor genoemde toets van Feldt (1965). Bij voorbaat weten we dan dat niet aan de aannamen voldaan is. We kunnen de toetsing echter ook op een andere wijze uitvoeren, namelijk via een *permutatie*- ofwel *randomisatietoets*. Een randomisatietoets is terug te voeren op de logica van het experimentele onderzoek. De basis-vorm van het experimentele onderzoek is, dat een proefgroep volgens een toevalsproces wordt verdeeld in een experimentele groep en een controlegroep. Vervolgens wordt het experiment uitgevoerd en wordt nagegaan of de experimentele groep verschilt van de controlegroep. Doordat de groepsindeling volgens toeval, oftewel *at random* is geschied, mogen we aannemen dat de beide groepen aan het begin van het experiment gelijk zijn. Wanneer we achteraf verschillen constateren, kunnen we die toeschrijven aan de experimentele manipulatie.

De rol van de statistische toets bij experi-menteel onderzoek is na te gaan of een eventueel gevonden verschil reëel is. Immers, wanneer we de proefpersonen at random over de experimentele en de controlegroep verdelen, zullen deze beide groepen doorgaans niet exact gelijk zijn. In ons voorbeeld is het erg onwaar-schijnlijk dat de coëfficiënt alfa in beide groe-

pen in alle decimalen gelijk zal zijn. Kleine, toevallig optredende verschillen zijn te verwachten. Ook hier zouden we de toetsing via een statistisch model in principe kunnen omzeilen door het experiment een groot aantal malen te herhalen, met steeds andere toewijzingen van personen aan groepen. Om allerlei redenen is dat niet erg praktisch. Wat we wel kunnen doen is de personen waarvan we gegevens hebben, in de computer op alle mogelijke manieren in twee groepen te verdelen, van beide groepen telkens de coëfficiënt alfa bepalen, en daarvan ten slotte het verschil berekenen. De verdeling van al die verschillen tussen twee alfa's is een indicatie van de mate van variabiliteit in het verschil die we op basis van toeval mogen verwachten. Wanneer het daadwerkelijk geobserveerde verschil, vergeleken met die referentieverdeling, uitzonderlijk groot is, concluderen we dat het geobserveerde verschil niet op toeval berust, maar reëel is. Analooq aan het 5% significantieniveau van de gebruikelijke statistische toets, kunnen wij afspreken dat we een gevonden verschil significant noemen, als het zich bij de 5% grootste verschillen bevindt.

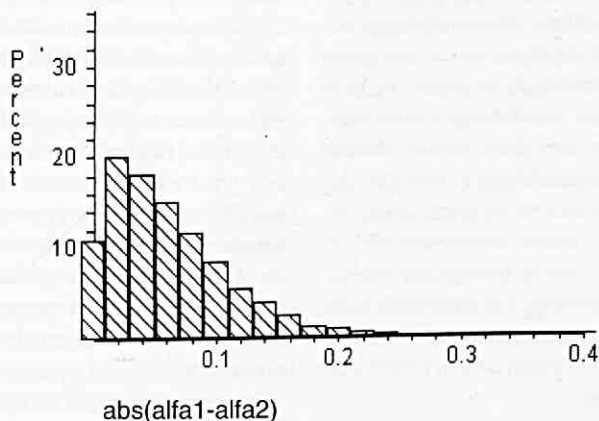
De zojuist beschreven toets heet een randomisatietoets, omdat deze het at random toewijzen van personen aan groepen nabootst. De randomisatie- of permutatietoets is al oud; hij is al in 1935 door Fisher beschreven. Nieuw is, dat de randomisatietoets tegenwoordig eenvoudig in de computer kan worden uitgevoerd en niet alleen meer voor eenvoudige grootheden zoals gemiddelden. Zolang het aantal personen

in de steekproef klein is, kunnen alle mogelijke manieren om twee groepen samen te stellen in de computer worden uitgerekend. Het aantal manieren om 50 personen over twee groepen te verdelen is echter onhanteerbaar groot. Gebruikelijk is om dan bij de constructie van de referentieverdeling een steekproef te trekken en bijvoorbeeld duizend random toewijzingen uit te voeren en door te rekenen. Dit heet een approximatieve randomisatietoets.

Figuur 2 bevat het histogram van de verdeling van de absolute verschillen tussen twee alfa's bij duizend at random samengestelde groepen uit de steekproefgegevens in Tabel B van de bijlage. Het verschil tussen de betrouwbaarheidscoëfficiënten in de twee groepen bedraagt 0.15. Het percentage gevonden verschillen van 0.15 en groter is in de randomisatietoets 6%. Dat is iets groter dan de gebruikelijke kritische grens van 5%, en dus mogen we niet concluderen dat de beide groepen significant van elkaar verschillen. Omdat de overschrijdingskans in de randomisatietoets dicht bij de kritische significantiegrens van 0.05 ligt, is de randomisatietoets herhaald, nu met tienduizend aselekt samengestelde groepen. De overschrijdingskans in de randomisatietoets is nu 6.2%, en de conclusie blijft dat de nulhypothese hier niet kan worden verworpen.

3 Software

Computerintensieve methoden behoren nog niet tot het statistische standaardpakket. Recen-



Figuur 2. Histogram van de verschillen tussen de coëfficiënten alfa van twee groepen in duizend bootstrapsteekproeven.

Tabel 1
Kruistabel van Tijdschrift en Soort probleemstelling

	PROBLEEMSTELLING (belangrijkste)					Totaal
	NAUWELIJKS AANWEZIG	EXPLO- RATIE	CONCRETE VRAAG	HYPO- THESE	INSTR. ONTWIKKELING	
TIJDSCHRIFTENCODI						
TS v. Opv., Vorming en Onderwijs	1	8	6	3	2	20
Pedagogisch Ts	4	9	11	9	4	37
Pedagogische Studiën		6	23	10	2	41
Kind en Adolescent	7	15	12	10	3	47
Ts v. Orthopedagogiek	11	15	13	11	11	61
Ts v. Onderwijsresearch	4	11	35	16	7	73
Ts v. Onderwijswetenschappen	3	6	21	2	3	35
Totaal	30	70	121	61	32	314

te versies van het statistisch pakket SPSS (vanaf 7.0) bevatten al enkele van de hierboven besproken methoden (Mehta & Patel, 1996). De module 'exact tests' voegt aan de kruistabel en de meeste nonparametrische toetsen de mogelijkheid toe om een 'exacte test' uit te voeren. In feite gaat het hierbij om de hierboven besproken randomisatietoets. SPSS berekent bij kleine steekproeven voor alle mogelijke steekproefindelingen de betreffende toetsingsgrootheid, en bepaalt vervolgens waar de geobserveerde toetsingsgrootheid zich bevindt in deze referentieverdeling. De proportiewaarden van de toetsingsgrootheid in de gesimuleerde randomisatieverdeling die groter is dan de gevonden waarde van de toetsingsgrootheid, levert de overschrijdingskans op voor de toets. Wanneer alle mogelijke steekproefindelingen kunnen worden berekend, is dit een exacte toets. In veel gevallen zal deze equivalent zijn aan een nonparametrische toets waarvoor de exacte overschrijdingskans bekend is. Bij grote steekproeven is het aantal mogelijke groepsindelingen te groot. In zo'n geval rekent SPSS de waarden uit van de toetsingsgrootheid voor een groot aantal (default 10000) random groepsindelingen. Het gaat dan om een approximatieve randomisatietoets, welke door SPSS wordt aangeduid als de 'Monte Carlo'-test². Dit is niet langer equivalent aan de overeenkomstige nonparametrische toets, omdat deze bij grote steekproeven de overschrijdingskans bepaalt via een (doorgaans normale) benadering.

deling van het soort vraagstelling aangetroffen in artikelen in een aantal Nederlandse wetenschappelijke tijdschriften. Deze tabel is afkomstig uit een vergelijking van Nederlandse wetenschappelijke tijdschriften door Hox en Van Peet (1998). In het oorspronkelijke artikel wordt deze tabel geïnterpreteerd zonder dat er een statistische toets is uitgevoerd om na te gaan of de ogenschijnlijke verschillen tussen de tijdschriften significant zijn. De reden daarvoor is duidelijk wanneer we naar de aantallen in de cellen kijken. Wanneer we een gewone chi-kwadraat-toets opvragen voor de gegevens in Tabel 1, dan geeft SPSS de waarschuwing dat veel cellen een verwachte frequentie hebben kleiner dan vijf en dat toepassing van de chi-kwadraat-kansverdeling voor de toets niet verantwoord is. Het is echter mogelijk een exacte toets op te vragen. Dan berekent SPSS alle mogelijke manieren om de 314 waarnemingen, bij vastliggende randtotalen, over de 35 cellen van de kruistabel te verdelen. Dit komt overeen met de nulhypothese dat de beide variabelen die de kruistabel definiëren, niet met elkaar samenhangen. Voor elk van de mogelijke random indelingen berekent SPSS een chi-kwadraat-waarde. De gevonden chi-kwadraat-waarde wordt vergeleken met de chi-kwadraat-waarden uit de exacte kansverdeling, en de proportie chi-kwadraat-waarden gelijk aan of groter dan de gevonden chi-kwadraat-waarde levert wederom de exacte overschrijdingskans.

Tabel 2 is een weergave van de SPSS-uitvoer voor de exacte toets op de gegevens van

Tabel 2

Monte Carlo significantietoets gebaseerd op tienduizend gesimuleerde kruistabellen

	Value	df	Asymp. Sig. (2-sided)	Sig.	Monte Carlo Sig. (2-sided)		Monte Carlo Sig. (1-sided)		
					9% Confidence Interval		9% Confidence Interval		
					Lower Bound	Upper Bound	Sig.	Lower Bound	Upper Bound
Pearson Chi-Square	47.474 ^a	24	.003	.003 ^b	.001	.004			
Likelihood Ratio	51.267	24	.001	.002 ^b	.001	.003			
Fisher's Exact Test	46.897			.002 ^b	.001	.002			
Linear-by-Linear Association	.140 ^c	1	.708	.713 ^b	.701	.724	.356 ^b	.343	.368
N of Valid Cases	314								

a 12 cells (34.3%) have expected count less than 5. The minimum expected count is 1.91.

b Based on 10 000 sampled tables with starting seed 2 000 000.

c The standardized statistic is .374.

Tabel 1. Bij deze steekproefgrootte blijkt het niet mogelijk alle mogelijke verdelingen van 314 waarnemingen over 35 cellen te bepalen. SPSS voert daarom een approximatieve randomisatietoets uit, gebaseerd op 10000 random verdelingen. Dit levert bij deze gegevens voor de standaard Pearson chi-kwadraat vrijwel dezelfde overschrijdingskans op als de asymptotische: p is 0.003. Dat is op zich niet verwonderlijk; de regel die SPSS toepast bij het vermelden van de waarschuwing (niet meer dan 20% van de cellen mag een verwachte frequentie hebben lager dan 5) is nogal conservatief. Bij de Monte Carlo-resultaten krijgen we bovendien nog een 99% betrouwbaarheidsinterval van de berekende p -waarde, die immers gebaseerd is op tienduizend trekkingen van random-verdelingen. Als we het 99% betrouwbaarheidsinterval van de randomisatie p -waarde te breed vinden, kunnen we de toets eventueel overdoen en dan meer dan tienduizend trekkingen specificeren. In ons geval zijn de resultaten nu reeds duidelijk; de nulhypothese dat de verdeling van soorten probleemstellingen over de tijdschriften gelijk is, moet worden verworpen.

Bootstrapmethoden zijn (nog) niet beschikbaar in SPSS. Ze zijn al wel beschikbaar in een aantal programma's voor complexe modellen. Zo bevatten de programma's Lisrel, Eqs en Amos (alle bedoeld voor het schatten van structurele modellen) mogelijkheden om een bootstrap uit te voeren. Bij Lisrel en Eqs moet de gebruiker zelf nog het gemiddelde en de standaardafwijking van de bootstrapverdelingen berekenen,

bij Amos doet het programma alles automatisch. Ook andere complexe programma's, zoals het multiniveau-analyseprogramma MlwiN, bevatten tegenwoordig bootstrapprocedures.

Voor het toepassen van bootstrapprocedures op variantieanalyses, regressieanalyses, en andere gebruikelijke analysetechnieken moet een speciaal programma gebruikt worden. Het expertisecentrum ProGamma in Groningen heeft daarvoor het programma BOJA (Boomsma, 1991), dat bootstrap-schattingen kan berekenen voor een groot aantal toetsen.

Een speciale vermelding verdient het programma Resampling Stats (Simon, 1997). Dit is in feite een programmeertaal, waarin de gebruikers zelf hun statistische analyse moeten programmeren. Dat blijkt in de praktijk mee te vallen, omdat het programma veel commando's bevat die speciaal voor resamplingmethoden bedoeld zijn. Zo zet het commando SHUFFLE X alle getallen in de kolom X in random volgorde. Voor het bepalen van de randomisatieverdeling van de correlatie tussen X en Y, onder de nulhypothese dat er geen verband is tussen beide variabelen, is het dus voldoende om X duizendmaal te SHUFFLEN, en van de duizend berekende correlaties (commando: CORREL X Y in r) de verdeling te bepalen. Voor complexe variantieanalyses is Resampling Stats niet erg geschikt, maar voor randomisatietoetsen en bootstrapverdelingen van eenvoudige statistische grootheden zoals (verschillen tussen) gemiddelden, medianen, en correlaties is het vrij eenvoudig hanteerbaar.³

4 Voor- en nadelen van bootstrap- en randomisatiemethoden

Bootstrap- en randomisatiemethoden bestaan al enige tijd. Ze worden dikwijls toegepast wanneer langs de weg van de mathematische statistiek geen oplossing kan worden gevonden voor een specifiek toetsingsprobleem. Een voorbeeld daarvan is een betrouwbaarheidsinterval voor het verschil tussen twee medianen. Ze worden ook veel toegepast, als aan allerlei aannamen van de statistische toets niet is voldaan. De gedachte daarachter is in principe eenvoudig. Wanneer bijvoorbeeld de geobserveerde variabelen een zeer vreemde verdeling hebben, zodat aan de gebruikelijke aanname van een normale verdeling niet voldaan is, dan zal deze vreemde verdeling in elke bootstrap-steekproef opnieuw opduiken. Het probleem wordt daarmee vanzelf in de berekeningen verdisconteerd.⁴ Omdat de bootstrap- en randomisatieprocedures alle informatie gebruiken die een steekproef bevat, hebben ze doorgaans ook een onderscheidingsvermogen (power) dat weinig onderdoet voor de gebruikelijke parametrische toetsen. Zie voor een bespreking van de power van randomisatieprocedures Onghe-
na (1994.) Bootstrap- en randomisatieprocedures volgen doorgaans de logica van de parametrische toets, zonder een beroep te doen op de aanname van een normale verdeling. Nonparametrische toetsen volgen doorgaans een sterk verschillende logica, en hebben bij normale verdelingen minder onderscheidingsvermogen. Computerintensieve methoden zijn daarom een aantrekkelijk alternatief voor de gebruikelijke nonparametrische toetsen, die in veel gevallen minder onderscheidingsvermogen hebben dan hun parametrische tegenhangers, en bovendien soms ook sterke verdelingsaannamen doen. Zo veronderstelt bijvoorbeeld de bekende Mann-Whitney-toets voor het verschil in gemiddelde tussen twee groepen weliswaar niet een normale verdeling, maar wel dat beide groepen een gelijkvormige verdeling hebben.

Bootstrap- en randomisatietoetsen lossen uiteraard niet alle statistische problemen op. Allereerst wordt alle informatie uit de voorliggende steekproef betrokken. De validiteit van de statistische conclusies hangt daarom volle-

dig af van de representativiteit van die steekproef. Dat is bij de klassieke statistiek overigens ook het geval. Men name bij kleine steekproeven kunnen bootstrapprocedures in de problemen komen, omdat de variabiliteit in de populatie onderschat wordt. Daarnaast zijn er enkele statistische grootheden, zoals de grootste waarde in een steekproef, die zich slecht lenen voor bootstrappen of randomisatietoetsen (Mooney & Duval, 1993, p. 60).

Het moge duidelijk zijn waarom hier gesproken wordt van computerintensieve methoden. Het berekenen van één enkele coëfficiënt alfa is betrekkelijk weinig werk, zeker op een steekproef van slechts vijftig personen. Het berekenen van duizend alfa's valt voor een computer ook nog wel mee. Wanneer het daarentegen gaat om bootstrappen of randomisatietoetsen bij complexe multivariate modellen voor grote steekproeven, dan staat zelfs een moderne PC enige tijd te rekenen. Naarmate de computers sneller en goedkoper worden, zullen dit soort methoden echter aan populariteit winnen. De hier gegeven voorbeelden werden bijvoorbeeld berekend op een kleine Pentium laptop waarop alle berekeningen hoogstens enkele minuten duurden.

Noten

- 1 Procedures gebaseerd op het simuleren van toevalsprocessen, waaronder de randomisatieprocedures vallen, worden dikwijls aangeduid met Monte Carlo procedures, naar het bekende casino.
- 2 Een demonstratieversie van Resampling Stats is te verkrijgen op het Internetadres <http://www.statistics.com>. Informatie over BOJA, inclusief een demonstratieversie, is verkrijgbaar bij Programma op het Internetadres <http://www.gamma.rug.nl>.
- 3 De hier beschreven bootstrap-methode wordt wel aangemerkt als de naïeve of nonparametrische bootstrap. Wanneer deze methode gebruikt wordt voor het simuleren van de passing van complexe modellen, zijn de resultaten doorgaans onbetrouwbaar (Bollen & Stine, 1992). Dan biedt de parametrische bootstrap uitkomst: de achtereenvolgende steekproeven worden daarbij niet uit de data getrokken, maar met behulp van de geschatte modelparameters uit een theoretische kansverdeling gegenereerd

(vgl. Van der Heijden, 't Hart & Dessens, 1997). De parametrische bootstrap maakt daarmee wel weer gebruik van de aanname dat de gegevens een theoretische kansverdeling volgen.

Peet, A.A.J. van, Wittenboer, G.L.H. van den & Hox, J.J. (1995). *Beschrijvende technieken*. Groningen: Wolters-Noordhoff.
Simon, J.L. (1997). *Resampling: the new statistics*. Arlington, VA: Resampling Stats.

Literatuur

Auteur

J.J. Hox is werkzaam aan de Universiteit van Utrecht als hoogleraar methodenleer. Dit artikel is voor een deel gebaseerd op zijn oratie, die op de homepage te verkrijgen is. Cap. groep Methodenleer en Statistiek, FSW, UU. Postbus 50140, 3508 TC Utrecht. E-mail j.hox@fss.uu.nl.
Homepage <http://www.fsw.ruu.nl/ms/jh>

Abstract

Computer intensive data analysis: bootstrap and randomization methods in statistics

J.J. Hox. Pedagogische Studiën, 1999, 76, 132-141.

This article describes two computer intensive methods for deciding on statistical generalizability: the bootstrap and randomization tests. In the bootstrap method, repeated samples are drawn from the available data, and subsequently the statistic of interest is calculated in each bootstrap sample. The variability of the results in the bootstrap samples is taken as indicative for the sampling variability of that specific statistic in the population at hand. Thus, the bootstrap sampling distribution is used to derive confidence intervals and significance tests. The randomization test is used to test group differences with respect to a specific statistic. The available data, ignoring the grouping structure, are randomly divided into groups following the group structure in the sample. For each random assignment, the difference of the specified statistic is calculated across the groups. The set of outcomes of the random assignments is then used as a null distribution for the observed difference. Bootstrap and randomization procedures are especially useful if there is no classical statistical procedure for the statistic of interest, or if the distribution of the data follows an unusual distribution. In the latter case, computer intensive methods are an alternative to classical nonparametric procedures.

Bollen, K.A., & Stine, R.A. (1992). Bootstrapping goodness-of-fit measures in the context of structural equation models. *Sociological Methods & Research*, 21, 205-229.
Boomsma, A. (1991). *BOJA. A program for bootstrap and jackknife analysis*. Groningen: ProGamma.
Brink, W.P. van den, & Mellenbergh, G.J. (1998). *Testleer en testconstructie*. Meppel: Boom.
Efron, B., & Tibshirani, R.J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
Feldt, L.S. (1965). The approximate sampling distribution of Kuder-Richardson coefficient twenty. *Psychometrika*, 30, 357-370.
Feldt, L.S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. *Psychometrika*, 34, 363-373.
Feldt, L.S., Woodruff, D.J., & Salih, F.A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, 11, 93-103.
Hakstian, A.R., & Whalen, T.E. (1976). A k-sample significance test for independent alpha coefficients. *Psychometrika*, 41, 219-231.
Heijden, P.G.M. van der, 't Hart, H., & Dessens, J. (1997). A parametric bootstrap procedure to perform statistical tests in a LCA of anti-social behavior. In: J. Rost & R. Langeheine eds. *Applications of latent trait and latent class models in the social sciences* (pp. 196-208). Münster/New York: Waxmann.
Hox, J.J. & Peet, A.A.J. van (1998). Van oude methoden en dingen die voorbij gaan. *Pedagogische Studiën*, jubileumnummer 1998.
Mehta, C.R. & Patel, N.R. (1996). *SPSS Exact Tests 7.0 for Windows*. Chicago: SPSS Inc.
Mooney, C.Z. & Duval, R.D. (1993). *Bootstrapping. A nonparametric approach to statistical inference*. Newbury Park, CA: Sage.
Noreen, E.W. (1989). *Computer-intensive methods for testing hypotheses: An introduction*. New York: Wiley.
Onghena, P. (1994). *The power of randomization tests for single-case designs*. Academisch proefschrift, Katholieke Universiteit Leuven.

Bijlage

De (gefingeerde) gegevens

Tabel A
Scores van 15 personen op 7 vragen

Personen	1	2	3	4	5	6	7	Totaal
1	5	4	5	6	7	6	2	35
2	1	2	3	2	4	1	7	20
3	4	3	4	5	4	3	7	30
4	1	1	1	1	1	1	6	12
5	2	2	2	2	2	2	5	17
6	7	7	7	7	7	7	1	43
7	1	2	3	2	2	1	6	17
8	6	5	6	4	6	7	2	36
9	2	1	4	1	2	6	7	23
10	3	2	3	2	3	3	7	23
11	5	6	4	6	7	6	2	36
12	4	4	5	4	6	5	1	29
13	2	2	2	2	2	2	6	18
14	5	4	3	5	6	5	1	29
15	7	6	5	7	6	5	2	38

N.B. Deze tabel is afkomstig uit Van Peet, Van den Wittenboer & Hox, 1995, p. 210).

Tabel B

Scores van 50 personen op 7 vragen, CASI vs. PAPI.

Personen	1	2	3	4	5	6	7	Totaal	Groep
1	5	6	5	5	5	5	6	37	1
2	6	6	6	6	6	6	6	42	1
3	5	3	3	4	4	5	6	30	1
4	6	5	4	6	4	5	6	36	1
5	6	5	6	5	5	4	6	37	1
6	5	6	5	5	5	6	6	38	1
7	5	5	5	5	5	6	6	37	1
8	6	6	5	6	4	5	6	38	1
9	5	5	7	5	4	5	6	37	1
10	6	5	6	5	5	4	6	37	1
11	4	5	6	4	5	4	7	35	1
12	2	3	3	2	4	3	5	22	1
13	3	4	3	4	4	4	6	28	1
14	6	5	6	5	5	5	7	39	1
15	5	4	6	6	6	4	6	37	1
16	3	3	1	2	2	2	6	19	1
17	6	5	4	5	4	5	6	35	1
18	6	5	4	5	4	6	7	37	1
19	3	4	3	3	4	3	6	26	1
20	5	3	5	3	4	5	6	31	1
21	6	6	5	6	5	5	7	40	1
22	2	3	3	3	4	2	6	23	1
23	1	3	2	3	1	3	6	19	1
24	4	6	5	4	4	4	6	33	1
25	2	2	3	3	3	3	6	22	1
26	3	4	2	3	2	5	7	26	2
27	4	4	8	2	3	5	7	33	2
28	6	2	4	3	5	4	4	28	2
29	3	6	3	2	1	3	7	25	2
30	3	4	5	1	4	5	7	29	2
31	1	3	3	1	2	3	7	20	2
32	3	5	4	5	5	5	6	33	2
33	3	4	2	4	3	3	7	26	2
34	6	7	3	4	5	5	7	37	2
35	4	3	4	5	4	5	7	32	2
36	7	4	4	4	3	5	7	34	2
37	5	5	6	3	7	5	6	37	2
38	5	4	2	4	3	5	7	30	2
39	5	4	6	4	5	4	7	35	2
40	2	2	0	2	1	3	5	15	2
41	3	4	2	0	2	2	6	19	2
42	5	7	6	6	3	7	2	36	2
43	3	2	2	2	2	2	6	19	2
44	6	7	4	6	5	7	7	42	2
45	3	3	4	3	7	3	7	30	2
46	4	4	3	3	5	9	7	35	2
47	3	4	4	3	5	3	6	28	2
48	0	1	2	3	1	2	6	15	2
49	0	4	3	3	6	3	7	26	2
50	2	4	1	3	2	2	6	20	2

Realiteit van wiskunde- onderwijs VWO ter discussie

Aad Goddijn

Freudenthal instituut, Universiteit van Utrecht

Inleiding

In het voorjaar van 1998 stak er in de pers een kleine storm op over de kwaliteit van het Nederlandse wiskundeonderwijs. Aanleiding was de inaugurale rede van 17 april die Frans Keune hield bij zijn benoeming tot hoogleraar algebra in Nijmegen (Keune, 1998a). Keune koos ervoor niet in te gaan op zijn wetenschappelijk specialisme maar sprak zijn zorg uit over ontwikkelingen in het Nederlandse wiskundeonderwijs onder de titel *Naar de knoppen*. Om te beginnen zette hij uiteen wat wiskunde volgens hem is: wiskunde gaat over abstracte objecten en zuivere deductie is de enige werkmethode. Toepassingen van de wiskunde zijn er uiteraard, maar wiskunde is geen natuurwetenschap. De bedoeling van wiskundeonderwijs moet zijn logisch te leren denken; de wiskunde is daar de juiste context voor, omdat die context volmaakt helder is.

Met voorbeelden uit wiskundemethoden laat Keune zien dat de 'nieuwste trend' op gebied van wiskunde (de zogenaamde realistische wiskunde) de plank mislaat. Keune is er duidelijk over dat het hem vooral gaat om de groep leerlingen die via de top van het VWO in een van de exacte vakken terecht zullen komen. Die groep komt in de onderbouw van het VWO niet aan zijn trekken en het gevolg is dat de belangstelling voor de bèta-vakken ernstig daalt.

De inaugurale rede – inclusief de publiciteit er omheen – had het karakter van een manifest en eindigde dan ook met een plan tot actie: het ontwikkelen van een via Internet beschikbare methode waarin het gaat om het juiste samenspel van abstractie, deductie en toepassing.

In de wetenschapsbijlage van de *NRC*

(Van Delft, 1998) brengt hij de kwestie ook direct in verband met het Freudenthal instituut:

Het Freudenthal instituut mag internationaal aan de weg timmeren, waar het abstractie en realiteit door elkaar klutst tot een mistig geheel is de betere leerling de klos.

Ook in de *Nieuwe Wiskrant* heeft Keune zijn standpunt kunnen toelichten (Keune, 1998b). In een korte reactie wijst Martin Kindt (Kindt, 1998) van het Freudenthal instituut erop dat de voorstelling van zaken die Keune van realistisch wiskundeonderwijs geeft niet strookt met wat in de literatuur hierover te vinden is, en gaat op de voornaamste verschillen in inzicht in. Kindt heeft ook twijfels over de uitwerking van 'Wiskunde 12-16' in de huidige schoolboeken. (Die worden overigens door onafhankelijke auteursgroepen ontwikkeld.)

Ook wordt een misverstand rechtgezet: de huidige stilte aan de poorten van de Mathematische Instituten heeft weinig met het wiskundeprogramma van de basisvorming te maken, dat is daar nog te vers voor, en ook niet met een realistische stroming in het VWO-programma; wiskundestudenten van nu hebben immers het VWO-B wiskundeprogramma in hun bagage, dat niemand met de term 'realistisch' in welke betekenis dan ook zal beschrijven. Bovendien blijkt de belangstelling voor de bèta-vakken in Amerika en andere Europese landen de laatste jaren minstens even hard terug te lopen als in Nederland. Verband zoeken met de aard van het Nederlandse wiskundeonderwijs of wijzen op (mogelijke) betrekkelijkheid van internationaal vergelijkingsonderzoek als TIMSS zet weinig zoden aan de dijk.

Een tegenstelling?

In het volgende uitgebreide fragment (Keune, F. 1998b, p. 48) komt een scherpe tegenstelling naar voren:

Het is het samenspel van abstractie, deductie en toepassing dat de wiskunde zo krachtig maakt. Goed wiskundeonderwijs is gericht op dit samenspel van deze aspecten van de wiskunde. Vroeger lag de nadruk op abstractie en deductie en liet men de toepassing geheel over aan andere vakken, voornamelijk aan de natuurkunde.

Tegenwoordig worden in het voortgezet onderwijs abstractie en deductie grotendeels genegeerd en heeft men alleen oog voor het gebruik van de wiskunde. Dit wiskundegebruik is door gebrek aan abstractie en deductie noodzakelijkerwijs van een laag niveau.

Door het wegwerken van abstractie en deductie, twee wezenskenmerken van de wiskunde, blijft iets over dat geen wiskunde meer is. Toch is dit precies wat in de realistische wiskunde gebeurt. Abstracte zaken worden vervangen door zaken uit de leefwereld. De deductie is vervangen door waarnemen.

De tegenstelling is die tussen concreetheid, laag niveau, waarneming, werkelijkheid en toepassingen aan de ene kant en abstractie, deductie en goede wiskunde aan de andere kant. Daarmee wordt de wiskunde als een geïsoleerd, goed herkenbaar gebied voorgesteld, waarin begrippen absoluut helder zijn en geschikt zijn (gemaakt) voor logisch redeneren. Vermoedelijk wordt deze opvatting over de afgrenzing van het vakgebied gedeeld door een fors deel van de wiskundige beroepsgroep, in ieder geval op het moment dat wiskundigen zich openbaar uitspreken over de aard van hun specialisme.

Het accentueren van deze tegenstelling treffen we al bij Plato aan in de Philebus-dialogoog (Plato, ongeveer 360 voor Christus), waarin twee soorten rekenen tegenover elkaar worden gezet. Plato legt Socrates in de mond:

Er is een belangrijk kenmerk van verschil, Protagoras. De gewone rekenaar handelt met ongelijke eenheden: zijn 'twee' kan zijn twee legers, twee koeien of twee wat dan ook, van het kleinste in de wereld tot het grootste; terwijl de filosoof met hem niets van doen wil hebben, tenzij deze toestemt dat elk voorkomen van zijn eenheid precies gelijk is aan alle andere.

In 'de filosoof' herkennen we uiteraard Plato zelf, die zich niet bezig wil houden met voorbijgaande vluchtigheden als legers en koeien, maar wel met de van alle verwaringen ontdane zuivere idee 'twee'. (Op de mogelijkheid ook deze 'twee' met de term 'concreet' aan te duiden, kom ik later nog terug).

Ik wil wijzen op twee andere aspecten

van dit fragment. Ook al voelt Plato zich verderop in de dialoog gedwongen de getallen van de filosoof 'beter' te vinden, hij legt ons in de opsomming 'legers, koeien, wat dan ook, het grootste en het kleinste in de wereld' juist ook de rijkdom van de 'twee' van de gewone rekenaar voor. Het tweede punt is dat de filosoof wel met de gewone rekenaar wil spreken, maar dan onder één voorwaarde: deze moet een bepaalde actie ondernemen, namelijk toestemmen dat elke eenheid gelijk is aan alle andere.

Opmerkelijk is de conclusie die Keune trekt direct na het bovengeciteerde fragment (Keune, 1998b, p. 48):

Realistische wiskunde is dus geen wiskunde. Men doet er beter aan om te spreken van rekenen.

Het is ongebruikelijk om van 'realistische wiskunde' te spreken. In zwang is wel de term 'realistisch wiskundeonderwijs'. Het verschil is essentieel: ook al zou wiskunde een van 'de werkelijkheid' onderscheidbare wereld zijn, dan nog kan het didactisch van groot belang zijn in het onderwijs meer aan bod te laten komen dan wiskunde in zijn uitgekristalliseerde vorm.

Heel grof – maar te beperkt – kan het uitgangspunt van realistisch wiskundeonderwijs toegelicht worden met de twee aspecten die zojuist bij het fragment uit de dialoog van Plato zijn aangegeven: de rijkdom van de werkelijkheid kan een goed uitgangspunt zijn om via een bewuste activiteit te komen tot de wiskunde van de filosoof. De alledaagse werkelijkheid is daarbij absoluut niet het enige uitgangspunt, de essentie van realistisch wiskunde onderwijs ligt veel meer in de nadruk op wiskunde als activiteit, niet als kant en klaar product dat in het onderwijs moet worden 'overgedragen'.

Realistisch wiskundeonderwijs

De theorie van het realistisch wiskundeonderwijs legt onder andere grote nadruk op het proces van mathematiseren, dat in eerste instantie de overgang van concrete ervaringen naar de abstractie omvat, maar ook meer is. Freudenthal gaat in *Revisiting Mathematics Education* (Freudenthal, 1991) dan ook diep in op dit begrip. Het omvat zowel axiomatiseren, formaliseren in de traditionele

zin, maar ook het modernere modelleren en het omvattende 'looking for essentials'. Het is de wiskundige activiteit bij uitstek:

I myself insist on including in this one term the entire organizing activity of the mathematician, whether it affects mathematical content and expression, or more naive, intuitive, say lived experience, expressed in everyday language.

De 'mathematician' is hier duidelijk niet alleen de vakspecialist. De nadruk op activiteit is hier een constante. Iets eerder staat:

To be sure, as popular as axioms and axiomatizing might be in modern mathematics, they are only the highlights and the finishing touches in the course of an activity where the stress is on form rather than on content.

Onder 'looking for essentials' staat een lijstje opgenomen dat er zo uit ziet:

within a situation and across situations,
within a problem and across problems,
within a procedure and across procedures,
within an organisation and across organisations,
within a scheme and across schemes,
within an algorithm and across algorithms,
within a structure and across structures,
within a formulation and across formulations,
within a symbolisation and across symbolisations,
within an axiomatic system and across axiomatic systems.

Waarmee de brede range waarop de activiteit 'mathematiseren' van toepassing is nog eens wordt aangegeven.

Over mathematiseren als de essentiële activiteit is De Lange (1987) zeer duidelijk waar het gaat om wiskunde A (voornamelijk gericht op gebruik van wiskunde in de sociale wetenschappen), maar ook binnen de didactiek van het rekenen is deze gedachte theoretisch verhelderd en door praktijk ondersteund, men zie bijvoorbeeld Streef-land, 1990. Realistisch rekenonderwijs propageert met name ook door reflectie geleide flexibiliteit van werkwijze, in tegenstelling tot het systematisch inslijpen van één enkel algoritme voor het oplossen van een specifiek probleem. Zo zal een leerling niet alleen leren dat $39 + 39$ onder elkaar gezet kan worden om cijferend tot een resultaat te

komen, maar ook merken dat er een route naar de uitkomst is via $40 - 1$ en $80 - 2$. De tegenstelling is weer dezelfde: die van rekenen leren als door reflectie geleide activiteit aan de ene kant en aan de andere kant die van het inslijpen van een algoritme, in de uitgekristalliseerde vorm die door anderen reeds is ontwikkeld. In dit licht gezien is 'rekenen' wiskundig veel rijker dan door Keune in bovenstaand citaat wordt gesuggerd.

Gezegd moet worden dat er op het gebied van wiskunde B (gericht op exacte wetenschappen en techniek) veel minder is gepubliceerd waarin de theorie van het realistisch wiskunde onderwijs expliciet wordt genoemd. Vanuit het Freudenthal instituut gezien is dat goed verklaarbaar: op dit gebied is het Freudenthal instituut betrekkelijk recent bij de ontwikkeling van een curriculum betrokken, in feite pas actief bij de ontwikkeling van de nieuwe bovenbouwprofielen vanaf 1995.

In het advies van de Vakontwikkelgroep Wiskunde, (Sjamaar, e.a., 1995) worden drie hoofdaspecten van de wiskunde genoemd: modelleren, abstraheren en redeneren. Binnen het project Profi van het Freudenthal instituut werd hieraan vormgegeven in het experimenteel ontwerp voor de wiskundeprogramma's voor de profielen Natuur en Techniek en Natuur en Gezondheid, maar men liet zich ook, net als de Vakontwikkelgroep, inspireren door hoofdstuk 3 van het rapport van de commissie Wiskunde B VWO (De Lange, 1994). Daarin wordt een visie op wiskunde leren uiteengezet; gesteld wordt dat begripsontwikkeling een proces is:

De eerste introductie van een bepaald begrip zal soms onvolledig of intuïtief moeten blijven; door bij alle gelegenheden waarin het begrip zich aandient expliciet op de betekenis in te gaan, wordt het begrip gaandeweg preciezer omschreven. Afhankelijk van de aard van het begrip kan een goede definitie in een later stadium worden gegeven, of zal dit wegens te grote moeilijkheidsgraad of complexiteit aan het vervolgonderwijs moeten worden overgelaten (p. 47).

Het gaat hier niet over een aanbeveling hoe

wiskunde aan vakgenoten moet worden gepresenteerd; daar mag, zo men dat wil, een definitie van een begrip aanvang zijn. Wiskundeonderwijs op het VWO is blijkbaar iets anders.

Wiskundedocent G. Huls doet in de *Nieuwe Wiskrant* van juni 1998 (Huls, 1998) verslag van een experiment Voortgezette Meetkunde, uitgaande van in het Profi-project ontwikkeld materiaal. Het verslag opent met een beschrijving van een toepassing maar gaat verderop vooral in op het leren bewijzen. Het verslag is bedoeld voor docenten wiskunde en verwijzingen naar de theorie van het realistisch wiskundeonderwijs zijn daar niet per se nodig. Het had makkelijk gekund; de gebruikte tekst in het experiment richt zich expliciet op de gang van praktische uitgangssituatie naar wiskundige theorie.

Door Keune wordt 'realistische wiskunde' opgevat als blijvend bij de realiteit, in genoemde publicaties worden 'toepassingen' juist als startpunt voor abstractie, formalisatie en deductie gezien. Een groter tegenstelling in interpretatie is nauwelijks denkbaar. Het is goed mogelijk dat de term 'realistisch' dit misverstand in de hand werkt. In de inleiding tot de bundel waarin de bijdrage van Streefland (Streefland, 1990) is opgenomen wordt het misverstand al op de eerste bladzijde genoemd en opgehelderd. In de openbare discussie van dit moment had de verwarring makkelijk vermeden kunnen worden.

Maar er is wel wat aan de hand: algebra

Zoals al vermeld in de reactie van Kindt, moet de huidige praktijk in de onderbouw niet vereenzelvigd worden met wat door voorstanders van realistisch wiskundeonderwijs wordt beoogd. De praktijk van de wiskunde in de onderbouw van het VWO is gevarieerd. Veel docenten kiezen momenteel voor extra activiteiten buiten de methoden om, in feite omdat ze het met Keune eens zijn dat de gangbare wiskundemethoden te weinig uitdaging bieden voor de meer getalenteerde leerling, en ook omdat er te weinig gewerkt wordt aan formalisatie en abstractie. Vaak wordt ook gesteld dat er te weinig aandacht is voor oefenen van vaar-

digheden, met name van algebraïsche aard. Het is nuttig een blik terug in de tijd te werpen, omdat het vaststellen van oorzaak en gevolg hier van belang is.

Vóór de periode waar we ons nu in bevinden werd het onderbouw wiskundeprogramma nog in grote mate bepaald door de verandering van de late jaren 60, die bekend staat als de New Math. De beweging van de New Math – gestart door wiskundigen van professie – beoogde wat nu ook weer voorgesteld wordt: de wiskunde zuiver te houden. Dat werd toen gedaan door het onderwijs op te zetten vanuit een reeds bereikte systematiek van de wiskunde zelf, waarin verzamelingenleer en structuur brengen in verzamelingen de leidraad was. Bij algebra ging het meer om de kennis van en inzicht in 'wetten' als de distributiviteit en bij meetkunde was het abstracte concept transformatie belangrijker dan het te transformeren object zelf.

Met veel idealisme werd dit concept voor het onderwijs uitgewerkt maar binnen dertig jaar wordt overal afstand genomen van dit gedachtegoed.

In een Cito-examenopgave uit de nadagen van de New Math werd gevraagd naar het aantal elementen van de verzameling S , gedefinieerd door

$$S = \{ z \in \mathbb{Z} \mid z^2 + z + \frac{1}{4} = 0 \}$$

Er was keuze uit 0, 1, 2, of meer.

De taal is die van de formele wiskunde; hoe ver de ontaarding gaat blijkt vooral uit de toelichting van het Cito, waarin verwacht werd dat de vergelijking opgelost zou worden met behulp van de algemene methode voor het oplossen van vierkantsvergelijkingen: de bekende wortelformule. Dat zelfs binnen de New Math-benadering een triviale oplossing voor de hand ligt (z is geheel, dus $z^2 + z$ ook, dus is $z^2 + z + \frac{1}{4}$ niet geheel en daarom ongelijk nul), werd door niemand meer gezien.

Het (extreme, maar niet alleenstaande) voorbeeld illustreert hoe door te veel nadruk te leggen op de abstracte kant van de wiskunde, de wiskunde in het onderwijs ont-aardde in een loos ritueel waar in de uiterlijke schijn van het vak – symbolen en

formules – werd opgehouden maar waarin in ieder geval wiskundigen hun vak niet herkennen.

De operatie W12-16 van de jaren 1987-1992 was bedoeld om een programma te ontwikkelen en uit te testen dat betekenisvolle wiskunde voor iedereen moest bevatten. In zo'n programma heeft de genoemde opgave geen plaats.

Uit het experiment W12-16 kies ik als voorbeeld de volgende opgave die tot het experimentele examen 1992 hoorde.

De tabel in de bijlage [hier niet afgedrukt] gaat over $x^3 + x$ en $\frac{1000}{x}$.

11. Kies voor x de waarden 2, 5 en 8 en vul de tabel in.
12. Reken uit hoeveel verschil er is tussen $x^3 + x$ en $\frac{1000}{x}$ als $x = 100$.
13. Gebruik de getallen uit de tabel en schets in het assenstelsel op de bijlage de grafieken van de functies $x \rightarrow x^3 + x$ en $x \rightarrow \frac{1000}{x}$.
14. Geef een waarde voor x waarvoor $x^3 + x$ en $\frac{1000}{x}$ ongeveer aan elkaar gelijk zijn. Bereken hoe groot het verschil dan is.
15. Schrijf op hoe je een waarde voor x kunt vinden waarvoor het verschil tussen $x^3 + x$ en $\frac{1000}{x}$ kleiner is dan in vraag 14.

De meerkeuzevorm is verlaten. De opgave begint met twee betrekkelijk gemakkelijke onderdelen, maar daarna wordt de leerling gevraagd zelf een methode onder woorden te brengen. Dat zal een vorm van steeds beter benaderen van de oplossing van de vergelijking zijn, een wiskundig proces dat hier misschien onschuldig aandoet, maar makkelijk aanleiding kan zijn voor voortzetting op hoger niveau.

De opgave bevat geen verwijzing naar een concrete werkelijkheid maar heeft wel de sfeer van 'realistisch wiskundeonderwijs' in de zin dat het gaat om een situatie waarin de leerling zich kan inleven, waarin de situatie dus realiteit wordt en waarin de leerling zelf constructies kan uitvoeren. Ook kan opgemerkt worden dat in het bedoelde oplossingsproces – het geleidelijk inklemmen van de gezochte oplossing – het concept van oplossen van een vergelijking helder zichtbaar blijft.

Ook de inklempogave heeft de mogelijkheid van ontarding in zich, kan aangediend worden als recept dat (bijna) altijd benaderen van een oplossing van een vergelijking toelaat. Dat is echter niet de opzet van het voorgestelde algebraprogramma, zoals bij inzage van *Achtergronden van het Nieuwe leerplan wiskunde 12-16* (Abels, e.a., 1992) snel blijkt.

Wie de opgave nu ziet, kan licht de indruk krijgen dat het niveau ervan niet hoog is. Zo wordt er alleen maar ingevuld in formules en worden geen typisch algebraïsche manipulaties vereist. Toch wezen veel docenten er in de projectperiode herhaaldelijk op dat het projectteam te hoge verwachtingen had van wat mogelijk was, als een dergelijke opgave werd voorgesteld. Die waarschuwingen hebben hun weerslag gehad.

Terugkijkend kan vastgesteld worden: in het project W12-16 is veel aandacht geweest voor de middengroep van leerlingen, die van de Mavo en ook voor de VBO-stroom, zij het in een later stadium, maar voor de groep leerlingen die het VWO ingaat betrekkelijk weinig. Dat is indertijd al binnen de Commissie Ontwikkeling Wiskundeonderwijs, die het project begeleidde, geconstateerd; de hoop was misschien dat de auteurs van leerboeken voldoende aanvulling zouden bieden. Voor dezen geldt echter nog meer dat ze luisteren naar wat docenten van nu met leerlingen van nu in het algemeen denken te kunnen bereiken.

Keune is ook niet de enige die vaststelt dat de toekomstige VWO-leerling hierdoor vaak niet voldoende uitgedaagd wordt, veel leerlingen uit die groep zeggen hetzelfde.

Of valt het mee: meetkunde en groei

In zijn rede citeert Keune enkele fragmenten uit *Netwerk*. De gekozen voorbeelden gaan over berekenen van oppervlakte van figuren met punten op een rooster en het gebruik daarvan in relatie tot de stelling van Pythagoras. Een bewijs was in *Netwerk* niet te vinden, aldus Keune. Hij vergezelt het voorbeeld van een bewijs, zoals hij dat zelf heeft overgeschreven van het bord van de heer Peters op de HBS.

Zonder de aanpak van *Netwerk* te verde-

digen – dat laat ik over aan de betreffende auteursgroep – wil ik toch wel de vraag stellen: wordt bewijzen dan geleerd door overschrijven?

Het antwoord op de vraag is vaak 'ja, maar' geweest. De leerlingen die het vroeger op deze manier inderdaad leerden, waren er wel. Maar wiskundedocenten die in de bedoelde periode actief waren, vertellen dat de groep die het leerde klein was. Overigens: niet alle leerlingen uit die kleine groep zijn zich daarvan achteraf nog bewust. De meetkunde die nu in de onderbouw van het VWO in de boeken verschijnt is minder direct op deductie en geheel niet op werken binnen een axiomatisch systeem gericht. Ruimtelijke oriëntatie speelt er daarentegen een grote rol, leren zien van patronen ook en het gebruik maken daarvan bij redeneringen en berekeningen. Reflectie is er zeker wel, activiteiten die voorbereiden op bewijzen ook. De gewraakte opgave van Netwerk hoort daar bij, al blijft het ontbreken van een verdergaande redenering kwalijk. Deductie blijft in de onderbouw vaak lokaal en is weinig gesystematiseerd. De vraag is hoe ernstig dat is, als we kijken naar het vervolg.

Zoals al gezegd werd, komt het klassieke bewijzen in de meetkunde weer terug in het VWO, maar nu wel in een breder kader waar ook toepassingen bij horen en alleen voor die groep waarvoor het redelijkerwijs zinvol en haalbaar lijkt: binnen het profiel Natuur en Techniek.

Opmerkelijk is dat bij het experimentele examen van 1997 ongeveer de helft van de huidige wiskunde-B groep tot de kern van de zaak komt bij de examenopgave van 1997. Zie de bespreking in de *Nieuwe Wiskrant* (Lagerwaard & Lent, 1997). Van de helft van deze groep mag verwacht worden dat ze, was de herprofilingsoperatie voltooid, het profiel N&T gekozen zou hebben; de andere helft zou vermoedelijk N&G gekozen hebben, waarin het onderwerp Voortgezette Meetkunde geen plaats heeft. Bij het begin van het experiment heb ik er al op gewezen dat het leeftijdsverschil tussen de meetkundeleerlingen van voor 1968 en die van nu een positief effect heeft op de houding ten opzichte van het bewijzen (Goddijn, 1995). Zo stellen leerlingen van

nu soms zelf kritische vragen over onderbouwing van de meetkunde: een niveau van actief met axiomatiek bezig zijn dat vroeger niet voorkwam.

Bij observeren van het meetkunde-experiment in de klas bleek dat de problemen die leerlingen met bewijzen in de meetkunde hebben zelden met het gebruik van de logica te maken hebben, maar meer met het herkennen van de juiste componenten in een figuur waardoor het bewijs gaat lopen. De leerlingen die nu meedoen in het experiment hebben nog een tamelijk kaal meetkunde-programma genoten in de onderbouw. De hoop bestaat dat de nieuwe onderbouw, waarin meer aan intuïtief meetkundig inzicht wordt gewerkt eerder een positieve uitwerking zal hebben dan een negatieve.

Een ander voorbeeld van een lijn van onderbouw naar VWO-top is het onderwerp 'groei'. In de nieuwe onderbouw komen expliciet lineaire, kwadratische en exponentiële groei ter sprake; het algebraprogramma is namelijk breder van opzet dan de vroeger heersende microscopie van de kwadratische functie. In het kader van toepassingen wordt verkend hoe deze groeitypen zich tot elkaar verhouden, welk groeitypen sterker zijn dan anderen. Men kan zich hier gemakkelijk talloze praktijksituaties bij voorstellen. In de Voortgezette Analyse van het VWO profiel N&T wordt daar nadrukkelijk op voortgebouwd, maar nu in de formele vorm van klassieke limieten als bijvoorbeeld

$$\lim_{x \rightarrow \infty} \frac{x^2}{2^x} = 0$$

die hier ook bewezen worden.

Vormende waarde

Keune wijst op de gevaren van het veronachtzamen van abstractie en deductie in de onderbouw van het voortgezet onderwijs. Hij noemt onder andere:

Er is geen training in logisch redeneren, hetgeen ook maatschappelijke consequenties heeft: logisch kunnen denken is voor iedereen van grote waarde (Keune, 1997c).

Er is nog nooit expliciet transfer aangetoond van leren logisch denken in de zuiver wiskundige context naar contexten daarbuiten. Het zijn vaak de wiskundigen geweest die dit

geloof in de vormende waarde van het vak naar voren brengen. Dijksterhuis, eminent historicus van de wiskunde die zich gepassioneerd met hervormen van het wiskundeonderwijs bezig hield, was een groot voorbeeld, zie de inmiddels bekende biografie van K. van Berkel uit 1996.

Er zijn ook argumenten naar voren te brengen waarom die transfer helemaal niet zo waarschijnlijk is. Keune zelf stelt bijvoorbeeld dat de abstracte en welgedefinieerde objecten van de (zijn) wiskunde de helderheid hebben waardoor zuiver logisch denken erover mogelijk is. Daarmee wordt dan feitelijk ook gesteld dat de minder welgedefinieerde onderwerpen waar men zich buiten de wiskunde mee bezighoudt, zich daar minder voor lenen. Het lijkt er dan ook eerder op dat het een gevaarlijke schijn van logisch denken oplevert als de redeneertechniek die bij de wiskunde opgeld doet, daarbuiten wordt toegepast.

Met weinig succes is ook geprobeerd door expliciet logica te onderwijzen, studenten op een hoger plan van denken over hun eigen vak – inclusief wiskunde – te brengen. Bij de beschrijving van ervaringen met de VWO-meetkunde werd al iets dergelijks gezegd: de logica zelf is zelden het probleem, de wiskundige inhoud (de meetkundige objecten) maakt het moeilijk een juiste bewijsweg te vinden.

Lang voor de termen 'vormende waarde' en 'transfer' bestonden sprak Blaise Pascal – eminent wiskundige uit de 17e eeuw – zich er in zijn *Esprit de Géométrie et Esprit de Finesse*, van vóór 1662, al over uit. Heeft men eenmaal, zo zegt hij, het hoofd zo gedraaid dat men de objecten van de meetkunde kan zien, dan is redeneren ermee gemakkelijk. Maar in het dagelijks leven is het andersom: daar zijn de objecten waar men over spreekt een ieder bekend, maar ze zijn met vele en zijn niet helder omschreven; daardoor is het redeneren ermee uiterst lastig. Pascal voegt eraan toe, dat wie geheel doortrokken is van de *Esprit de Géométrie*, niet noodzakelijk ook een meester is van de *Esprit de Finesse*, en andersom.

Nu bedoelt Keune misschien ook meer dat toepassingen van de wiskunde van groot maatschappelijk belang zijn. Dat is juist,

maar het is een veel minder vergaande claim.

Abstractie en de realiteit van de wiskunde

In zijn rede stelt Keune:

Wat is er volgens mij nu echt mis met ons wiskundeonderwijs? De meest diepgaande fout is het verwarren van abstractie en realiteit. Abstractie is ontstaan uit de behoefte aan helderheid. In de hersenspinsels van de wiskunde is er sprake van absolute helderheid. In ons onderwijs wordt dat mistig gemaakt en dat is zonde. Het is een gemiste kans voor het leren van exact denken (Keune, 1998a).

Ook Schalkwijk (1998) stelt dat bij het leren bewijzen het scheiden van realiteit en wiskunde van belang is:

Vermenging van leefwereld en wiskunde frustriert het bewijzen; de grens daartussen moet duidelijk worden gemarkeerd, zeker niet worden verdoezeld, om te kunnen begrijpen wat de betekenis van beide werelden voor elkaar is (p. 22).

Schalkwijk citeert met instemming hierbij Freudenthal (1973):

... Sie [Wiskunde en werkelijkheid] sollen ihm [de leerling] nicht getrennt vorgesetzt werden. Er soll es lernen, sie selber zu trennen, so bald ihre gegenseitige Beziehungen so innig geworden sind, dass die Trennung keiner Seite schadet (p. 527).

Er is nog wel groot verschil tussen 'de grens daartussen moet worden gemarkeerd' en 'soll es lernen, sie selber zu trennen', maar Schalkwijk, die probeert leerlingen (een selectie van de VWO-B leerlingen) in een speciale cursus actief en onderzoekend met wiskunde bezig te laten zijn, blijkt het het principe van actief mathematiseren in de praktijk van zijn onderzoek wel te huldigen. Kan men bij het voorgaande nog spreken van een gang van concreet naar abstract, bij het bestuderen van wiskunde lijkt vaak ook het omgekeerde plaats te hebben, zij het in een iets andere zin.

Professor N. Kuiper heeft tijdens een college Lineaire algebra voor mij eens mooi verwoord wat er mis is met het begrippenpaar concreet/abstract. Wat concreet of ab-

abstract is, hing volgens hem niet van het onderwerp af, maar van de betrokken persoon. Zo kan een getal als 5 voor sommige kinderen nog abstract zijn, ook in de zin dat het nog realiteit-gebonden is. Voor anderen kan 5 al zeer concreet zijn, ze weten precies wat er mee gedaan kan worden en wat niet. Voor hen kan een getal als de wortel uit min één echter weer onoverkomelijk abstract zijn, terwijl een geschoold wiskundige zich in een omgeving van dit getal even thuis voelt als een schilder bij zijn kwast.

Ofwel: abstractie is weliswaar iets dat veroverd kan worden, maar kan dan ook verduwen zijn: de verovering van het abstracte maakt de zaken concreet. Zo ook het al genoemd rapport Wiskunde B VWO:

'Beauty is in the eye of the beholder', en abstractheid ook. Het leven is vol concepten die bij eerste kennismaking abstract aandoen maar waar men gaandeweg mee vertrouwd raakt, zodat het gevoel van abstractie verdwijnt (p. 51).

De ander zin is: Keune spreekt over het a priori vaststaan van de abstractheid van de objecten waar de wiskunde zich mee bezighoudt; Kuiper en het rapport Wiskunde B spreken over de subjectieve veranderlijkheid van de abstractheid; abstractheid is dan meer een kwalificatie hoe een object of begrip gezien of beleefd wordt.

Realistisch wiskundeonderwijs gaat dan ook niet – zoals gesuggereerd wordt door Keune – over concrete materiële realiteiten, maar wil steeds uitgaan van wat voor de leerling realiseerbaar is. Dat kunnen zaken zijn die tot normaliter 'concreet' worden genoemd, maar ook begrippen die traditioneel tot de zuivere wiskunde behoren.

Vanuit filosofisch standpunt gezien is het onderscheid tussen abstractie en realiteit ook al een hachelijke zaak, omdat natuurlijk ook de vraag beantwoord moet worden: wat wordt met realiteit bedoeld? Een naïeve beperking tot wat materieel is, voldoet natuurlijk niet. Zonder hier op dit gebied vol voetangels en klemmen in te gaan, wil ik nog wel wijzen op het feit dat diverse filosofen die zich momenteel bezig houden met de filosofie van de wiskunde, ontevreden zijn met het onderscheid dat gemaakt wordt over de status van de wiskundige beweringen

enerzijds en de beweringen binnen andere disciplines anderzijds. Met status bedoel ik: welk waarheidscriterium moet er voor dergelijke uitspraken worden aangelegd? De wiskunde afdoen als afzonderlijk zuiver formeel systeem leidt ertoe dat waarheid niet meer kan zijn dan interne consistentie. De praktizerend wiskundige hoeft daar niet van wakker te liggen, maar vanuit filosofisch standpunt blijft het een kunstgreep die het onderliggende probleem vermijdt.

Tegenover deze formalistische visie op wiskunde (veelal geassocieerd met Hilbert) staat de platonistische, die in zekere zin het bestaan van de wiskundige werkelijkheid erkent, ongeveer zoals Plato de Ideeën als werkelijkheid en koeien en legers als vluchtige afschaduwingen ervan beschouwde.

Tot deze groep behoort onder ander Kurt Gödel; hij stelde, sprekend over 'the objects of transfinite set theory':

But, despite their remoteness from sense experience, we do have a perception also of the objects of set theories, as is seen from the fact that the axioms force themselves upon us as being true. I don't see why we should have less confidence in this kind of perception, i.e. in mathematical intuition, than in sense perception, which induces us to build up physical theories and to expect that future sense perceptions will agree with them and, more, to believe that a question not decidable now has meaning and may be decided in the future.

Aldus geciteerd in 'Mathematical Truth, door Paul Benacerraf, zie *The Philosophy of Mathematics* (ed. Hart 1997). In deze bundel zijn ook de bijdragen van o.a. Hart, Maddy en Tait op dit punt interessant voor de filosofisch geïnteresseerde.

Tot slot

In het voorgaande zijn diverse visies op wiskunde en wiskundeonderwijs aan bod geweest. De formalistische visie en de platonistische zijn – weer, nog steeds – punt van discussie.

Davis & Hersch (1981) halen dezelfde passage van Gödel aan bij hun beschrijving van de platonistische visie, maar noemen ook René Thom, die niet vanuit de verzamelingenleer, maar vanuit de meetkunde

spreekt. Ze stellen ook dat het in de dagelijkse wiskundige praktijk niet zo veel uitmaakt en vatten dat samen met:

Most writers on the subject seem to agree that the typical working mathematician is a Platonist on weekdays and a formalist on Sundays. That is, when he is doing mathematics he is convinced that he is dealing with an objective reality whose properties he is attempting to determine. But then, when challenged to give a philosophical account of this reality, he finds it easiest to pretend that he does not believe in it after all (p. 321).

Evenzo kan men eerst benadrukken dat beroepswiskundigen en onderwijsontwikkelaars met een realistische visie op wiskundeonderwijs fundamenteel verschillende visies op het vak kunnen hebben, om er geleidelijk achter te komen: beiden gaat het om 'goed' wiskundeonderwijs. Laten beide groepen in overleg blijven met respect voor elkaars uitgangspunten, werkelijke en belede-

Literatuur

- Abels, M. e.a. (1992) *Achtergronden van het Nieuwe leerplan wiskunde 12-16*. Freudenthal instituut, RU Utrecht.
- Davis, P. J. & Hersch, R. (1981). *The Mathematical Experience*. Penguin Books.
- Delft, D. van (1998). Pannenkoekwiskunde. In: *Bijlage Wetenschap en Onderwijs van de NRC*, 25 april 1998.
- Freudenthal, H. (1991). *Revisiting Mathematics Education; China lectures*. Kluwer Academic Publishers, Dordrecht, Boston, London.
- Freudenthal, H. (1973). *Mathematik als pädagogische Aufgabe, I & II*. Klett Studiebücher, Stuttgart.
- Goddijn, A. (1995). De diepte van het platte vlak. *Nieuwe Wiskrant* 15-3, maart 1995, pp. 6-12, Freudenthal instituut, Utrecht.
- Hart, W. D. ed. (1997) *The Philosophy of Mathematics*. Oxford University Press.
- Huls, G. (1998) Voronoidiagrammen en bewijzen in de klas. *Nieuwe Wiskrant* 17-4, juni 1998, pp. 18-22, Freudenthal instituut, Utrecht.
- Keune, F. (1998a). *Naar de knoppen*. Inaugurale rede 17 april 1998, Katholieke Universiteit Nijmegen. <http://www.sci.kun.nl/math-keune>.
- Keune, F. (1998b). *Naar de knoppen*. *Nieuwe Wiskrant* 17-4, juni 1998, p. 48, Freudenthal instituut, Utrecht.
- Kindt, M. (1998). Beknopt weerwoord. *Nieuwe Wiskrant* 17-4, juni 1998, p. 49, Freudenthal instituut, Utrecht.
- Lagerwaard, K. & Lent, G. van. (1997). Q. E. D. Meetkundebewijzen in het eindexamen. *Nieuwe Wiskrant* 17-2, december 1997, p. 4-9, Freudenthal instituut, Utrecht.
- Lange, J. de. (1987). *Mathematics, Insight and Meaning*. Vakgroep Onderzoek Wiskundeonderwijs en Onderwijscomputercentrum, Rijksuniversiteit Utrecht.
- Lange, J. de e.a. (1994). *Rapport studietoetscommissie Wiskunde B VWO*.
- Pascal, B. (voor 1662). Difference entre l'esprit de géométrie et l'esprit de finesse. Nr. 512 (nummering Lafuma) van de *Pensées; Oeuvres Complètes*, Editions du Seuil, 1963).
- Plato, (-360). Philebus, geciteerd in Fauvel J. & Gray, J.: *The History of Mathematics, a reader*; Macmillan Press 1988.
- Schalkwijk, L. T. J. M. van. (1998). *Onderzoekend Wiskunde Lereren*. Proefschrift Katholieke Universiteit van Nijmegen.
- Sjamaar, M. e.a. (1995). *Advies examenprogramma's havo/vwo wiskunde*. Den Haag, 1995.
- Streefland, L. (1990) Realistic Mathematics Education (RME), What does it mean? In: *Context, Free productions, Tests and Geometry in Realistic Mathematics Education*, K. Gravemeijer, M. van den Heuvel, L. Streefland; researchgroup for Mathematical Education and Educational Computer Centre, State University of Utrecht.

Kwaliteit wiskundeonderwijs

VWO ter discussie

Frans Keune

Onderwijsinstituut Wiskunde, Katholieke Universiteit Nijmegen

Inleiding

In de door mij op 21 april 1998 uitgesproken inaugurele rede (Keune, 1998a) heb ik betoogd dat het huidige wiskundeonderwijs niet voldoende recht doet aan de aard van de wiskunde en dat het getalenteerde leerlingen geen uitdaging biedt. Verder kondigde ik aan actie te zullen ondernemen om de situatie te verbeteren. De rede heeft aanleiding gegeven tot discussies over de doelstellingen van het wiskundeonderwijs. Zulke discussies zijn hard nodig, want er is te weinig belangstelling voor de exacte vakken en we zullen ervoor moeten zorgen dat het onderwijs zodanig is dat dit niet aan het onderwijs zal liggen. Ik heb vele positieve reacties ontvangen, waaronder van veel van wiskundeleraars. Negatieve reacties waren er van toenmalig staatssecretaris Netelenbos en van het Freudenthal instituut. In *NRC Handelsblad* werden mijn ideeën door Dirk van Delft op uitstekende wijze verwoord (Van Delft, 1998). Zijn artikel was vergezeld van een reactie van de directeur van het Freudenthal instituut, Jan de Lange. In de *Nieuwe Wiskrant*, een blad van het Freudenthal instituut, mocht ik een samenvatting van mijn standpunt geven. Deze werd gevolgd door een commentaar van Martin Kindt van het Freudenthal instituut (Kindt, 1998). Beide reacties kwamen erop neer dat ik niet goed weet waar ik het over heb. De tweede verwoordde dat overigens op een meer diplomatieke wijze dan de eerste. Nu is er ook het uitvoerige standpunt van Aad Goddijn van het Freudenthal instituut (Goddijn, 1998). Dit artikel is een reactie op zijn beschouwingen.

De aard van de wiskunde

Voordat we een idee kunnen hebben van wat goed wiskundeonderwijs inhoudt, moeten we eerst een duidelijk beeld hebben van de wiskunde zelf: van de werkwijze van de wiskunde, van het gebruik van de wiskunde en niet in de laatste plaats een beeld van waar wiskunde over gaat. Iedere wetenschap heeft een object. Het object van de wiskunde bestaat uit abstracties. Het gaat om eigenschappen van abstracte zaken. Een lijn in de wiskunde is iets anders dan een potloodstreep op een vel papier. De wereld van de wiskunde is een bedachte wereld. Nu is het moeilijk te communiceren over bedachte zaken, want waarom zou iedereen daar hetzelfde bij denken? Om dat probleem te omzeilen leggen wiskundigen zich beperkingen op: alles moet gebaseerd zijn op harde afspraken over de eigenschappen van deze abstracte zaken. Van de buitenkant heeft de wiskunde daardoor een zeer formeel karakter gekregen. Deze formele verpakking moet niet verward worden met de inhoud. De filosofische kwestie of het object van de wiskunde bestaat of niet is voor wiskundigen weinig interessant en heeft geen invloed op de praktijk van de wiskundebeoefening. Als je vindt dat abstracties bestaan, dan ben je een platonist. Vind je dat zij niet bestaan en dat alleen de formele verpakking bestaat, dan ben je een formalist. Essentieel is dat het abstracte karakter niet wordt ontkend. Het is juist door de abstractie dat de wiskunde zo veelvuldig toepasbaar is, dat hij van pas komt bij zeer uiteenlopende vakgebieden. Sommigen voelen zich gelukkig in de abstracte wereld van de wiskunde en hebben weinig behoefte aan het toepassen. Anderen zijn juist erg gericht op het toepassen. Het gaat in beide gevallen wel om dezelfde wiskunde, al is er een verschil in belangstelling.

Goed wiskundeonderwijs en de huidige praktijk

De abstractie in de wiskunde is niet uit de lucht komen vallen. Zij vindt haar basis in de werkelijkheid. Ook zijn er verschillende niveaus van abstractie te onderscheiden. Bij het leren van wiskunde doet men er goed aan pas tot een hoger abstractieniveau over te gaan als men op

het actuele niveau met de materie vertrouwd is. In het onderwijs moet men daar goed rekening mee houden. Deze opvatting verschilt in essentie niet van hetgeen Goddijn ziet als realistisch wiskundeonderwijs. Verder is hij ook in overeenstemming met de door Goddijn aangehaalde opvattingen van Kuiper.

De kritiek in mijn oratie betrof vooral de praktijk van het wiskundeonderwijs. Het kan best zijn dat mijn kritiek "niet strookt met wat er in de literatuur hierover (het realistisch wiskundeonderwijs) te vinden is". Mogelijk is het "ongebruikelijk" om van realistische wiskunde te spreken, maar in de praktijk gebeurt dat wel: er was zelfs een wiskundemethode die zo heette. In het wiskundeonderwijs moet er een voorbereiden op een hoger abstractieniveau zijn. Hoe lang dat voorbereiden moet duren hangt af van de capaciteiten van de leerlingen en daarmee van het schooltype. Zeker voor VWO-leerlingen duurt dit voorbereiden momenteel zo lang dat het moeilijk valt dit als voorbereiden te zien. Er is sprake van infantilisering: jaren achtereenvolgend gaan opgaven over Pietje en Marietje. Abstractie is ver te zoeken: als een opgave gaat over een chocoladereep, dan staat er een kleurenafbeelding van een chocoladereep bij. De voor de wiskunde belangrijke logische samenhang is niet aanwezig. Anders gezegd: aan bewijzen wordt niet gedaan. Nu kan het best zijn dat men op het Freudenthal instituut ook 'twijfels heeft' over de uitwerking van Wiskunde 12-16 in de huidige schoolboeken, aan het destijds door dit instituut gestelde voorbeeld kleven precies dezelfde bezwaren.

Natuurlijk is het programma van de basisvorming nog te recent om het direct in verband te brengen met de geringe belangstelling voor bètastudies. Wel is er sprake van een ontwikkeling in het wiskundeonderwijs die al veel langer gaande is. Ook zijn Wiskunde I en II vervangen door Wiskunde A en B, dus moet men niet doen alsof er decennia lang niets gebeurd is. Om mijn ongelijk aan te tonen wordt er een TIMSS-rapport bijgehaald. Daarop wees ik (Keune 1998c) op de betrekkelijkheid van dat rapport, maar volgens Goddijn kan ik dat maar beter niet doen, want dat "zet weinig zoden aan de dijk". Dat de belangstelling voor bètastudies in andere landen ook terugloopt zou verder aantonen dat er in Nederland geen verband is tussen deze afnemende belangstelling en het

wiskundeonderwijs. Er zijn overigens landen waar het duidelijk beter gaat. Laten we proberen daar wat van te leren. Als er iets te verbeteren is, dan moeten we geen redenen gaan zoeken waarom we het zouden nalaten.

Tegenstelling

Door Goddijn wordt herhaaldelijk een tegenstelling geschetst: enerzijds wiskunde als activiteit (het mathematiseren) en anderzijds als kant en klaar product dat in het onderwijs moet worden 'overgedragen'. Dit met de suggestie dat ik voor het laatste zou zijn. Onder mathematiseren valt het proces van overgang van concrete ervaringen naar abstractie, het axiomatiseren, het formaliseren, het modelleren en ook 'looking for essentials'. Volgens Goddijn is dit de wiskundige activiteit bij uitstek. Het belang van dit mathematiseren wil ik geenszins ontkennen. Waar ik mij tegen verzet is het idee dat, als je dit zou weglaten, er alleen maar hapklare brokken wiskunde overblijven. Dat het alleen maar zou gaan om nadoen van wat anderen je hebben voorgedaan. Wiskundige opgaven die beginnen met 'Bewijs dat...' hebben als oplossing een *redenering* waaruit blijkt dat hetgeen te bewijzen was waar is. Zulke opgaven zijn er in vele gradaties van moeilijkheid. Het kan inderdaad gaan om nadoen van wat al eerder is voorgedaan. Het kan gaan om een test of theorie begrepen is. Het kan een opgave zijn die inzicht vereist en waar creativiteit voor nodig is. Een opgave kan heel goed een uitdaging zijn zonder dat er sprake is van mathematiseren. Uitdagingen worden nu in het geheel niet geboden. Dit soort uitdagingen valt geheel buiten de door Goddijn beschreven tegenstelling. Het lijkt wel of het buiten zijn blikveld valt.

New Math en andere vernieuwing

In de jaren zestig is de New Math in het wiskundeonderwijs opgekomen. Verzamelingen hadden daarin een centrale rol. Daar is men terecht van teruggekomen. Zeker voor de onderbouw betrof het een verkeerd niveau van abstractie. Goddijn geeft een voorbeeld van een Cito-examenopgave uit die tijd en verge-

lijkt dat met een examenopgave van het experimentele examen 1992 van het experiment W12-16. De Cito-opgave vraagt naar het aantal elementen van de verzameling $\{z \in \mathbb{Z} \mid z^2 + z + \frac{1}{4} = 0\}$. Ontdaan van de verzamelingsnotatie is dit de vraag naar het aantal gehele getallen z waarvoor $z^2 + z + \frac{1}{4} = 0$. Ook als de vraag zo gesteld is, is het heel goed mogelijk dat het Cito verwacht dat de vergelijking eerst met de bekende wortelformule wordt opgelost. Het is mij helemaal niet duidelijk dat zo'n onhandige oplossing het resultaat is van de formele verpakking. Ook nu is het wiskundeonderwijs vol met loze rituelen waar de uiterlijke schijn van het vak wordt opgehouden en waar wiskundigen hun vak niet in herkennen.

Van het voorbeeld uit het experimentele examen 1992 is het mij niet duidelijk wat daar zo goed aan is. Als het een eerste bezinning is op iets dat nog gaat komen, dan kan ik mij er wel iets bij voorstellen. Het is echter een examenopgave. Het is niet de bedoeling van een examen om er iets van te leren, maar om iets te testen. Deze opgave gaat over $x^3 + x$ en $\frac{1000}{x}$. De opgave test of de leerling voor x de getallen 2, 5 en 8 kan invullen, of hij voor $x = 100$ het verschil van de twee waarden kan uitrekenen, of hij grafiekjes van deze functies van x kan schetsen, of hij met de geschetste grafiekjes een x kan vinden waar de waarden ongeveer gelijk zijn en of hij ook nog een betere x kan vinden. Dit is al met al niet meer dan: kan de leerling een grafiekje van een eenvoudige functie schetsen en begrijpt hij een beetje wat hij gedaan heeft. Ik ben benieuwd hoeveel punten een leerling hier zou krijgen als hij de vergelijking $x^3 + x = \frac{1000}{x}$ echt opgelost had, want dat was niet de bedoeling van de opgave. Is dit nu een voorbeeld van hoe het moet zijn? Ik ben bang dat Goddijn en ik op geheel verschillende golflijnen zitten.

De titel van de paragraaf waarin Goddijn deze opgaven beschrijft is: "maar er is wel wat aan de hand: algebra". Hoe moet ik dat nu lezen? Is er iets mis met algebra? Is dat de boosdoener? Zijn het juist de algebraïci die er zulke bizarre ideeën op na houden? Heb ik misschien een verkeerde kijk op het wiskundeonderwijs juist omdat ik een algebraïcus ben? Ik houd het er maar op dat de titel wat ongelukkig gekozen is. De examenopgave van het experimentele examen 1992 gaat over een in wezen

algebraïsch probleem zonder dat de vragen ook maar iets met algebra te maken hebben.

Valt het mee?

In mijn oratie geef ik enkele fragmenten uit de wiskundemethode Netwerk die betrekking hebben op de Stelling van Pythagoras. Die fragmenten heb ik gekozen omdat ze duidelijk illustreren hoe bedroevend het met het wiskundeonderwijs gesteld is. Daar heb ik voor de aardigheid een bewijs bij gegeven dat ik vroeger op de 2e Openbare Handelsschool te Amsterdam had gekregen. Dat was van het bord overgeschreven in een fotokopieeloos tijdperk. Het is flauw om daaruit de conclusie te trekken dat ik vind dat bewijzen worden geleerd door overschrijven. Het klinkt heel interessant als men stelt dat deductie in de onderbouw vaak lokaal en weinig gesystematiseerd is. Ook de opgave uit Netwerk ziet Goddijn er als één die voorbereidt op deductie. Het betreft een opgave die ik heb genoemd als een voorbeeld waar men de plank mis slaat. Maar niet getreurd, het bewijzen komt weer terug op het VWO en wel precies waar het volgens Goddijn thuis hoort: binnen het profiel Natuur en Techniek. Het plan is nu om in de zesde klas van het VWO iets aan bewijzen, te doen. Het gaat om wiskunde waar men vroeger in de eerste klas van de MULO al kennis mee kon maken: de vlakke meetkunde. Let wel: dit is alleen voor de 'bollebozen' die voor het profiel Natuur en Techniek hebben gekozen. Op grond waarvan kies je voor dit profiel? Misschien omdat je goed bent in exacte vakken waar weinig exacts in te beleven valt en daar dan nog meer van wil of omdat het zo goed voor je carrière is? Hopelijk is het de goede invloed van een leraar die snapt wat er aan de hand is. Een andere positieve ontwikkeling is volgens Goddijn de nieuwe behandeling van het onderwerp groei in een lijn van onderbouw naar VWO-top. Zolang, zoals nu in de nieuwe boeken voor de vierde klas, eigenschappen van logaritmen 'bewezen' worden door de leerlingen een vraagje te stellen over de groei van een bacteriekolonie en een paar toetsen van een rekenapparaat te laten indrukken, zal het met het redeneren in de wiskunde niets worden.

Als leerlingen niet leren redeneren over een-

voudige wiskundige begrippen, dan missen ze een voorbereiding op het redeneren in het algemeen. Het gaat niet om het expliciet onderwijzen van logica, dat is een stap te ver. Ook is het kunnen redeneren in de wiskunde geen garantie voor verstandig handelen daarbuiten. Daar komt meer bij kijken. Het leren bewijzen in de wiskunde is een vorm van hersengymnastiek, een training van de hersenen. Het is een spel met 'waar' en 'onwaar', van nut voor iedereen. Zeker voor hen die zich gaan begeven in wetenschap, rechtspraak, politiek, journalistiek en onderwijs. Het is dus van groot maatschappelijk nut.

Beroepswiskundigen

De discussie gaat over onderwijs in wiskunde. Het valt op dat hierbij steeds opnieuw aan de deskundigheid van beroepswiskundigen wordt getwijfeld. Er wordt beweerd dat wiskundigen zich daar niet mee zouden moeten bemoeien. Ook bij Goddijn gebeurt dat, zij het dat het slechts om suggestieve opmerkingen gaat. Enkele citaten:

Daarmee wordt de wiskunde als een geïsoleerd, goed herkenbaar gebied voorgesteld, waarin begrippen absoluut helder zijn en geschikt zijn (gemaakt) voor logisch redeneren. Vermoedelijk wordt deze opvatting over de afgrenzing van het vakgebied gedeeld door een fors deel van de wiskundige beroepsgroep, in ieder geval op het moment dat wiskundigen zich openbaar uitspreken over hun specialisme.

De beweging van de New Math is gestart door wiskundigen van professie en deze beoogde wat nu ook weer voorgesteld wordt: de wiskunde zuiver houden.

Het zijn vaak de wiskundigen geweest die dit geloof in de vormende waarde van het vak naar voren brengen.

De suggestie is dat ^o beroepswiskundigen wereldvreemde individuen zijn die overdreven veel belang aan de wiskunde hechten en deze wiskunde zien als een kant en klaar product dat in het onderwijs moet worden overgedragen, zonder aandacht voor de toepassingen.

Als het echt zo is dat beroepswiskundigen

en onderwijsontwikkelaars fundamenteel verschillende visies op het vak hebben, dan verstaan beiden ook iets anders onder 'goed' wiskundeonderwijs. Dat is echt een groot probleem. Goddijn wil dat beide groepen met elkaar in overleg blijven. 'Blijven' is sterk uitgedrukt gelet op de grote verwijdering die er de laatste decennia heeft plaatsgevonden. Hij roept op tot respect voor elkaars uitgangspunten, maar zijn toevoeging "werkelijke en belezene" geeft aan dat hij een discrepantie ziet tussen mijn denken en handelen. Dat ervaar ik niet als respect.

Literatuur

- Delft, D. van (1998) Pannenkoekwiskunde. In: *Bijlage Wetenschap en Onderwijs, NRC Handelsblad*, 25 april 1998.
- Goddijn, A. (1998) Realiteit van wiskundeonderwijs ter discussie, *Pedagogische studiën* 76, Wolters-Noordhoff, Groningen.
- Keune, F. (1998a) Naar de knoppen. Inaugurale rede, 21 april 1998, Katholieke Universiteit Nijmegen.
(<http://www.sci.kun.nl/math/keune/>)
- Keune, F. (1998b) Naar de knoppen. *Nieuwe Weekrant* 17-4, juni 1998, p. 4-9, Freudenthal instituut, Utrecht.
- Keune, F. (1998c) Ingezonden brief NRC Handelsblad, 2 mei 1998.
(<http://www.sci.kun.nl/math/keune/>)
- Kindt, M. (1998) Beknopt weerwoord. *Nieuwe Weekrant* 17-4, juni 1998, p. 49, Freudenthal instituut, Utrecht.

Aad Goddijn

Freudenthal instituut, Universiteit van Utrecht

In mijn artikel laat ik duidelijk merken van mening te zijn dat de toekomstige VWO-leerling in de onderbouw van nu onvoldoende wordt uitgedaagd. In zijn reactie laat Keune merken dat hij oog heeft voor de nuancering van het begrip abstractie, zoals ik dat voorstel. Hiermee zou op belangrijke punten een goede uitgangssituatie voor verdere discussie kunnen zijn gevonden, al wijst de verdere reactie van Keune niet overduidelijk in die richting. Ik probeer toch nog enkele overeenkomsten en verschillen in visie aan te geven.

1 *Abstractie*. Keune: de wereld van de wiskunde is een bedachte wereld. Dat sluit goed aan bij wat in de wiskundendidactiek tegenwoordig constructivisme wordt genoemd, en ook een belangrijke component is van de realistische visie op wiskundeonderwijs. Binnen de realistische visie wordt wel de consequentie getrokken dat het proces van construeren van begrippen en abstracties in het onderwijs een plaats moet hebben. Juist van de zijde van beroepswiskundigen is vaak zeer positief op dit onderdeel gereageerd.

Keune vervolgt met onder andere: alles moet gebaseerd zijn op harde afspraken over de eigenschappen van deze abstracte zaken. Hier wordt op zijn minst vermeden te zeggen wie de afspraken maakt.

De consequentie van de constructivistische opvatting is zeker niet – dat wordt weleens, zij het hier niet door Keune, gesuggereerd – dat bewijzen en definities nu bij meerderheid van stemmen worden aanvaard of niet. Juist het samen construeren lokt het redeneren en argumenteren uit. Imre Lakatos beschrijft precies dit fenomeen uitermate boeiend in zijn *Proofs and Refutations* (1976).

2 *Mathematiseren of bewijzen*. Ik heb in mijn bijdrage 'mathematiseren' als een zeer ruim scala van activiteiten beschreven. Bewijzen – als activiteit – valt daar uiteraard onder. Ik heb dan ook niet het idee dat 'als je dit

[mathematiseren] zou weglaten, er alleen maar hapklare brokken wiskunde overblijven'. Je houdt misschien wel helemaal niets over.

Een essentie van het bewijzen is dat al bekende (of, zo men wil: axiomatisch gestelde of bewezen) zaken zodanig met elkaar in verband worden gebracht, dat 'hetgeen te bewijzen is' er uit volgt, en zeker: duidelijk wordt. Viel dit niet te lezen uit mijn aanhalingen uit *Revisiting Mathematics Education* van Freudenthal?

In recente experimenten met Vlakke Meetkunde in de top van het VWO valt op dat veel leerlingen (niet allen) gevoelig zijn voor de uitdaging een bewijs te vinden en dan ook het genoeg ontdekken om 'het allemaal prachtig te laten passen', zoals ze het zelf formuleren. Voor leerling, docent en leerplanontwerper zijn dit natuurlijk de mooie momenten.

3 *Niveau*. Hoewel ik het er mee eens ben, dat de toekomstige VWO-leerling in de basisvorming niet voldoende aan zijn/haar trekken komt als we afgaan op de leerboeken van de onderbouw, wil ik de inhoud van het onderbouwprogramma toch niet slechter voorstellen dan het is. Een karikatuur als 'jaren achtereensommen over Pietje en Marietje' is pijnlijk voor wie zich onder moeilijke omstandigheden inzetten voor wiskundeonderwijs aan een brede heterogene groep leerlingen; voor veel docenten is dat de situatie van de dag, en reacties in deze sfeer waren er dan ook in april 1998.

Ook bij de opgave van het experimentele examen van 1992 gaat er iets mis in inschatting van het niveau. Keune ziet niet veel meer dan het maken van een 'grafiekje van een eenvoudige functie' en gaat eraan voorbij dat in de opgave in feite gevraagd werd naar het beschrijven van een benaderingsproces. Weinig algebra, inderdaad; maar zeker wel analyse. Precies dat verschil werd toen door velen in het experiment en de omgeving eromheen herkend en wordt nu door Keune genegeerd. Ik wees in mijn artikel overigens al op reacties van ervaren docenten. De opgave was overigens bedoeld voor het MAVO-D-examen; dat is misschien niet uit de aangegeven W12-16 context opge maakt.

Een andere onjuiste inschatting betreft het niveau van de Vlakke Meetkunde, die in het huidige experiment VWO-B wordt gebruikt.

Leerlingen en docenten in de experimenteerscholen wijzen er krachtig op dat hier toch wel erg veel van ze gevraagd wordt, terwijl de waarde ervan wel wordt erkend. Ik heb de indruk dat Keune met zijn 'vroeger in de eerste klas MULO' aangeeft, dat hij niet op de hoogte is van wat er nú werkelijk gebeurt en hoe het vroeger in de MULO ging. Oudere docenten die dat laatste nog als docent hebben meemaakt, geven in de huidige discussie over de meetkunde onomwonden aan dat het voor de meeste leerlingen van de MULO onbereikbaar was. Ze worden daarin overigens gesteund door publicaties van Van Hiele en Van Hiele-Geldof (1957).

4 *Beroepswiskundigen*. Aan het eind van mijn artikel haal ik Davis en Hersch aan, beroepswiskundigen. Het fragment heeft een licht ironische toon en misschien klinkt er ook wel wat zelfspot in door. Het 'werkelijke' en 'beledene' sloot daar direct opaan. Mijn laatste oproep was gemeend, blijft dat en moet dan ook niet verdraaid worden tot een persoonlijke aanval.

Pedagogische onzekerheid en overbescherming bij ouders van kinderen met hemofilie. Verklarende factoren en gevolgen*

S. Banis, D.R. van Peer, Th.P.B.M. Suurmeijer

Samenvatting

Dit artikel gaat over pedagogische onzekerheid en overbescherming bij ouders van kinderen met hemofilie. Nagegaan wordt (1) wat de (relatieve) invloed is van klinische kenmerken van het kind en emotionele reacties van de ouders op de opvoedingskenmerken en (2) of er verbanden zijn tussen het beschermend gedrag van de ouders en de sociale participatie van het kind. Honderd en acht moeders van een zoon met hemofilie in de leeftijd van 0 tot en met 12 jaar werden geïnterviewd. De emotionele reacties van de moeder bleken meer invloed te hebben op pedagogische onzekerheid en overbeschermend gedrag dan klinische kenmerken van de zoon. Verder was er een positief verband tussen het aantal gedragsbelemmeringen dat de zoon werd opgelegd en negatieve bejegening door leeftijdsgenoten.

1 Inleiding

Tussen 10% en 20% van alle kinderen lijdt aan een chronische aandoening (Midence, 1994). Kenmerkend voor chronische aandoeningen is dat ze lang duren en niet zelden ongeneeslijk zijn. Patiënten moeten vaak levenslang behandeld worden en contacten onderhouden met artsen en andere hulpverleners. De ziekte en de behandeling vormen een belasting voor het chronisch zieke kind en vragen veel van zijn aanpassingsvermogen. De chronische ziekte

van een kind is echter niet alleen van invloed op het functioneren van het kind, maar ook op het functioneren van het gezin waarvan het deel uitmaakt (Bruhn, 1977; Kazak, 1989; Schwenk & Hughes, 1983; Shapiro, 1983).

Met name de manier waarop de ouders het kind opvoeden kan worden beïnvloed door het feit dat het chronisch ziek is. Uit onderzoek blijkt dat pedagogische onzekerheid en overbeschermend gedrag vaker voorkomen bij ouders van chronisch zieke kinderen dan bij ouders van gezonde kinderen (Bos, 1977; Gresnigt & Gresnigt-Strengers, 1973; Gustafsson, Kjellman, Ludvigsson, & Cederblad, 1987; Van Peer & De Vries, 1986; Suurmeijer, 1980). De ouders hebben vragen over de beperkingen en mogelijkheden van hun kind en over de manier waarop ze met hun kind moeten omgaan. Verder kan de fysieke kwetsbaarheid van het kind ertoe leiden dat de ouders hun kind allerlei activiteiten gaan verbieden, om letsel te voorkomen.

Hoewel ouders van chronisch zieke kinderen als groep een verhoogd risico op opvoedingsproblemen hebben, zijn er grote individuele verschillen in de respons van de ouders op de chronische ziekte van het kind (Eiser, 1993; Garrison & McQuiston, 1989; Kazak, 1989; Schwenk & Hughes, 1983; Shapiro, 1983). In traditionele, medische modellen tracht men deze verschillen te verklaren in termen van medische kenmerken, zoals de diagnose of de mate waarin de patiënt aan een bepaalde ziekte lijdt. Subjectieve factoren worden daarin buiten beschouwing gelaten. Meerdere onderzoekers vonden echter dat niet zozeer de objectieve situatie, maar vooral de subjectieve betekenis van die situatie voor de ouders hun reactie op de chronische ziekte van hun kind bepaalde (Athreya & McCormick, 1987; Eiser, 1993; Van Peer & De Vries, 1986; Suurmeijer, 1980).

Verschillende auteurs stellen dat de wijze waarop individuen hun situatie interpreteren

* We zijn zeer erkentelijk voor de waardevolle bijdragen van de leden van de werkgroep 'Hemofilie in Nederland 3' (E. Briët, A. Bröcker-Vriends, H. van Dijk, F.R. Rosendaal, C. Smit en I. Varekamp) en voor de medewerking van de ouders en de hemofiliecentra. De studie werd financieel ondersteund door het Praeventie Fonds, subsidies nr. 28-1099 en nr. 28-1773, en door de Stichting Haemophilia.

van cruciaal belang is voor de manier waarop ze op die situatie reageren (Kelly, 1991; Lazarus, 1966, 1977; Lazarus & Folkman, 1984; Stal & Scheffer, 1982). In de theorie van Lazarus over 'stress and coping' spelen cognitieve processen een belangrijke rol. Deze theorie biedt een verklaring voor het feit dat mensen verschillend reageren op stressoren. Het begrip 'coping' verwijst naar "the constantly changing cognitive and behavioral efforts to manage specific external and/or internal demands that are appraised as taxing or exceeding the resources of the person" (Lazarus & Folkman, 1984, p. 141). Lazarus stelt dat 'an individual's coping behavior is organized not by emotions but by the cognitive process that leads to the emotional response' (Boss, 1987, p. 703). Het individu denkt na over zijn situatie en zijn mogelijkheden aan die situatie het hoofd te bieden. Hij stelt dus vast wat de situatie voor *hem* betekent. De uitkomst van dit cognitieve proces bepaalt zijn gedrag en zijn emotionele respons, die bovendien niet los van elkaar staan: de emotionele respons is van invloed op het gedrag.

De theorie van Lazarus kan worden toegepast op de situatie van ouders van chronisch zieke kinderen. De chronische ziekte van het kind is in deze situatie de stressor waarmee de ouders moeten 'copen'. Lazarus' theorie suggereert dat het gedrag van de ouders jegens hun kind sterk wordt beïnvloed door hun ideeën over de ziekte. De invloed van deze subjectieve ideeën is mogelijk zelfs groter dan de invloed van objectieve factoren. In het onderhavige onderzoek wordt deze hypothese getoetst in een groep ouders van jonge hemofiliepatiënten. Het onderzoek focust op de invloed van klinische kenmerken van het kind (objectieve factoren) en emotionele reacties van de ouders (als expressie van hun interpretatie) op de opvoedingskenmerken. Beschouwde klinische kenmerken zijn de ernst van de hemofilie en de fysieke mobiliteit van de zoon; beschouwde emotionele reacties zijn gevoelens van angst en schuld. Voordat nader wordt ingegaan op deze factoren, wordt een korte beschrijving gegeven van de ziekte en de behandeling.

Hemofilie is een chronische ziekte die wordt gekenmerkt door een verhoogde neiging tot bloedingen wegens een afwijking van het

bloedstollingsmechanisme. De ziekte heeft in Nederland een prevalentie van 8 tot 10 per 100.000 inwoners (Triemstra, 1996). Hemofilie treft vrijwel uitsluitend mannen, omdat het gen dat de recessieve stoornis veroorzaakt op het X-chromosoom ligt. De meeste moeders van patiënten zijn draagster van de ziekte (Rosendaal et al., 1990).

Er worden drie vormen van hemofilie onderscheiden. Bij de ernstige vorm treden spontaan of door een geringe aanleiding bloedingen op in de gewrichten en de spieren. Bij de matig-ernstige vorm treden ook gemakkelijk bloedingen op, maar is er in het algemeen een duidelijk aanwijsbare reden. Bij de lichte vorm van hemofilie vinden alleen bloedingen plaats bij grote verwondingen en operaties en na het trekken van tanden of kiezen. In de Nederlandse patiëntenpopulatie heeft 39% ernstige hemofilie, heeft 18% matig-ernstige hemofilie en heeft 43% lichte hemofilie (Triemstra, 1996).

Bloedingen kunnen zowel intern als extern plaatsvinden, maar de interne bloedingen veroorzaken de grootste lichamelijke problemen. Ze treden meestal op in de gewrichten en de spieren en leiden op korte termijn tot zwellingen, pijn en bewegingsbeperking. Op lange termijn kunnen deze met name bij ernstige vormen van hemofilie steeds weer terugkerende bloedingen schade veroorzaken aan het bewegingsapparaat. De behandeling is daarom gericht op het vermijden en controleren van bloedingen. De patiënt moet zijn activiteiten aanpassen en krijgt thuis of in het ziekenhuis intraveneus stollingsproducten toegediend. Door de moderne behandelingsmethoden is de levensverwachting voor mensen met hemofilie tegenwoordig vrijwel normaal. Het aantal bloedingen bij de matig-ernstige en ernstige vorm van hemofilie is echter nog steeds te groot om gewrichtsschade uit te kunnen sluiten (Triemstra, 1996).

Hoewel er in de literatuur wel aandacht wordt besteed aan de gezinnen rond kinderen met hemofilie (Bruhn, 1977; Jonas, 1977; Mattsson, 1984; Suurmeijer, 1985; Varekamp, Suurmeijer, Rosendaal, Van Dijck, Bröcker-Vriends, & Briët, 1990; Wincott, 1977), konden we slechts één onderzoek traceren dat betrekking had op de opvoedingskenmerken

van ouders van kinderen met hemofilie (Markova, MacDonald, & Forbes, 1980). Markova en haar collega's onderzochten of bepaalde opvoedingskenmerken van de ouders samenhangen met de ernst van de hemofilie van het kind. De steekproef bestond uit de ouders van acht 3 tot 5-jarige jongens met hemofilie en de ouders van acht 8 tot 13-jarige jongens met hemofilie. De jongens stonden geregistreerd bij een ziekenhuis. De onderzoekers vonden dat de ouders het minder eens waren over de opvoeding, minder samenwerkten, minder toekomstplannen maakten voor hun kind en het kind minder verantwoordelijkheid gaven, naarmate de hemofilie ernstiger was. Verder vonden zij dat de moeders onzekerder waren over hoe zij met hun zoon moesten omgaan, naarmate hij een ernstiger vorm van hemofilie had. Deze resultaten wijzen op positieve verbanden tussen enerzijds de ernst van de hemofilie en anderzijds pedagogische onzekerheid en overprotectie.

Volgens Athreya en McCormick (1987) is het aantal beperkingen in het alledaagse functioneren van het kind een belangrijke voorspeller van de invloed van de chronische ziekte op het gezin. In het geval van hemofilie zijn voorkomende beperkingen meestal van fysieke (versus mentale) aard. Er is ons geen onderzoek bekend waarin de relatie tussen de fysieke mobiliteit van het kind met hemofilie en de opvoedingskenmerken van de ouders is onderzocht. We veronderstellen dat ouders pedagogisch onzekerder zijn en meer overbeschermend gedrag vertonen, naarmate hun kind meer fysieke beperkingen heeft.

De ernst van de hemofilie en de fysieke mobiliteit van het kind zijn objectieve factoren die mogelijk van invloed zijn op het opvoedingsgedrag van de ouders. Op basis van de in het voorgaande beschreven theorie van Lazarus veronderstellen we echter, dat de subjectieve betekenis die de ouders verlenen aan de ziekte een grotere invloed heeft op de opvoedingskenmerken. Ouders die de bloedingen als gevaarlijk beschouwen, zullen reageren met angst voor het optreden van bloedingen en voor de gevolgen ervan op langere termijn. Het idee dat de bloedingen gevaarlijk zijn en de daaruit voortvloeiende angst kunnen leiden tot pedagogische onzekerheid en overbeschermend gedrag. Naast angstgevoelens kunnen moeders

schuldgevoelens ervaren, omdat zij de ziekte aan hun kind hebben overgedragen (Jonas, 1977; Markova et al., 1980; Simon, 1984). Simon (1984) suggereert dat moeders hun schuldgevoel compenseren met overprotectief gedrag. Volgens Madden, Terrizzi en Friedman (1982) vormen schuldgevoelens echter maar een klein probleem voor de moeders in vergelijking met hun angst voor de consequenties van bloedingen.

Het opvoedingsgedrag van de ouders beïnvloedt de ontwikkeling van het kind. Zoals hiervoor is beschreven zijn ouders van chronisch zieke kinderen geneigd hun kind meer te beschermen dan ouders van gezonde kinderen. Dit kan onder andere tot uiting komen in het opleggen van gedragsbelemmeringen aan het kind. Dergelijke regels verkleinen de kans op lichamelijk letsel, maar beperken ook de mogelijkheden van het kind deel te nemen aan het sociale verkeer en zich daarin te ontwikkelen (Bruhn, 1977). Uit onderzoek blijkt dat chronisch zieke kinderen een verhoogde kans op sociaal-emotionele problemen hebben (Eiser, 1990, 1993; Garrison & McQuiston, 1989; Pless & Nolan, 1991). Met betrekking tot hemofilie rapporteert Mattsson (1984) dat in de leeftijdscategorie van 7 tot 18 jaar minstens 15% van de kinderen gedragsproblemen heeft, tegenover 7% van de kinderen in de gezonde populatie.

Sociaal-emotionele problemen kunnen zich onder andere uiten in een verminderde sociale participatie, die een verdere belemmering vormt voor de sociaal-emotionele ontwikkeling. Kinderen met epilepsie bleken bijvoorbeeld minder vaak lid te zijn van een club, minder vriendjes te hebben en vaker negatief bejegend te worden dan gezonde kinderen (Suurmeijer, 1980). Behalve de ouders speelt de omgeving buiten het gezin een belangrijke rol. Andere kinderen bijvoorbeeld kunnen moeite hebben zich aan te passen aan de beperkingen van het kind en het daardoor gaan buitensluiten. Wanneer de ouders het kind bovendien allerlei gedragsbelemmeringen opleggen, wordt nog eens benadrukt dat het 'anders dan andere kinderen' is (Suurmeijer, 1985). Dit stempel kan leiden tot een negatieve bejegening door leeftijdsgenootjes. Negatieve bejegening van het kind kan de neiging van de ouders

het kind te beschermen op haar beurt weer versterken.

Onderzoeksdoelen en onderzoeksvragen

Dit onderzoek heeft twee doelen. Het eerste doel is na te gaan welke factoren van invloed zijn op de opvoedingskenmerken (pedagogische onzekerheid en overbescherming) van de ouders van een kind met hemofilie. De centrale vraag daarbij is of de emotionele reacties van de ouders een grotere rol spelen wat betreft hun opvoedingsgedrag, dan klinische kenmerken van het kind. Uit deze algemene vraag zijn vier onderzoeksvragen afgeleid: (1) Hebben klinische kenmerken van het kind invloed op de pedagogische kenmerken van de ouders van een kind met hemofilie? (2) Hebben de emotionele reacties van de ouder invloed op de pedagogische kenmerken van de ouders van een kind met hemofilie? (3) Hebben klinische kenmerken van het kind invloed op de emotionele reacties van de ouders van een kind met hemofilie? (4) Zijn de emotionele reacties van de ouder betere voorspellers van zijn/haar pedagogische kenmerken dan klinische kenmerken van het kind met hemofilie?

Het tweede doel is na te gaan of er een verband is tussen het opvoedingsgedrag van de ouders en het gedrag van het kind met hemofilie. Meer toegespitst richten we ons op een mogelijk verband tussen het beschermend gedrag van de ouders en de sociale participatie van het kind: (5) Is er een verband tussen het beschermend gedrag van de ouders en de sociale participatie van het kind met hemofilie?

2 Methode

De onderzoeksvragen werden beantwoord door middel van een secundaire data-analyse. Het gebruikte gegevensbestand is afkomstig van een onderzoek bij gezinnen met een kind met hemofilie (Hutten et al., 1988; Varekamp et al., 1990). Doel van dat onderzoek was na te gaan wat de effecten van hemofilie op het kind en het gezin waren. In het hier beschreven onderzoek werden gegevens geanalyseerd die niet eerder waren gebruikt.

2.1 De onderzoeksgroep

In 1987 werden 126 moeders/vaders van kinde-

ren met hemofilie in de leeftijd van 0 tot en met 12 jaar gevraagd te participeren in het onderzoek. Hun adressen waren voor een groot deel afkomstig van een groot adressenbestand dat was gebruikt voor een survey onder Nederlandse hemofiliepatiënten in 1985. Verder was aan grote hemofiliecentra een lijst gevraagd van patiënten geboren na 1983. Honderd en elf ouders namen deel, de respons was derhalve 88% (Varekamp et al., 1990). De gegevens die in het onderhavige onderzoek werden geanalyseerd, zijn afkomstig van 108 moeders van een zoon met hemofilie.

Tabel 1 geeft een overzicht van algemene kenmerken van de moeders in de steekproef en hun gezinnen. Om een indicatie te krijgen van de sociaal-economische status van het gezin werd gevraagd naar het beroep van de vader. Als de moeder alleenstaand was, werd gevraagd naar het beroep van de moeder. De indeling naar beroepsniveau werd gemaakt aan de hand van de Beroepenklapper (Kropman, 1975).

In 12 gezinnen waren twee of meer kinderen met hemofilie. Tijdens de interviews met de betreffende moeders werd gesproken over hun ervaringen met het jongste kind met hemofilie. Tabel 2 geeft een overzicht van de gemiddelde leeftijd, de plaats in de kinderrij en enkele medische gegevens van de zoons.

De frequentieverdeling van de ernst van de

Tabel 1

Algemene kenmerken van de moeders en de gezinnen

Leeftijd moeders		$M = 34.4$ ($SD = 4.8$)
Gezins situatie (%)		
Eén-ouder-gezin		6.5
Twee-ouder-gezin		93.5
Gezinsgrootte (%)		
Eén kind		26.9
Twee of meer kinderen		73.1
Aantal kinderen		$M = 2.0$ ($SD = 0.8$)
Beroepsniveau (%)		
Ongeschoolde arbeid		7.4
Geschoolde arbeid		24.1
Lagere bediende		15.7
Kleine zelfstandige		9.3
Middelbare bediende		23.3
Hogere beroepen		20.4

Noot. $N = 108$

Tabel 2
Leeftijd, plaats in de kinderrij en enige medische kenmerken van de zoons

Leeftijd	<i>M</i> = 6.9 (<i>SD</i> = 3.5)
Plaats in de kinderrij (%)	
1	52.8
2	32.4
3	14.8
Type hemofilie (%)	
Hemofilie A	85.2
Hemofilie B	14.8
Ernst van de hemofilie (%)	
Licht	22.2
Matig-ernstig	25.9
Ernstig	51.9
Behandeling (%)	
Geen transfusiebehandeling	10.2
Profylactische behandeling	38.0
Transfusiebehandeling alleen bij bloedingen	51.8
Thuisbehandeling	37.4
HIV-test (%)	
Getest	29.6
Seropositief	3.7
Uitslag onbekend	5.6
Seronegatief	20.4

Noot. *N* = 108

hemofilie in de steekproef week significant af van de frequentieverdeling in de Nederlandse patiëntenpopulatie ($\chi^2 = 13.53$, *df* = 2, *N* = 108, *p* = .001).

Verder waren vier zoons seropositief. Het is niet onwaarschijnlijk dat de seropositiviteit van een kind met hemofilie van invloed is op de emotionele reacties en pedagogische kenmerken van de moeder. Herhaling van de uitgevoerde analyses zonder de seropositieve cases leverde echter geen significant andere resultaten op.

2.2 De vragenlijst

De vragenlijst werd voor een groot deel afgeleid van bestaande lijsten, ontwikkeld voor onderzoek onder gezinnen met een gehandicapt kind of een kind met een chronische ziekte (Bos, 1977; Gresnigt & Gresnigt-Strengers, 1973; Suurmeijer, 1980). De onderzoekers pasten sommige items aan de onderzoeksgroep aan en voegden nieuwe items toe (Varekamp et al., 1990). Hier wordt alleen dat deel van de vragenlijst besproken, dat relevant is voor het

onderhavige onderzoek. Tabel 3 geeft een overzicht van de schalen, de index en de items die de onderzoeksvariabelen operationaliseren.

De items van de schaal voor fysieke mobiliteit hebben vier antwoordcategorieën: 1 = kan het niet, 2 = veel moeite, 3 = enige moeite, 4 = geen moeite. Alleen de moeders van een zoon met hemofilie van 3 jaar of ouder (*n* = 93) moesten deze items beantwoorden.

De items van de schalen voor pedagogische onzekerheid, overprotectie, angst voor bloedingen en angst voor invaliditeit, en het item voor schuldgevoel hebben vijf antwoordcategorieën, lopend van *helemaal mee eens* tot *helemaal mee oneens*.

Naast de schaal voor overprotectie werd een index gebruikt om overbeschermend gedrag te meten. De belemmeringsindex is gebaseerd op 11 items over door de ouders aan hun kind opgelegde gedragsbelemmeringen. Negen items hebben twee antwoordcategorieën: 0 = nee, 1 = ja; twee items hebben daarnaast een extra categorie: 2 = niet van toepassing. Alleen de moeders van een zoon met hemofilie van 4 jaar of ouder (*n* = 71) moesten deze 11 items beantwoorden.

De sociale participatie van de zoon werd geoperationaliseerd door middel van een schaal en twee losse items. De items van de schaal voor negatieve bejegening hebben drie antwoordcategorieën: 0 = vrijwel nooit, 1 = soms, 2 = vrij vaak. Het antwoord op het item over het aantal vriendjes wordt ingedeeld in één van de volgende drie categorieën: 0 = geen vriendjes, 1 = 1 of 2 vriendjes, 2 = 3 of meer vriendjes. Het item over het lidmaatschap van een vereniging heeft twee antwoordcategorieën: 1 = nee, 2 = ja. De zes items betreffende de sociale participatie moesten alleen door de moeders van een zoon met hemofilie van 4 jaar of ouder worden beantwoord.

Tabel 3 geeft voor elke schaal de waarde van Cronbach's alpha in dit onderzoek. Alleen de schalen voor pedagogische onzekerheid en negatieve bejegening werden eerder gebruikt in de huidige samenstelling (Suurmeijer, 1980). Er is echter geen informatie beschikbaar over de betrouwbaarheid van de schalen in dat onderzoek.

Cronbach's alpha is een schatting van de ondergrens van de betrouwbaarheid (Drenth & Sijsma, 1990). De vereiste hoogte van alpha is

Tabel 3

Schalen, index en items: aantal items, bereik, betrouwbaarheidscoëfficiënt Cronbach's α in dit onderzoek, auteur(s) van item(s) en itemvoorbeeld

Variabele	Items	Bereik	Alpha	Auteurs	Itemvoorbeeld
Klinische kenmerken					
Ernst van de hemofilie	1	1-3	-	V	Wat is de ernst van de hemofilie?
Fysieke mobiliteit	3	3-12	.71	V	Heeft uw kind moeite met 400 meter lopen?
Emotionele reacties					
Angst voor bloedingen	5	5-25	.84	S*, V	Ik ben vrij vaak bang dat ons kind valt.
Angst voor invaliditeit	2	2-10	.68	V	Ik maak me vrij vaak zorgen dat ons kind moeite met lopen zal krijgen.
Schuldgevoel	1	1-5	-	V	Ik denk wel eens dat hij zijn aandoening een beetje aan mij te wijten heeft.
Opvoedingskenmerken					
Pedagogische onzekerheid	6	6-30	.68	GG, S	Ik vind het moeilijk dit kind op de juiste wijze aan te pakken.
Overprotectie	2	2-10	.65	GG, S	Je laat dit kind niet zo gauw los als een ander kind.
Belemmeringsindex	11	0-11	-	S	Mag dit kind met gymnastiek meedoen?
Sociale participatie					
Negatieve bejegening	4	0-8	.69	S	Plagen de kinderen in de buurt of op school uw kind nogal eens?
Aantal vriendjes	1	0-2	-	S	Als u nagaat met wie uw kind zoal speelt en veel omgaat hier in de buurt en op school en elders, hoeveel echte vriendjes of vriendinnetjes heeft hij dan?
Lidmaatschap vereniging	1	1-2	-	S	Is uw kind actief lid van een sportvereniging of van een hobbyvereniging?

Noot. Hoe hoger de score, des te meer het kenmerk aanwezig is.

V = Varekamp et al. (1989); S = Suurmeijer (1980); GG = Gresnigt & Gresnigt-Strengers (1973).

* formulering items aangepast.

afhankelijk van de toepassing van het instrument. Volgens Nunnally (1978) zijn alpha's van .70 of hoger voldoende voor onderzoek bij groepen met nieuwe schalen. In een eerdere publicatie (Nunnally, 1967) stelde hij echter dat voor dit doel alpha's tussen .50 en .60 voldoende zijn. Uit Tabel 3 blijkt dat in ons onderzoek alpha varieert van .65 tot .84. Alle schalen voldoen aan het tweede criterium, terwijl twee schalen ook voldoen aan het eerste criterium. Verder zijn alle item-restcorrelaties positief en voldoende hoog (.22 - .81). Dit betekent dat de schalen intern consistent zijn.

3 Resultaten

Tabel 4 geeft een overzicht van de gemiddelden en de standaarddeviaties of de frequentie-

verdelingen van de variabelen in de steekproef.

Om de onderzoeksvragen 1, 2 en 3 te kunnen beantwoorden, werden Pearson correlaties berekend tussen de betreffende variabelen. Voor de beantwoording van onderzoeksvraag 4 maakten we gebruik van hiërarchische multiple regressie analyses. Voor onderzoeksvraag 5 werden Spearman correlaties berekend. Omdat het bereik van de leeftijd van de zoons tamelijk groot was (0 tot en met 12 jaar), werd leeftijd als variabele opgenomen in de analyses. Bij de toetsingen werd steeds een significantieniveau van 5% gehanteerd. Het aantal cases waarop berekeningen konden worden gebaseerd, varieerde per toetsing, maar was steeds voldoende om relevante effecten te kunnen detecteren. De resultaten van de analyses staan in de Tabellen 5, 6 en 7.

De resultaten worden per onderzoeksvraag

Tabel 4
Gemiddelden en standaarddeviaties of frequentieverdelingen van de variabelen in de steekproef

Variabele	M	SD	f (%)
Ernst van de hemofilie			
licht			22.2
matig-ernstig			25.9
ernstig			51.9
Fysieke mobiliteit	11.6	.7	
Angst voor bloedingen	13.6	5.7	
Angst voor invaliditeit	5.3	2.5	
Schuldgevoel	2.1	1.5	
Pedagogische onzekerheid	15.4	5.3	
Overprotectie	5.7	2.6	
Belemmeringsindex	3.8	1.1	
Negatieve bejegening	1.2	1.5	
Aantal vriendjes			
geen vriendjes			6.8
1 of 2 vriendjes			33.8
3 of meer vriendjes			59.5
Lidmaatschap vereniging			
nee			41.0
ja			59.0

besproken. Vanwege de samenstelling van de onderzoeksgroep spreken we in het volgende steeds van moeders en zoons in plaats van ouders en kinderen.

Klinische kenmerken en pedagogische kenmerken

Moeders beschermden hun zoon meer (schaal voor overprotectie), naarmate hij een ernstiger vorm van hemofilie had. Er waren geen andere significante correlaties tussen klinische kenmerken en pedagogische kenmerken.

De belemmeringsindex correleerde niet met de ernst van de hemofilie, terwijl dit instrument, net als de schaal voor overprotectie, werd gebruikt als operationalisatie van overbescherming. Blijkbaar meten de instrumenten verschillende kenmerken. Dit idee wordt bevestigd door de matige correlatie tussen beide ($r = .27$).

Emotionele reacties en pedagogische kenmerken

De pedagogische onzekerheid van de moeder was groter, naarmate haar angst voor bloedingen, haar angst voor invaliditeit en/of haar schuldgevoel sterker was. Ze beschermde haar zoon meer (schaal voor overprotectie), naarmate haar angst voor bloedingen en/of haar angst voor invaliditeit sterker was. Het aantal gedragsbelemmeringen dat de moeder haar

Tabel 5
Pearson correlaties tussen klinische kenmerken, emotionele reacties, pedagogische kenmerken en leeftijd zoon

	Ernst hemofilie	Fysieke mobiliteit	Angst voor bloedingen	Angst voor invaliditeit	Schuldgevoel	Pedagogische onzekerheid	Overprotectie	Belemmeringsindex
Ernst hemofilie	-.09							
Fysieke mobiliteit	.05	-.25**						
Angst voor bloedingen	.11	-.40***	.47***					
Angst voor invaliditeit	.04	-.03	.16	.17*				
Schuldgevoel	.06	-.16	.40***	.35***	.23**			
Pedagogische onzekerheid	.23**	-.14	.45***	.29**	.11	.49***		
Overprotectie	.06	.06	.27*	.04	-.03	.13	.27*	
Belemmeringsindex	-.07	-.10	-.28**	.02	-.12	-.08	-.06	
Leeftijd zoon								-.45***

Noot. $n_{\max} = 108$, $n_{\min} = 71$.

* $p < .05$. ** $p < .01$. *** $p < .001$ (eenzijdig).

zoon oplegde, was afhankelijk van haar angst voor bloedingen.

Zowel de belemmeringsindex als de angst voor bloedingen correleerden negatief met de leeftijd van de zoon (zie Tabel 5). Om de unieke invloed van angst op het aantal opgelegde gedragsbelemmeringen te bepalen, berekenden we een partiële correlatie. De partiële correlatie tussen de belemmeringsindex en de angst voor bloedingen, gecontroleerd voor leeftijd, week niet af van de ongecontroleerde correlatie.

Klinische kenmerken en emotionele reacties

De angst voor bloedingen en de angst voor invaliditeit van de moeder waren beide groter, naarmate de fysieke mobiliteit van haar zoon slechter was. Er was geen significante correlatie tussen schuldgevoel en fysieke mobiliteit, noch tussen één van de emotionele reacties en de ernst van de hemofilie.

Emotionele reacties versus klinische kenmerken

Uit Tabel 5 blijkt dat pedagogische onzeker-

heid, overprotectie en belemmeringsindex in het algemeen sterker correleerden met de emotionele reacties van de moeder dan met de klinische kenmerken van de zoon. Beide soorten angstgevoelens correleerden echter met de fysieke mobiliteit. Om te bepalen welke groep variabelen, klinische kenmerken of emotionele reacties, de meeste variantie verklaarde, gebruikten we hiërarchische regressie analyses. Drie analyses werden uitgevoerd met pedagogische onzekerheid, overprotectie en belemmeringsindex als opeenvolgende afhankelijke variabelen. Om de invloed van de leeftijd van de zoon uit te sluiten werd deze variabele in de analyses opgenomen als controlevariabele. In drie achtereenvolgende stappen werden leeftijd, klinische kenmerken en emotionele reacties in de analyse gebracht. De resultaten staan in Tabel 6.

Overeenkomstig onze verwachtingen droegen emotionele reacties in belangrijke mate bij aan pedagogische onzekerheid. Ze verhoogden de hoeveelheid verklaarde variantie met 18%,

Tabel 6

Hiërarchische multiple regressie analyses van pedagogische kenmerken op leeftijd zoon, klinische kenmerken en emotionele reacties

Pedagogische onzekerheid (n = 93) ¹		Cum R ²	F	Change R ²	F
Stap	Variabelen				
1	Leeftijd	.01	.91		
2	Ernst hemofilie				
	Fysieke mobiliteit	.04	1.30	.03	1.49
3	Angst voor bloedingen				
	Angst voor invaliditeit				
	Schuldgevoel	.22	4.06**	.18	6.58***
Overprotectie (n = 92) ¹		Cum R ²	F	Change R ²	F
Stap	Variabelen				
1	Leeftijd	.00	.13		
2	Ernst hemofilie				
	Fysieke mobiliteit	.07	2.27	.07	3.34*
3	Angst voor bloedingen				
	Angst voor invaliditeit				
	Schuldgevoel	.27	5.29***	.20*	7.78***
Belemmeringsindex (n = 71) ²		Cum R ²	F	Change R ²	F
Stap	Variabelen				
1	Leeftijd	.20	17.14***		
2	Ernst hemofilie				
	Fysieke mobiliteit	.22	6.21***	.02	.80
3	Angst voor bloedingen				
	Angst voor invaliditeit				
	Schuldgevoel	.28	4.06**	.06	1.71

¹ leeftijd zoons 3-12; ² leeftijd zoons 4-12. * p < .05. ** p < .01. *** p < .001.

Tabel 7

Spearman correlaties tussen enerzijds overprotectie en belemmeringsindex en anderzijds de sociale participatie van de zoon

	Controlevariabele		
	Leeftijd	Ernst	
<i>Overprotectie</i>			
Negatieve bejegening	.18	.18	.18
Aantal vriendjes	-.22*	-.21*	-.19
Lidmaatschap vereniging	.02	.01	.05
<i>Belemmeringsindex</i>			
Negatieve bejegening	.09	.22*	.09
Aantal vriendjes	-.07	-.16	-.06
Lidmaatschap vereniging	-.08	.10	-.07

Noot. $n_{\max} = 77$, $n_{\min} = 67$.

* $p < .05$ (eenzijdig).

na controle voor de leeftijd van de zoon en klinische kenmerken.

In de voorspelling van overprotectie namen emotionele reacties 20% van de te verklaren variantie voor hun rekening. De leeftijd van de zoon verklaarde geen variantie. Klinische kenmerken verklaarden 7% van de variantie.

In tegenstelling tot pedagogische onzekerheid en overprotectie kon het aantal opgelegde gedragsbelemmeringen niet worden verklaard door emotionele reacties. De leeftijd van de zoon verklaarde 20% van de variantie. Klinische kenmerken noch emotionele reacties droegen significant bij aan de hoeveelheid verklaarde variantie.

Het beschermend gedrag van de moeder en de sociale participatie van de zoon

Er was een significante, negatieve correlatie tussen overprotectie en het aantal vriendjes (zie Tabel 7). De partiële correlatie, gecontroleerd voor de ernst van de hemofilie, was echter niet significant. Verder was er een significante, positieve correlatie tussen de belemmeringsindex en de mate waarin de zoon negatief werd bejegend, als er werd gecontroleerd voor de leeftijd van de zoon.

We hebben gekozen voor berekening van partiële correlaties met als controlevariabelen de leeftijd van de zoon en de ernst van de hemofilie, omdat deze twee variabelen beide correleerden of verondersteld werden samen te hangen met bepaalde aspecten van het functioneren van zowel moeder als zoon (zie ook Tabel 5). Omdat de fysieke mobiliteit van de

zoon noch met overprotectie noch met de belemmeringsindex bleek te correleren, werden geen partiële correlaties berekend met deze variabele als controlevariabele.

4 Discussie

Dit onderzoek had een tweeledig doel. Het eerste doel was na te gaan welke factoren van invloed zijn op pedagogische onzekerheid en overbescherming bij ouders van een kind met hemofilie. Het tweede doel was na te gaan of er een verband is tussen het opvoedingsgedrag van de ouders en het gedrag van het kind met hemofilie. Bij beide doelen werd een hoofdvraag geformuleerd. Om de vragen te kunnen beantwoorden werden gegevens geanalyseerd over jongens met hemofilie en hun moeders.

Bij het eerste doel was de centrale vraag of de emotionele reacties van de ouders meer invloed hebben op hun opvoedingskenmerken dan klinische kenmerken van het kind. Uit de resultaten kon worden geconcludeerd dat emotionele reacties van de moeder betere predictoren zijn van haar pedagogische kenmerken dan klinische kenmerken van de zoon. Deze conclusie is niet van toepassing op de belemmeringsindex. Hoewel het aantal gedragsbelemmeringen dat aan de zoon wordt opgelegd, was gerelateerd aan de angst voor bloedingen, was de leeftijd van de zoon een veel belangrijker voorspeller.

De klinische kenmerken bleken slechts een beperkt effect te hebben op het opvoedingsgedrag van de moeders van een zoon met hemofilie. De moeders beschermen hun zoon meer, naarmate hij lijdt aan een ernstiger vorm van hemofilie. Dit resultaat is in overeenstemming met eerder onderzoek waaruit bleek dat moeders hun zoon minder verantwoordelijkheid geven, naarmate de hemofilie ernstiger is (Markova et al. 1980). Omdat de fysieke mobiliteit van invloed is op de angstgevoelens van de moeder, heeft ze een indirect effect op de opvoedingspraktijken. Maar dit effect is waarschijnlijk erg beperkt, gezien de afwezigheid van significante correlaties tussen mobiliteit en opvoedingskenmerken. Bovendien laten de regressieresultaten zien, dat de unieke bijdrage van de emotionele reacties van de moeders aan pedagogische onzekerheid en overprotectie

veel groter is dan de bijdrage van klinische kenmerken, inclusief de fysieke mobiliteit.

Overeenkomstig eerder onderzoek (Athreya & McCormick, 1987; Eiser, 1993; Van Peer & De Vries, 1986; Suurmeijer, 1980), suggereren onze resultaten dat ouderlijke interpretaties hun reactie op de chronische ziekte, i.c. hun opvoedingsgedrag, bepalen. Individuele emotionele reacties op een situatie worden immers bepaald door een voorafgaand proces van betekenisverlening aan die situatie (Lazarus & Folkman, 1984). In dit proces spelen objectieve factoren een beperkte rol. De afwezigheid van significante correlaties tussen gevoelens van angst en de ernst van de hemofilie in onze steekproef is opmerkelijk. In het algemeen gaan ernstiger vormen van hemofilie gepaard met een hogere bloedingsfrequentie dan lichtere vormen van hemofilie. Men zou daarom verwachten dat de angstgevoelens sterker zouden zijn, naarmate de hemofilie ernstiger was.

Bij het tweede doel spitsten we ons toe op de vraag of er een verband is tussen het beschermend gedrag van de ouders en de sociale participatie van het kind. Uit de resultaten bleek dat er een zwak verband is tussen het beschermend gedrag van de moeder en de sociale participatie van de zoon met hemofilie.

Het lijkt er op dat zoons die relatief meer worden beschermd door hun moeder, op het sociale vlak minder goed functioneren. De correlaties in onze steekproef tussen enerzijds overprotectie en anderzijds negatieve bejegening, respectievelijk het aantal vriendjes, weken net niet significant af van nul. De correlatie tussen het aantal gedragsbelemmeringen dat aan de zoon werd opgelegd en de mate waarin de zoon negatief werd bejegend door andere kinderen, gecontroleerd voor leeftijd, was wel significant. Suurmeijer (1980) deed een soortgelijke bevinding in zijn onderzoek bij kinderen met epilepsie en hun gezinnen. De negatievere bejegening van kinderen met epilepsie in vergelijking met gezonde kinderen ging gepaard met een groter aantal gedragsbelemmeringen.

Het cross-sectionele karakter van dit onderzoek staat eigenlijk geen causale interpretatie van gevonden verbanden toe. Toch veronderstellen wij, evenals Suurmeijer (1980), dat de ouders een bepalende rol spelen. Naarmate de zoon minder 'mag' van zijn ouders, is het voor

zijn leeftijdsgenootjes moeilijker zich aan hem aan te passen en wordt meer de indruk gevestigd dat hij 'anders' is. Dit werkt sociale isolatie van het kind in de hand. Het kind krijgt onvoldoende gelegenheid ervaringen op te doen en daarvan te leren, waardoor zijn sociaal-emotionele ontwikkeling wordt belemmerd.

Methodologische overwegingen

In de onderzoeksgroep waren relatief meer kinderen met een matig-ernstige of ernstige vorm van hemofilie dan in de Nederlandse populatie hemofiliepatiënten. Dit is te wijten aan de gevolgde selectieprocedure. Met name in de populatie kinderen met hemofilie die een hemofiliecentrum bezoekt, is de lichte vorm van hemofilie ondervertegenwoordigd. Hoewel we de herkomst van de afzonderlijke adressen niet konden achterhalen, vermoeden we dat de ernst-bias wordt veroorzaakt door de subgroep kinderen in de steekproef die geregistreerd staat bij een hemofiliecentrum. De ernst-bias is mogelijk van invloed geweest op de omvang van besproken problemen, maar niet noodzakelijk op de verbanden tussen variabelen.

De bevinding dat klinische kenmerken nauwelijks van invloed zijn op opvoedingskenmerken, impliceert niet dat klinische kenmerken onbelangrijk zijn. De aanwezigheid van hemofilie op zich kan opvoedingspraktijken beïnvloeden. Onderzoek waarin ook een controlegroep is opgenomen, zou deze effecten aan het licht kunnen brengen.

In dit onderzoek stonden de opvoedingskenmerken van *moeders* van een zoon met hemofilie centraal. Uit onderzoek blijkt dat vaders van chronisch zieke kinderen de situatie vaak anders interpreteren en anders reageren dan moeders (Eiser, 1993; Van Peer & De Vries, 1986; Suurmeijer, 1980, 1995). Opname van zowel moeders als vaders in de onderzoeksgroep zou het mogelijk maken cognities, emotionele reacties en gedrag van beide ouders met elkaar te vergelijken. Ook zou dan voor moeders en vaders afzonderlijk kunnen worden onderzocht door welke factoren hun gedrag wordt beïnvloed.

Aanbevelingen voor de hulpverlening

De betekenis die ouders van een kind met

hemofilie verlenen aan de ziekte is van groot belang, omdat deze hun gedrag jegens het kind beïnvloedt. Onze bevindingen suggereren dat de interpretatie van de ouders slechts gedeeltelijk gebaseerd is op objectieve factoren. Hulpverleners die het opvoedingsgedrag van deze ouders willen ondersteunen of verbeteren, moeten dan ook aandacht schenken aan het proces van betekenisverlening bij ouders, in plaats van te focussen op medische factoren. Ouders moeten goed worden voorgelicht over de ziekte en de behandeling, zodat zij zich een reëel beeld kunnen vormen van de situatie. Bovendien moeten zij worden ondersteund bij het zoeken naar oplossingen voor (opvoedings)problemen, zodat zij het hoofd kunnen bieden aan de situatie.

Het belang van goede informatie wordt bevestigd door onderzoek bij ouders van kinderen met epilepsie (Suurmeijer, 1985, 1995). De ouders bleken veel waarde te hechten aan informatie over zowel medische als psychosociale aspecten van de ziekte. Volgens de ouders kwam het de relatie met hun kind ten goede goed geïnformeerd te worden. Verder bleek dat de ouders die positief oordeelden over het informatieve gedrag van de behandelende arts, hun kind één tot drie (van de tien onderzochte) gedragsbelemmeringen minder oplegden dan de ouders die niet positief oordeelden.

Literatuurlijst

Athreya, B.H., & McCormick, M.C. (1987). Impact of chronic illness on families. *Rheumatic Disease Clinics of North America*, 13, 123-131.

Bos, J.M. (1977). *Het gezin met een hartekind: Een sociaal-wetenschappelijk onderzoek in gezinnen met een kind met een aangeboren hartafwijking*. Proefschrift, Amsterdam: Swets & Zeitlinger.

Boss, P. (1987). Family stress. In M.B. Sussman & S.K. Steinmetz (Eds.), *Handbook of marriage and the family* (pp. 695-723). New York: Plenum Press.

Bruhn, J.G. (1977). Effects of chronic illness on the family. *The Journal of Family Practice*, 4, 1057-1060.

Drenth, P.J.D., & Sijtsma, K. (1990). *Testtheorie: Inleiding in de theorie van de psychologische test en zijn toepassingen*. Houten: Van Loghum Slaterus.

Eiser, C. (1990). Psychological effects of chronic disease. *Journal of Child Psychology and Psychiatry*, 31, 85-98.

Eiser, C. (1993). *Growing up with a chronic disease: The impact on children and their families*. London: Kingsley.

Garrison, W.T., & McQuiston, S. (1989). *Chronic illness during childhood and adolescence: Psychological aspects*. Newbury Park, CA: Sage Publications.

Gresnigt, H.A.A., & Gresnigt-Strengers, A.M.C. (1973). *Ouders en gezinnen met een diepzwakzinnig kind*. Proefschrift, Amsterdam: Swets & Zeitlinger.

Gustafsson, P.A., Kjellman, N.-I.M., Ludvigsson, J., & Cederblad, M. (1987). Asthma and family interaction. *Archives of Disease in Childhood*, 62, 258-263.

Hutten, J.B.F., Varekamp, I., Rosendaal, F.R., Smit, C., Dijck, H. van, Bröcker-Vriends, A.H.J.T., Briët, E., & Suurmeijer, Th.P.B.M. (1988). *Hemofilie in Nederland 3. Kinderen met hemofilie: Het functioneren van kind, ouders en gezin*. Groningen: RUG/RUL/NVHP.

Jonas, D.L. (1977). Psychiatric aspects of hemophilia. *The Mount Sinai Journal of Medicine*, 44, 457-463.

Kazak, A.E. (1989). Families of chronically ill children: A systems and social-ecological model of adaptation and challenge. *Journal of Consulting and Clinical Psychology*, 57, 25-30.

Kelly, G. (1991). *The psychology of personal constructs*. London: Routledge.

Kropman, L. (1975). *Beroepenklapper*. Niet gepubliceerd, Instituut voor Toegepaste Sociologie, Nijmegen.

Lazarus, R.S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.

Lazarus, R.S. (1977). Cognitive and coping processes in emotion. In A. Monat & R.S. Lazarus (Eds.), *Stress and coping*. New York: Columbia University Press.

Lazarus, R.S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Madden, N.A., Terrizzi, J., & Friedman, S.B. (1982). Psychological issues in mothers of children with hemophilia. *Developmental and Behavioral Pediatrics*, 3, 136-142.

Markova, I., MacDonald, K., & Forbes, C. (1980). Impact of haemophilia on child-rearing practices and parental cooperation. *Journal of Child Psychology and Psychiatry*, 21, 153-162.

- Mattsson, A. (1984). Hemophilia and the family: Life-long challenges and adaptation. *Scandinavian Journal of Haematology*, 33(Suppl. 40), 65-74.
- Midence, K. (1994). The effects of chronic illness on children and their families: An overview. *Genetic, Social, and General Psychology Monographs*, 120, 309-326.
- Nunnally, J.C. (1967). *Psychometric theory*. New York: McGraw-Hill.
- Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Peer, D.R. van, & Vries, A.F. de (1986). De invloed van een kind met een chronische ziekte (CARA) op de opvoedingsattituden van de ouders. *Tijdschrift voor Sociale Gezondheidszorg*, 64, 309-312.
- Pless, I.B., & Nolan, T. (1991). Revision, replication and neglect: Research on maladjustment in chronic illness. *Journal of Child Psychology and Psychiatry*, 32, 347-365.
- Rosendaal, F.R., Bröcker-Vriends, A.H.J.T., Houwelingen, J.C. van, Smit, C., Varekamp, I., Dijk, H. van, Suurmeijer, Th.P.B.M., Vandenbroucke, J.P., & Briët, E. (1990). Sex ratio of the mutation frequencies in haemophilia A: Estimation and meta-analyses. *Human Genetics*, 86, 139-146.
- Schwenk, T.L., & Hughes, C.C. (1983). The family as patient in family medicine: Rhetoric or reality? *Social Science and Medicine*, 17, 1-16.
- Shapiro, J. (1983). Family reactions and coping strategies in response to the physically ill or handicapped child: A review. *Social Science and Medicine*, 17, 913-931.
- Simon, S. (1984). Hemophilia and the family system. *Psychosomatics*, 25, 845-849.
- Stal, P., & Scheffer, W. (1982). *Omgaan met problemen: Theorie en empirie over reaktiewijzen van mensen op psychosociaal belastende omstandigheden; perspectieven voor onderzoek*. Instituut voor Toegepaste Sociologie, Nijmegen.
- Suurmeijer, Th.P.B.M. (1980). *Kinderen met epilepsie: Een onderzoek naar de invloed van een ziekte op kind en gezin*. Proefschrift, Groningen: Veenstra Visser.
- Suurmeijer, Th.P.B.M. (1985). Ouders van kinderen met een chronische aandoening *Gezondheid & Samenleving*, 6, 113-122.
- Suurmeijer, Th.P.B.M. (1995). The impact of epilepsy on social integration and 'quality of life': Family, peers, and education. In A.P. Aldenkamp, F.E. Dreifus, W.O. Renier & Th.P.B.M. Suurmeijer (Eds.), *Epilepsy in children and adolescents* (pp. 251-270). Boca Raton: CRC Press.
- Triemstra, A.H.M. (1996). *Medical and psychosocial aspects of haemophilia*. Proefschrift, Enschede: CopyPrint 2000.
- Varekamp, I., Smit, C., Rosendaal, F.R., Bröcker-Vriends, A., Briët, E., Dijk, H. van, & Suurmeijer, Th.P.B.M. ((1989). Employment of individuals with haemophilia in The Netherlands. *Social Science and Medicine*, 28, 261-270.
- Varekamp, I., Suurmeijer, Th.P.B.M., Rosendaal, F.R., Dijk, H. van, Bröcker-Vriends, A., & Briët, E. (1990). Family burden in families with a hemophilic child. *Family Systems Medicine*, 8, 291-301.
- Wincott, E.A.I.M.S.W. (1977). Psychosocial aspects of hemophilia: Problems, prevention, treatment modalities, research, and future directions. *The Mount Sinai Journal of Medicine*, 44, 438-455.

Auteurs

S. Banis studeerde orthopedagogiek aan de Rijksuniversiteit Groningen.

D.R. van Peer is als universitair docent orthopedagogiek verbonden aan de Basiseenheid Orthopedagogiek van de Rijksuniversiteit Groningen.

Th.P.B.M. Suurmeijer is als universitair hoofd-docent medische sociologie verbonden aan de Basiseenheid Sociologie, het Noordelijk Centrum voor Gezondheidsvraagstukken (NCG) en het Interuniversitair Centrum voor Sociaalwetenschappelijke theorievorming en methodenontwikkeling (ICS), van de Rijksuniversiteit Groningen.

Adres: D.R. van Peer, Rijksuniversiteit Groningen, Basiseenheid Orthopedagogiek, Grote Rozenstraat 38, 9712 TJ Groningen. Tel.: 050-3636592

Abstract

Child-rearing uncertainty and overprotection among parents of children with haemophilia: Explanatory factors and consequences

This study addresses (1) the relative importance of clinical characteristics of the child and parental emotional reactions, to child-rearing practices towards children with haemophilia, and (2) possible relationships between overprotective behaviour of the parents and social participation features of the child. Hundred and eight mothers of 0- to-12-year old boys with haemophilia were interviewed. Mother's emotional reactions appeared to have a stronger influence on child-rearing uncertainty and overprotective behaviour than clinical characteristics of the son. Furthermore, a positive association was found between the number of behavioural restrictions imposed on the son and negative treatment by peers.

Wiskunde- en taalvaardigheid als voorspeller van B-keuzen in het voortgezet onderwijs

D. Uerz, H. Dekkers & J. Dronkers

Samenvatting

Een groeiend tekort aan technisch en exact opgeleide afgestudeerden houdt de vraag naar determinanten van verschillen in prestaties en keuzen op dat terrein naar onder andere maatschappelijke achtergrondkenmerken (seks, sociaal-economische status, etnische herkomst, gezinssamenstelling) actueel. In dit onderzoek wordt, met behulp van secundaire analyses op het VOCL-cohort 1989, nagegaan in hoeverre de afstand tussen wiskunde- en taalprestaties van leerlingen in het voortgezet onderwijs (ofwel de relatieve wiskundevoorsprong ten opzichte van taal) intermedieert in de samenhang tussen de genoemde achtergrondkenmerken en exacte en technische keuzen. Uit de analyses blijkt dat, naast de belangrijkste verklarende variabele geslacht, de afstand tussen wiskunde- en taalvaardigheid een relevante bijdrage levert in het voorspellen van de keuze voor exacte en technische richtingen in het voortgezet onderwijs. Deze afstand tussen wiskunde- en taalvaardigheid wordt op zijn beurt voorspeld door het opleidingsniveau van de ouders, seks en, voor leerlingen afkomstig uit de Antillen, Molukken of Suriname, door gezinssamenstelling.

Inleiding

De laatste decennia is internationaal sprake van een groeiend tekort aan technisch en exact afgestudeerde betreders van de arbeidsmarkt. Behalve de toegenomen technologisering, die een grotere behoefte aan technisch geschoolden zou genereren (o.a. SER, 1995), wordt het tekort ook veroorzaakt door juist afnemende belangstelling voor technische en exacte vakken in het onderwijs. Om toch aan de groeiende vraag te kunnen voldoen worden door de overheid en het bedrijfsleven veelal ad hoc maatregelen getroffen of wordt ad hoc beleid gevoerd. Daarnaast bestaat behoefte aan zicht

op het potentieel aan technici en B-wetenschappers, dus op studenten die mogelijk deze richtingen zouden kunnen kiezen. Overduidelijk lijkt het potentieel onder meisjes en vrouwen, omdat verschillen in keuze en deelname aan studierichtingen met name seksespecifiek zijn. Recent wordt, ook in Nederland, gezocht naar reserves aan technisch potentieel onder allochtone groepen, omdat ook verschillen naar etnische herkomst worden vastgesteld. Ook van andere 'herkomst' variabelen, zoals economisch milieu en gezinssamenstelling wordt bestudeerd of ze (eventueel in interactie) de keuze voor technische en exacte studierichtingen beïnvloeden.

Deze zoektocht naar potentieel leidt uiteraard ook naar de vraag welke factoren intermediairen in de samenhang tussen seks, etnische herkomst, sociaal-economische status en gezinssamenstelling met exacte en technische keuzen. Veel factoren in de persoonlijke, gezins- en schoolsfeer zijn, met name in gender-onderzoek, onderwerp van onderzoek geweest; een belangrijke determinant bleek de prestatiescore van de leerlingen in het vak wiskunde te zijn.

Omdat is gebleken dat naast seks ook de meer socio-economische en etnische achtergrondkenmerken verantwoordelijk zijn voor het niet exact of technisch kiezen, is het zinvol onderzoek te doen naar de invloed van mogelijke culturele determinanten op de relatie tussen de genoemde achtergrondkenmerken en de keuze voor exacte en technische richtingen in het voortgezet onderwijs. In dit onderzoek betekent dat, dat de prestatiescore wiskunde als voorspeller wordt vervangen door een score die de verhouding tussen wiskunde en taal uitdrukt (de afstand tussen wiskunde- en taalscore). De leidende idee daarbij is dat het (selectie)vak wiskunde een meer stabiele maat is inzake intellectuele mogelijkheden en dat taalbeheersing meer variabel, cultureel beïnvloed is. Via de afstandscore incalculeert men dan subcultu-

rele invloeden in voorafgaande leerprestaties.

Het databestand van het in 1989 gestarte longitudinale cohort voorgezet onderwijs (VOCL '89) biedt mogelijkheden om hypothesen ten aanzien van verschillen in deelname aan exacte vakken en technische opleidingen tussen groepen met (een combinatie van) de vier genoemde achtergrondkenmerken te onderzoeken. Het bestand bevat ook data om de samenhang daarvan met het *verschil* tussen taal- en wiskundevaardigheid, in plaats van een beperking tot de twee afzonderlijke maten voor taal en wiskunde, te analyseren.

1 (Inter)nationaal onderzoek

In het onderzoek naar de prestaties in en de keuze voor exacte vakken in het voortgezet- en beroepsonderwijs neemt, tot aan het eind van de jaren tachtig, het thema sekseongelijkheid een centrale plaats in. Zowel nationaal als internationaal is zeer veel gepubliceerd over sekseverschillen in schoolloopbanen. Ook is rijkelijk gedocumenteerd dat, hoewel er grote verschillen blijken te bestaan tussen verschillende subgroepen, keuzen en prestaties in het voortgezet onderwijs voor alle groepen relevant van invloed zijn op de uiteindelijke keuze voor een exacte vervolgopleiding (Dekkers, 1993; Maple & Stage, 1991). We vatten in deze theoretische verantwoording samen op welke wijze de (ontwikkeling van de) prestaties verschillen, niet alleen naar sekse, maar ook naar etnische herkomst, sociaal-economische achtergrond en samenstelling van het gezin van de leerlingen.

Uit de Third International Mathematics and Science Study (TIMSS) komt naar voren dat prestaties in en attitudes ten opzichte van exacte vakken, ook internationaal, sterk samenhangen met *seks*. Jongens presteren beter in de science-vakken natuurkunde en scheikunde en vinden deze vakken ook leuker dan meisjes. Voor mathematics geldt dat de verschillen tussen jongens en meisjes, zowel in prestaties als in attitude in de afgelopen jaren zijn afgenomen (Bos, Knuver & Kuiper, 1997).

Uit een longitudinaal onderzoek naar de determinanten van exacte vakkenkeuze blijkt

dat voor jongens en meisjes dezelfde factoren samenhangen met de (voorgenomen) keuze voor exacte vakken (Dekkers, 1993). Met name de geschiktheid volgens relevante anderen en het leuk, interessant en gemakkelijk vinden van een vak hangen samen met de voorgenomen vakkenkeuze. De samenhangen verschillen dus niet tussen jongens en meisjes, maar liggen wel op een ander scoreniveau; jongens scoren hoger op (de predictoren van) wiskunde en natuurkunde, meisjes op (die van) Frans.

In een studie van Johnson (1996) wordt een overzicht gegeven van een aantal grootschalige, nationale en internationale onderwijsprestatietellingen, waarin onder andere sekseverschillen tot uiting komen. Uit de beschrijving van de resultaten van deze prestatietellingen valt af te leiden dat, ondanks grote verschillen in populatie, opzet en methode van onderzoek, grote overeenkomsten bestaan in de uiteindelijke conclusies ten aanzien van sekseverschillen in prestaties in en attitudes ten opzichte van de verschillende vakken. Die conclusie is dat er op het gebied van lezen en schrijven in het basisonderwijs prestatieverschillen bestaan in het voordeel van meisjes. Deze verschillen verdwijnen tijdens het voortgezet onderwijs. Jongens presteren gemiddeld beter dan meisjes in 'mathematics' en 'science', deze voorsprong neemt toe tijdens de schoolcarrière. Belangstelling voor en attitude tegenover de betreffende vakken lopen hiermee gelijk op.

Vanaf 1990 wordt in studies naar verschillen in schoolloopbanen, mede onder invloed van het (dreigende) tekort aan technisch en exact geschoold personeel, in toenemende mate naar de positie van sociaal-culturele groepen gekeken. Ook sekseverschillen in prestaties in en keuzen voor exacte vakken en richtingen worden steeds meer bestudeerd in relatie tot *sociaal-economische en sociaal-culturele herkomstkenmerken*. Zo beschrijven Maple & Stage (1991) in een overzichtsstudie in hoeverre de keuze voor een exacte vervolgopleiding wordt beïnvloed door achtergrondkenmerken van leerlingen (zoals sekse, sociaal economische achtergrond en etnische herkomst), aanleg van leerlingen en eerdere ervaringen en gemaakt keuzen in het voortgezet onderwijs. Zoals gezegd luidt hun conclusie dat, hoewel er grote verschillen blijken te

bestaan tussen de verschillende subgroepen, keuzen en prestaties in het voortgezet onderwijs voor alle groepen van invloed zijn op de uiteindelijke keuze voor een exacte vervolgopleiding. Ook Ainsley & Dale (1997) rapporteren over een studie waarin behalve de samenhang van sekse met de keuze voor en prestaties in wiskunde ook de rol van sociaal-economische en etnische herkomst is onderzocht. Na controle voor eerdere prestaties bleek alleen sekse van belang. De andere achtergrondkenmerken van leerlingen (opleiding ouders en etnische herkomst) zijn alleen indirect, via de eerdere prestaties, van invloed op de keuzen voor en prestaties in wiskunde in het voortgezet onderwijs.

In een vervolgonderzoek op het leerlingen-cohort uit het 'School Matters'-onderzoek beschrijft Sammons (1995) de omvang en stabiliteit van sekseverschillen en etnische en sociaal-economische verschillen in (absolute) onderwijsprestaties en de (relatieve) vooruitgang in lezen en wiskunde, vanaf het begin van het basisonderwijs tot aan het eind van het voortgezet onderwijs. Uit de resultaten blijkt dat sekseverschillen en socio-economische verschillen in prestaties in beide vakken tijdens het voortgezet onderwijs constant zijn gebleven of zelfs zijn toegenomen en daarmee een afspiegeling vormen van al in het basisonderwijs vastgestelde verschillen tussen jongens en meisjes, en hogere en lagere sociale milieus. Het patroon van etnische verschillen is echter gedurende deze periode, zowel in de absolute prestaties als in de relatieve vooruitgang van leerlingen, aanzienlijk veranderd. Achterstanden van allochtone leerlingen in het basisonderwijs blijken in het voortgezet onderwijs te zijn omgekeerd. In de laatste jaren van het voortgezet onderwijs, op zestienjarige leeftijd, bleken leerlingen uit de verschillende etnische groepen sneller vooruit te zijn gegaan in deze periode en gemiddeld beter te presteren in beide vakken dan hun autochtone medeleerlingen.

Nederlands onderzoek, waarin expliciet een relatie gelegd wordt tussen geslacht en etnische herkomst in samenhang met diverse aspecten van schoolloopbanen, is schaars. Als dat al gebeurt dan betreft het veelal deelaspecten van onderwijsloopbanen, zoals voortijdig school-

verlaten, zittenblijven, et cetera. In de cohort-studies van de Landelijke Evaluatie van het Onderwijsvoorrrangsbeleid (LEO) worden sinds 1988 leerlingen gevolgd in hun gang door het basis- en voortgezet onderwijs. Resultaten uit deze studies die hier van belang zijn kunnen als volgt worden samengevat: meisjes halen lagere resultaten voor wiskunde dan jongens, met name in de eerste jaren v.o.; de sekseverschillen worden bij alle etnische groepen teruggevonden, maar zijn groter in de allochtone populatie (Suhre, De Wit & Mulder, 1996; Dekkers, 1996). Daarbij wordt veelal aangenomen dat verschillen tussen allochtone en autochtone leerlingen in prestaties in en keuzen voor exacte en technische vakken niet langer relevant zijn wanneer rekening wordt gehouden met sociaal-economische verschillen (De Fraiture, Dronkers & Van Erp, 1997). In diverse onderzoeken (o.a. Van 't Hof & Dronkers, 1993, Wolbers & Driessen 1996) wordt geconstateerd dat verschillen in schoolsucces tussen de belangrijkste etnische categorieën in het voortgezet onderwijs, verklaard kunnen worden met behulp van dezelfde factoren als voor Nederlandse leerlingen. Ook voor deze leerlingen is het sociaal milieu, bepaald aan de hand van het beroeps- en opleidingsniveau van de ouders, de belangrijkste factor voor een succesvolle schoolloopbaan.

In een vervolgonderzoek op het VOCL-cohort waarin de relatie tussen de achtergrondkenmerken sekse, sociaal milieu en etnische herkomst en schoolsucces is onderzocht, blijkt echter dat sprake is van complexe interactie-effecten tussen de drie factoren (Dekkers, Bosker & Driessen, 1999). Uit de onderzoeksresultaten blijkt dat schoolsucces niet eenvoudig voorspeld kan worden door verwachte additieve of multiplicatieve effecten van de genoemde achtergrondvariabelen. Zo is de situatie van allochtone meisjes beter dan verwacht. Zij kiezen weliswaar minder vaak exacte en technische vakken in het voortgezet onderwijs, maar presteren in het algemeen niet slechter dan autochtone meisjes. Nederlandse jongens uit lagere sociale milieus presteren echter minder dan verwacht. Zij hebben na zes jaar voortgezet onderwijs een achterstand ten opzichte van de overige groepen, ook in vergelijking met allochtone jongens uit hetzelfde milieu.

De Fraiture, Dronkers en Van Erp (1997) bekijken in hun onderzoek op het cohort primair onderwijs (PRIMA) de verschillen tussen groepen in de *afstand tussen rekenen en taal* in het basisonderwijs; geslacht en etnische achtergrond blijken de belangrijkste voorspellende factoren zijn. Ouderlijke opleiding verklaart slechts een gering percentage van de afstand tussen taal- en rekenprestaties. Daarnaast blijkt de samenstelling van het gezin een verklarende factor. Zoals ook uit een onderzoek van McNab en Murray (1985) naar de invloed van de gezinssamenstelling op wiskunde-prestaties is gebleken, lijkt de aanwezigheid van een vader in het gezin de rekenvaardigheid van zowel zonen als dochters te stimuleren.

Geconcludeerd wordt dat taal- en rekenprestaties door verschillende factoren worden beïnvloed. Rekenen is daarbij veel minder een cultuurgebonden vaardigheid dan taal waardoor de taalvoorsprong ten opzichte van de rekenvaardigheid voor verschillende groepen leerlingen verschilt. De afstanden tussen taal- en rekenscores variëren afhankelijk van de mate van integratie in een bepaalde samenleving, het opleidingsniveau van de moeder en de aanwezigheid van een vader in het gezin.

De vraag kan worden gesteld in hoeverre dit verschil tussen taal- en rekenprestaties een indicatie is voor het verschil in de latere voorkeur voor alfa, beta en gamma opleidingen. Bij de studiekeuze gaat het immers om een afwijking van alternatieven, in plaats van een keuze voor één richting. Bij het onderzoek naar studiekeuze is het daarom onjuist zich op de aantrekkelijkheid van één van de alternatieven te concentreren (bijvoorbeeld een exacte- of een talenstudie). Men moet juist beide alternatieven tegelijkertijd analyseren. Daarom staat in deze studie, in navolging van De Fraiture, Dronkers & Van Erp (1997) de verhouding tussen wiskunde- en taalvaardigheid centraal. Deze verhouding analyseren wij als het verschil tussen taal- en wiskunde-prestaties.

2 Vraagstellingen

De beschreven verschillen in taal- en wiskunde-prestaties en de afstand daartussen, leiden tot een tweetal onderzoeksvragen met betrekking tot de omvang van die afstandscores in het

voortgezet onderwijs voor de verschillende (sub)groepen en de samenhang daarvan met exacte of technische keuzen. Deze afstandscore (wiskunde – taal, omdat keuzen voor exacte en technische richtingen worden voorspeld) is dan de mate waarin hoger gescoord wordt op wiskunde dan op taal, ongeacht het scoreniveau. De score kan uit een positief of negatief getal bestaan; het laatste doet zich voor als de taalscore hoger is.

Vraagstellingen:

1. Zijn er verschillen in de afstandscore wiskunde – taal (ofwel de wiskundevoorsprong ten opzichte van taal) in het derde jaar van het voortgezet onderwijs naar geslacht, etnische herkomst, sociaal milieu en gezinssamenstelling en combinaties van deze variabelen, rekening houdend met de afstandscores in het eerste leerjaar?
2. In hoeverre beïnvloeden verschillen in de wiskundevoorsprong op taal in het derde jaar van het voortgezet onderwijs de vakkenkeuze in het AVO/VWO en de richtingkeuze in het LBO, wanneer gecontroleerd wordt voor verschillen in achtergrondkenmerken en voor de afstandscores in het eerste leerjaar?

3 Onderzoekopzet

Databestand¹

Ter beantwoording van de in dit artikel gestelde onderzoeksvragen is gebruik gemaakt van onderzoeksgegevens zoals die oorspronkelijk zijn verzameld ten behoeve van het cohort VOCL '89. Dit betreft een longitudinaal onderzoek naar schoolloopbanen van leerlingen in het voortgezet onderwijs. De eerste metingen voor het onderzoek zijn verricht in het schooljaar 1989/90 bij bijna 20.000 leerlingen die op dat moment in het eerste leerjaar zaten. Deze leerlingen waren afkomstig van 381 scholen, waarbij alle typen met uitzondering van het voortgezet speciaal onderwijs vertegenwoordigd waren. In grote lijnen kan de steekproef als representatief worden beschouwd voor alle leerlingen die in het betreffende schooljaar van het voortgezet onderwijs zaten (Driessen, Dekkers & Bosker, 1997). In het eerste jaar van het onderzoek is via vragenlijsten bij scholen en

ouders informatie verzameld met betrekking tot de sociaal-economische achtergrond, etnische herkomst en het geslacht van de leerlingen. Daarnaast is jaarlijks vastgesteld welke onderwijspositie (de combinatie van leerjaar en schooltype) de leerlingen innemen, zijn zowel in het eerste als in het derde leerjaar toetsen afgenomen voor taal en wiskunde en is van alle leerlingen informatie verzameld over het vakkenpakket in het examenjaar.

Variabelen

Voor de analyses in dit onderzoek zijn de volgende variabelen van belang.

Geslacht. Deze variabele behoeft geen verdere toelichting: (1) jongens, (2) meisjes.

Sociaal milieu. Het sociaal milieu waaruit de leerling afkomstig is, wordt geïndiceerd via het opleidingsniveau van de ouders. Hierbij wordt uitgegaan van het hoogst behaalde onderwijsniveau binnen het gezin. De indeling is: (1) geen lager onderwijs; (2) lager onderwijs; (3) voortgezet onderwijs, lagere trap (lbo, mavo en de leerjaren 1-3 havo, atheneum en gymnasium); (4) voortgezet onderwijs, hogere trap (mbo en de leerjaren 4 en hoger van havo, atheneum en gymnasium); (5) hoger onderwijs, eerste fase (hbo en kandidaats w.o.); (6) hoger onderwijs, tweede fase (doctoraal w.o.).

Etnische herkomst. Voor de bepaling van dit kenmerk is uitgegaan van het geboorteland van de ouders. Daarbij gold als uitgangspunt dat een kind wordt beschouwd als zijnde van allochtone herkomst wanneer tenminste één van de ouders allochtoon is. Kinderen uit gemengde gezinnen worden dus als allochtoon beschouwd. De resulterende variabele kent de categorieën: (1) Turkije/Marokko; (2) ASM (Aruba, Antillen, Suriname en Molukken); (3) Overige OVB landen (categorieën die in het kader van het Onderwijsvoorrrangsbeleid, OVB, met onderwijsachterstanden worden geassocieerd, met name de Zuideuropese landen, zoals Spanje, Italië, Griekenland); (4) Overig allochtoon (overige geïndustrialiseerde, westerse landen zoals Duitsland, België, Engeland, VS); (5) Nederland.

Samenstelling gezin. (1) beide ouders aanwezig; (2) eenoudergezin.

Toetsprestaties. In het eerste en het derde leerjaar zijn prestatietoetsen afgenomen. In het eerste leerjaar was dat een speciaal door het CITO ontwikkelde versie van de Entreetoets voor de leerstofgebieden taal, wiskunde en informatieverwerking. In het derde leerjaar zijn twee leerstofgebieden opnieuw getoetst, Nederlandse taal en wiskunde. Deze toetsing is beperkt gebleven tot de onvertraagde leerlingen die op dat moment in leerjaar 3 van de cohortscholen zaten.

Afstandscores. Voor iedere leerling is de afstand tussen de (gestandaardiseerde) wiskunde- en taalscore berekend in het eerste en het derde leerjaar;

$$\text{Afstandscore} = \text{Wiskundescore} - \text{Taalscore}$$

Vakkenpakket. Binnen de variabele 'vakkenpakket' kunnen twee aspecten worden onderscheiden. Voor de leerlingen in het *beroepsonderwijs* is voor de jaren 1992 tot en met 1995 bekend op welke afdeling ze zaten. Omdat voor deze leerlingen de vakkenpakket-gegevens in strikte zin ontbreken, zal de indeling in afdelingen (richtingen) worden gebruikt als indicator voor het vakkenpakket. Onderscheiden worden (1) techniek en landbouw; (2) dienstverlening, economie en handel.

Voor de leerlingen in het *algemeen voortgezet onderwijs* zijn de vakkenpakketten wel bekend. Uit deze vakkenpakket-gegevens is een variabele geconstrueerd waarin het aantal exacte vakken, wiskunde, natuurkunde en scheikunde is gesommeerd (vgl. Driessen, Dekkers & Bosker, 1997). Daarbij dient wel rekening te worden gehouden met het feit dat het aantal gekozen vakken per onderwijstype verschilt. Voor MAVO en HAVO geldt dat leerlingen examen dienen te doen in minimaal zes vakken, voor het VWO daarentegen is het minimum zeven vakken. Bij de analyse van de gegevens zal daarom niet alleen een onderscheid worden gemaakt tussen LBO en AVO/VWO, maar zullen de analyses binnen de laatste groep ook per onderwijstype worden uitgevoerd.

Analyses

Binnen de analyses die gepresenteerd worden ter beantwoording van de eerste onderzoeksvraag kunnen twee fasen worden onderschei-

den. In de eerste fase wordt de relatie tussen de afstand tussen wiskunde en taal in het derde leerjaar en de afzonderlijke achtergrondkenmerken (sekse, sociaal milieu, etnische herkomst en gezinssamenstelling) beschreven met behulp van frequentieverdelingen en zogenaamde 'one-way'-variantie-analyses². Vervolgens wordt aan de hand van covariantie-analyses en stepwise-regressie-analyses nagegaan of er sprake is van interactie-effecten tussen deze achtergrondkenmerken en in hoeverre de afstand tussen wiskunde en taal in het derde leerjaar voorspeld kan worden wanneer alle achtergrondkenmerken tegelijkertijd in beschouwing worden genomen, ook wanneer gecontroleerd wordt voor de afstandscores in het eerste leerjaar.

Op dezelfde wijze zijn de richtingkeuze in het lager beroepsonderwijs en de vakkenkeuze in het algemeen voortgezet onderwijs geanalyseerd. Met behulp van variantie-analyses wordt de relatie tussen de achtergrondkenmerken, afstandscores in het eerste en het derde jaar en de keuze voor exact beschreven. Vervolgens wordt getoetst in hoeverre hierbij sprake is van een lineaire relatie tussen de afstand tussen wiskunde en taal en de keuze voor een exacte/technische richting in het voortgezet onderwijs, daartoe is gebruik gemaakt van lineaire, stepwise-regressie-analyses. Hiermee kan de tweede onderzoeksvraag beantwoord worden.

In de analyses ter beantwoording van de onderzoeksvragen zijn alleen die leerlingen opgenomen die zowel in het eerste als in het derde leerjaar deel hebben genomen aan de prestatietoetsen. Hierbij moet rekening gehouden worden met het feit dat de toetsen in het derde leerjaar alleen bij de onvertraagde leerlingen zijn afgenomen. Dit kan een vertekend beeld geven. Om te controleren in hoeverre er sprake is van selectieve uitval, is de groep leerlingen die aan de toetsen in beide leerjaren hebben deelgenomen vergeleken met de totale onderzoeksgroep. Uit de resultaten blijkt dat de groep leerlingen waarbij in het derde jaar geen toets is afgenomen zowel op wiskunde als op taal significant lager scoort dan de onvertraagde leerlingen. Deze verschillen zijn groter op de wiskundetoets dan op de taaltoets (Wiskunde; $t = -7.23$; $p = .000$; Taal; $t = -4.60$; $p = .000$). De groep onvertraagde leerlingen scoort gemiddeld iets hoger op wiskunde dan op taal

(gem. = .02). Voor de groep leerlingen die voor het derde jaar zijn uitgevallen is dit net andersom, zij scoren gemiddeld hoger op taal dan op wiskunde (gem. = -.03). De verschillen tussen de diverse subgroepen zijn voor beide groepen gelijk.

Bij het trekken van conclusies zal rekening moeten worden gehouden met het feit dat het hier dus de 'betere leerlingen' betreft. De in de analyses betrokken groep leerlingen heeft gemiddeld een grotere wiskundevoorsprong ten opzichte van de taalvaardigheid dan in de totale groep het geval is.

4 Resultaten

4.1 De afstand tussen wiskunde en taal in het derde jaar van het voortgezet onderwijs

Ter beantwoording van de eerste onderzoeksvraagstelling is voor de afstandscores in het derde leerjaar een variantie-analyse uitgevoerd, waarbij de hoofd- en interactie-effecten van de achtergrondkenmerken en de afstandscores in het eerste leerjaar op de afstand tussen wiskunde en taal in het derde jaar stapsgewijs zijn getoetst. Dit houdt in dat in het eerste model alleen de effecten van de achtergrondkenmerken zijn getoetst. Vervolgens is getoetst in hoeverre deze effecten significant blijven wanneer gecontroleerd wordt voor de afstandscores in het eerste leerjaar (model 2), waarna tot slot alle niet-significante interactie- en hoofdeffecten uit het model zijn weggelaten (model 3). De resultaten hiervan worden weergegeven in Tabel 1.

Uit de tabel blijkt dat zowel in het eerste als in het laatste model voor het derde leerjaar sprake is van interactie tussen de variabelen *etnische herkomst* en *samenstelling gezin*. De groep ASM-leerlingen uit eenoudergezinnen blijkt af te wijken van de overige groepen. Zij hebben een taalvoorsprong ten opzichte van de wiskundevaardigheid (Afstand3; gem. = -.297), terwijl in de overige etnische groepen, net als in de gezinnen met beide ouders, gemiddeld even hoog op wiskunde als op taal wordt gescoord (Afstand3; gem. = .001). De wiskundevaardigheid blijft voor deze leerlingen dus niet achter ten opzichte van de taalvaardigheid. Voor leerlingen uit de ASM-groep is het effect van het ontbreken van een ouder, in veel geval-

Tabel 1

De invloed van achtergrondkenmerken op de afstand tussen wiskunde en taal in het derde leerjaar, gecontroleerd voor afstandsscores in het eerste jaar

Bron	DF	Model 1		Model 2		Model 3	
		F	Eta ²	F	Eta ²	F	Eta ²
Afstand 1	-	-	-	13.48	.002	310.31	.036
Geslacht	1	274.37	.032	1.90	n.s	120.36	.014
Opleiding	5	2.64	.002	.50	n.s	2.79	.000
Etn.herkomst	4	1.35	n.s	.75	n.s	1.29	n.s
Sam.gezin	1	.57	n.s	2.02	n.s	1.05	n.s
Gesl*Opl	5			.86	n.s		
Gesl*Etn	4			.94	n.s		
Gesl*Sam	1			.35	n.s		
Opl*Etn	20			.97	n.s		
Opl*Sam	5			.59	n.s		
Etn*Sam	4	3.47	.002	2.30	n.s	3.40	.002
Model		20.33		10.79		39.16	
% R ²		3.50		7.40		6.90	

Model 1: wordt niet gecontroleerd voor afstand 1; Model 2; volledige model met afstand 1; Model 3: model met afstand 1 zonder niet-significante effecten

len de vader, dus sterker van invloed op de wiskundevaardigheid dan voor de overige groepen het geval is. Leerlingen uit deze groep scoren gemiddeld iets lager op taal dan Nederlandse leerlingen (respectievelijk 9.51 en 9.94), maar het verschil in wiskundevaardigheid is aanzienlijk groter (9.22 versus 9.98). Voor deze groep geldt dat de vader in veel gevallen vaker en al gedurende langere tijd ontbreekt. Nederlandse eenoudergezinnen ontstaan veelal als gevolg van een scheiding, waardoor de vader pas op latere leeftijd ontbreekt.

Naast dit interactie-effect is er sprake van een hoofdeffect van de variabele *geslacht* en van de variabele *ouderlijke opleiding (ses)* waarbij voor jongens en leerlingen uit de middelste opleidingsniveaus sprake is van een wiskundevoorsprong ten opzichte van de taalvaardigheid. Voor meisjes en leerlingen uit de overige opleidingscategorieën geldt het omgekeerde; binnen deze groepen is sprake van een taalvoorsprong.

Wanneer rekening wordt gehouden met verschillen in de afstand tussen wiskunde en taal in het eerste jaar van het voortgezet onderwijs wordt het percentage verklaarde variantie bijna verdubbeld. De invloed van de achtergrondvariabelen geslacht, opleiding en etnische herkomst x gezinssamenstelling wordt daarmee bijna gehalveerd. Deze variabelen zijn dus met

name van invloed op de afstand tussen wiskunde en taal in het eerste leerjaar, maar zijn daarnaast ook nog direct van invloed op de afstandsscores in het derde jaar.

In de tot slot uitgevoerde regressie-analyse blijken de afstandsscores in het eerste leerjaar en het geslacht van de leerlingen de belangrijkste voorspellers voor de afstand tussen wiskunde en taal in het derde jaar te zijn. Waarbij voor de variabele *afstand 1* een positief effect wordt gevonden, hoge afstandsscores in het eerste jaar gaan samen met hoge afstandsscores in het derde leerjaar ($Beta = .194$). Voor de variabele *geslacht* wordt een negatief verband gevonden met de afstandsscores in het derde jaar, dit heeft te maken met de wijze van codering. Jongens (code 1) hebben gemiddeld een hogere, positieve afstandscore dan meisjes (code 2) ($Beta = -.121$). Het in het derde model vastgestelde effect van de *ouderlijke opleiding* is weliswaar significant, maar voegt weinig toe aan de verklaarde variantie.

4.2 Keuze van vakken en studierichtingen in het voortgezet onderwijs

Bij de analyses met betrekking tot de vakkenkeuze is onderscheid gemaakt tussen richtingkeuze in het LBO (dichotome variabele: wel/niet technische richting gekozen) en de pakketkeuze in het AVO/VWO, waarbij een indeling is gemaakt naar het aantal exacte vak-

ken dat gekozen is in het pakket. Aangezien de analyses tot nu toe alleen betrekking hadden op de onvertraagde leerlingen, is ook voor de vakkenkeuze alleen uitgegaan van deze groep.

Relaties tussen achtergrondkenmerken, de afstand tussen wiskunde en taal en de richting-keuze in het LBO

Via variantie-analyses is nagegaan in hoeverre de keuze voor techniek in het LBO verklaard kan worden uit de afstand tussen wiskunde en taal in het derde leerjaar³. Ook wanneer gecontroleerd wordt voor verschillen in afstandscores in het eerste leerjaar en voor verschillen in achtergrondkenmerken. De resultaten hiervan worden weergegeven in Tabel 2.

Uit de tabel blijkt dat de afstand tussen wiskunde en taal in het eerste en het derde jaar samen ruim 10% van de variantie verklaren (model 1). Wanneer de achtergrondkenmerken aan het model worden toegevoegd stijgt de verklaarde variantie naar ruim 70%. Van deze kenmerken blijkt alleen *geslacht* significant van invloed op de keuze voor techniek, jongens kiezen gemiddeld vaker voor een technische opleiding dan meisjes. Daarnaast is er sprake van interactie tussen de variabelen *ouderlijke opleiding* en *geslacht* ($F=2.94$; $p=.012$). Deze interactie is nader onderzocht door het effect van opleiding van de ouders voor zowel jongens als meisjes apart te bekijken. Daaruit blijkt dat alleen bij de

jongens een effect van de ouderlijke opleiding op de keuze voor techniek kan worden gevonden: jongens uit het hoogste opleidingsniveau kiezen daarbij minder vaak voor een technische opleiding.

In de aansluitend uitgevoerde stepwise regressie-analyse blijken de variabelen *geslacht* en de afstand tussen wiskunde en taal in het derde jaar het belangrijkste deel van de variantie verklaren, de overige variabelen voegen hier verder niets aan toe. De variabele *geslacht* verklaart de meeste variantie (71%), de afstandscore in het derde jaar voegt daar nog 1% aan toe.

De variabele *geslacht* is dus de belangrijkste voorspeller voor het al dan niet kiezen voor een technische richting in het LBO. Jongens kiezen vaker voor deze richting dan meisjes. Daarnaast is de afstand tussen wiskunde en taal in het derde jaar van het voortgezet onderwijs van invloed op de richtingkeuze. Ook wanneer gecontroleerd wordt voor het geslacht van leerlingen blijkt dat de keuze voor een technische richting samenhangt met een hoge, positieve afstandscore in het derde leerjaar. Leerlingen met een wiskundevoorsprong kiezen dus vaker voor techniek.

Vakkenkeuze in het AVO/VWO

Voor de drie onderwijstypen in het algemeen voortgezet onderwijs (MAVO, HAVO en VWO) is een variabele berekend waarin het

Tabel 2
De invloed van de afstand tussen wiskunde en taal in het derde jaar van het voortgezet onderwijs op de keuze voor een technische richting in het LBO, gecontroleerd voor afstandscores in het eerste leerjaar en achtergrondkenmerken

Bron	DF	Model 1		Model 2	
		F	Eta ²	F	Eta ²
Afstand 3	1	78.86	.031	5.10	.002
Afstand 1	1	138.13	.054	5.90	.002
Geslacht	1			976.61	.260
Opleiding	5			.23	.000
Etn.herkomst	4			-	-
Sam. Gezin	1			-	-
Gesl*Opl				2.94	.005
Model		143.83		557.03	
% R ²		10.60		72.20	

NB. In de tabel zijn alleen de significante effecten opgenomen.

Model 1: alleen afstandscores; Model 2: gecontroleerd voor achtergrondkenmerken en afstandscores in het eerste jaar.

aantal exacte vakken, wiskunde, natuurkunde en scheikunde is gesommeerd. Hierbij is, zoals gezegd, rekening gehouden met het feit dat het aantal gekozen vakken per onderwijstype verschilt. Ook voor deze onderwijstypen is via variantie-analyses nagegaan in hoeverre het aantal gekozen exacte vakken verklaard kan worden uit de afstand tussen wiskunde en taal in het derde jaar van het voortgezet onderwijs. Ook wanneer gecontroleerd wordt voor verschillen in afstandscores in het eerste leerjaar en voor de verschillen in achtergrondkenmerken. De belangrijkste resultaten hiervan worden voor de totale groep AVO/VWO weergegeven in Tabel 3.

De afstand tussen wiskunde en taal in het eerste en het derde jaar van het voortgezet onderwijs verklaren samen ruim 7% van de variantie in het aantal gekozen exacte vakken. De afstandscores in het eerste leerjaar leveren hierin de grootste bijdrage (5.6%), de afstand tussen wiskunde en taal in het derde jaar voegt hier nog eens bijna 2% aan toe. Wanneer de achtergrondkenmerken geslacht, ouderlijke opleiding, etnische herkomst en gezinssamenstelling aan het model worden toegevoegd stijgt de verklaarde variantie naar 19.5%.

In het model blijken twee interactie-effecten relevant te zijn voor het aantal exacte vakken in het pakket (model 2), de variabelen *afstand 1* en *gezinssamenstelling* interacteren met de

variabele *geslacht*. Deze interacties zijn nader onderzocht door voor zowel jongens als meisjes afzonderlijk te bekijken welke relatie er bestaat tussen gezinssamenstelling, respectievelijk de afstand tussen wiskunde en taal in het eerste leerjaar, en het aantal gekozen exacte vakken. De samenstelling van het gezin blijkt alleen voor jongens relevant te zijn (jongens; $F=5.95$; $p=.015$; $Eta^2=.001$; meisjes; $F=.24$; $p=.625$; $Eta^2=.000$), waarbij jongens uit eenoudergezinnen gemiddeld meer exacte vakken in het pakket kiezen dan jongens uit gezinnen waarin beide ouders aanwezig zijn. Wanneer we de afzonderlijke onderwijstypen bekijken blijkt gezinssamenstelling alleen voor de MAVO relevant te zijn, ook hier is sprake van een interactie met de variabele *geslacht*. Voor HAVO en VWO is gezinssamenstelling, noch als hoofdeffect, noch in interactie met *geslacht* relevant voor de keuze voor exact. De afstand tussen wiskunde en taal in het eerste leerjaar blijkt voor zowel meisjes als jongens relevant te zijn (jongens; $F=12.37$; $p=.000$; $Eta^2=.002$; meisjes; $F=70.59$; $p=.000$; $Eta^2=.013$). Voor beide groepen geldt dat leerlingen met een grote wiskundevoorsprong (ten opzichte van de taalvaardigheid) gemiddeld meer exacte vakken in het pakket kiezen. Dit effect is voor meisjes wel belangrijker dan voor jongens. Hetzelfde interactie-effect wordt teruggevonden in het model voor de MAVO-leerlingen; de afstandscores in het eerste jaar zijn, met name

Tabel 3

De invloed van de afstand tussen wiskunde en taal in het derde jaar van het voortgezet onderwijs op het aantal gekozen exacte vakken in de AVO/VWO, gecontroleerd voor afstandscores in het eerste leerjaar en achtergrondkenmerken

Bron	DF	Model 1		Model 2	
		F	Eta ²	F	Eta ²
Afstand 3	1	73.22	.016	58.11	.013
Afstand 1	1	215.80	.047	65.70	.015
Geslacht	1			160.98	.036
Opleiding	5			17.75	.020
Etn.herkomst	4			n.s	n.s
Sam. Gezin	1			2.64	.001
Gesl*Sam.gezin				4.77	.001
Gesl*Afstand 1				8.48	.002
Model		180.04		96.09	
%R ²		7.60		19.50	

NB. In de tabel zijn alleen de significante effecten opgenomen.

Model 1: alleen afstandscores; Model 2: gecontroleerd voor achtergrondkenmerken en afstandscores in het eerste jaar.

voor meisjes, positief verbonden met het aantal gekozen exacte vakken. Voor de HAVO- en VWO-leerlingen blijkt sprake te zijn van een hoofdeffect van de variabele 'afstand 1': leerlingen met hoge, positieve afstandscores in het eerste leerjaar kiezen gemiddeld meer exacte vakken in het pakket.

De variabelen geslacht, ouderlijke opleiding en afstand tussen wiskunde en taal in het derde leerjaar zijn direct van invloed op het aantal exacte vakken in het pakket. Jongens, leerlingen uit hogere sociale milieus en leerlingen met een grote wiskundevoorsprong ten opzichte van de taalbeheersing in het derde jaar kiezen uiteindelijk vaker voor exact. Deze effecten van geslacht en 'afstand 3' worden ook teruggevonden in de modellen voor de afzonderlijke onderwijstypen, waarbij geldt dat de gevonden sekseverschillen groter zijn binnen de MAVO dan binnen HAVO en VWO en de afstand tussen wiskunde en taal in het derde jaar het belangrijkste is voor de keuze van exacte vakken in de HAVO. De invloed van de ouderlijke opleiding blijkt alleen voor MAVO- en VWO-leerlingen relevant te zijn, deze variabele is binnen de HAVO niet significant.

Uit de aansluitend uitgevoerde stepwise-regressie-analyse blijkt dat de variabele geslacht (net als in de modellen voor de afzonderlijke onderwijstypen) het meest van de variantie in het aantal exacte vakken in het pakket verklaart (bijna 15%). De afstandscores in het eerste- en het derde leerjaar voegen daar nog bijna 3% aan toe. Hoewel de variabele ouderlijke opleiding in de modellen per schooltype nauwelijks extra variantie verklaart wanneer rekening wordt gehouden met geslacht en afstandscores in het eerste en derde jaar, verklaart deze variabele voor de totale groep bijna 1.5% van de variantie (dit is ruim 7% van het totale percentage verklaarde variantie).

5 Samenvatting en conclusie

Binnen het *beroepsonderwijs* is de variabele geslacht de belangrijkste voorspeller voor de keuze voor een technische richting in het examenjaar. Deze variabele verklaart ruim 70% van de variantie in de gekozen richtingen. De afstand tussen wiskunde en taal in het eerste

en derde jaar is weliswaar significant gerelateerd aan richtingkeuze, maar verklaart weinig extra variantie. Daarnaast is voor jongens in het LBO de ouderlijke opleiding relevant, maar niet zoals verwacht in de zin van hoe hoger de ouderlijke opleiding, des te meer technische keuzen; jongens uit de twee middencategorieën kiezen vaker voor een technische richting dan jongens in zowel de laagste als de hoogste opleidingsgroepen.

Ook binnen het *algemeen voortgezet onderwijs* is het geslacht van leerlingen de belangrijkste factor bij de verklaring van verschillen in het aantal gekozen exacte vakken in het pakket; daarbij dient wel opgemerkt te worden dat het percentage totaal verklaarde variantie binnen deze groep beduidend lager is dan in het LBO. Sekseverschillen zijn het grootst binnen de MAVO, dan volgt HAVO en tenslotte VWO. Ook de afstand tussen wiskunde en taal is positief gecorreleerd met de keuze voor exact in het voortgezet onderwijs. Leerlingen met een grote wiskundevoorsprong ten opzichte van de taalvaardigheid (in het eerste en het derde leerjaar) kiezen meer exacte vakken in het pakket dan leerlingen waarbij de wiskundevaardigheid minder voorloopt op de taalbeheersing. Hoewel het absolute percentage verklaarde variantie relatief laag is (gemiddeld 2%) blijkt dit afgezet tegen het totaal percentage variantie hoger te zijn dan in het LBO. De afstandscores in het eerste en het derde jaar zijn dan verantwoordelijk voor ruim 15% van de totaal (door het model) verklaarde variantie. Voor het VWO is de bijdrage van de afstand tussen wiskunde en taal (m.n. in het eerste jaar) nog groter, bijna een kwart van de totaal verklaarde variantie. Daarbij is de invloed van de afstand tussen wiskunde en taal in het eerste jaar groter voor meisjes dan voor jongens. De ouderlijke opleiding is binnen het AVO/VWO, in tegenstelling tot het LBO, positief gecorreleerd met het aantal exacte vakken in het pakket. Leerlingen uit de hogere opleidingsniveaus kiezen gemiddeld meer exacte vakken; alleen voor HAVO-leerlingen is deze relatie niet significant.

De overige achtergrondkenmerken zijn minder belangrijk voor de keuze van een exacte of technische richting. De etnische herkomst van leerlingen is alleen in het MAVO relevant. Allochtone leerlingen kiezen minder vaak voor

een exacte/technische richting dan Nederlandse leerlingen. De gezinssamenstelling is alleen binnen de MAVO in interactie met geslacht en etnische herkomst van belang. Voor de ASM-groep geldt dat leerlingen uit eenoudergezinnen gemiddeld minder exacte vakken kiezen dan leerlingen uit gezinnen waarin beide ouders aanwezig zijn. Voor jongens uit eenoudergezinnen in het algemeen geldt het omgekeerde, zij kiezen juist voor meer exacte vakken in het pakket.

Uit het voorafgaande kan geconcludeerd worden dat van de variabelen uit het analysemodel de variabele geslacht en de afstand tussen wiskunde- en taalvaardigheid de belangrijkste bijdrage leveren in het voorspellen van de keuze voor een exacte of technische richting in het voortgezet onderwijs. Er is sprake van een lineaire relatie tussen de afstand tussen wiskunde en taal en de keuze voor exacte en technische vakken, waarbij leerlingen met een grotere wiskundevoorsprong de 'beste' B-kiezers zijn.

Op zijn beurt wordt de afstand tussen taal- en wiskundevaardigheid in de eerste klas van het voortgezet onderwijs voorspeld door geslacht en door de combinatie van geslacht en ouderlijke opleiding en in de derde klas door geslacht en opleiding, en voor leerlingen afkomstig uit Suriname, de Antillen of de Molukken, door gezinssamenstelling.

Wat betekenen nu deze bevindingen voor het huidige onderzoek naar en debat over de maatregelen tegen het tekort aan technisch en exact personeel in Nederland en daarbuiten? De belangrijkste conclusie uit het onderzoek is dat de door ons onderzochte variabele 'afstand wiskunde - taal' een toegevoegde waarde heeft bij de verklaring van exacte en technische keuzen.

Een goed voorbeeld van het belang van deze relatieve verhouding tussen taal- en wiskundevaardigheid is het effect van het ouderlijk milieu op die verhouding. Hoe hoger het ouderlijk opleidingsniveau, hoe hoger de absolute wiskundescore. Op grond van alleen de relatie tussen wiskundeprestaties en exacte keuzen zou men kunnen concluderen dat leerlingen met ouders met de hoogste opleiding vaker exacte of technische vakken zouden kiezen dan

vergelijkbare leerlingen met ouders met een lage opleiding. Dat is echter niet altijd het geval; als de factor taalvaardigheid in het model wordt meegenomen, krijgen we een genuanceerder beeld. Leerlingen uit de hoogste opleidingsniveaus scoren niet alleen hoger op wiskunde, maar relatief nog hoger op taal dan de overige leerlingen. Met andere woorden: de taalvaardigheid van leerlingen is gevoeliger voor het ouderlijk milieu dan hun wiskundevaardigheid. Deze samenhang tussen ouderlijke opleiding en verhouding tussen taal- en wiskundevaardigheden leidt er toe dat leerlingen met ouders met de hoogste opleiding, met name in het LBO, vaker niet-exacte of niet-technische vakken kiezen dan vergelijkbare leerlingen met ouders met weinig opleiding.

De conclusie heeft, als hij in andere databestanden opnieuw wordt bevestigd, duidelijke implicaties voor de ontwikkeling van de voorkeur van toekomstige studenten voor exacte of technische vakken of studierichtingen. Gezien de stijging van de onderwijsdeelname gedurende de laatste decennia zal het onderwijsniveau van ouders de komende tientallen jaren blijven stijgen. Gegeven onze resultaten zou dit kunnen betekenen dat de verhouding tussen taal- en wiskundevaardigheid zal verschuiven ten gunste van de taalvaardigheid. Niet omdat de wiskundevaardigheid van deze toekomstige leerlingpopulaties met steeds hoger opgeleide ouders achteruitgaat, maar omdat hun taalvaardigheid sneller vooruit gaat. En indien de hier gevonden relatie tussen de verhouding taal- en wiskundevaardigheid en de keuze van exacte of technische vakken in stand blijft, betekent het stijgende opleidingsniveau van toekomstige ouders dat steeds minder belangstelling zal bestaan voor de exacte en technische vakken. Dit pleit ervoor in toekomstig onderzoek en beleid ter versterking van technische en exacte keuzen de aandacht niet alleen te richten op prestaties in de meest voor de hand liggende vakken, maar op de verhouding tussen (prestaties in) verschillende vakken. Het feit dat de huidige B-kiezers juist hoog scoren op wiskunde maar laag op taal leidt, in combinatie met het bovenstaande, tot de conclusie dat er potentiële B-kiezers te vinden zijn bij degenen die zowel boven gemiddeld op wiskunde als op taal presteren.

Noten

1. De gegevens zijn verzameld en beschikbaar gesteld door het Centraal Bureau voor Statistiek (CBS) te Heerlen.
2. Dezelfde analyses zijn ook uitgevoerd voor de afstandscores in het eerste leerjaar. Omdat de relaties met de achtergrondkenmerken voor beide leerjaren vergelijkbaar zijn wordt in dit artikel alleen ingegaan op de resultaten van het derde leerjaar. De resultaten van deze analyses zijn bij de eerste auteur opvraagbaar.
3. Vanwege het dichotome karakter van de afhankelijke variabele (richtingkeuze) zijn dezelfde effecten ook met behulp van een logistische regressie-analyse getoetst. Dit leverde geen extra informatie op. Met het oog op de vergelijking met de resultaten binnen het AVO/WVO worden in dit artikel alleen de resultaten van de variantie-analyses weergegeven.

Literatuur

- Ainsley, J., & Daly, P. (1997). *Student participation in mathematic courses in Australian secondary schools*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Bos, K., Knuver, A., & Kuiper, W. (1997). Prestaties exacte vakken op de internationale meetlat. *Didaktief & School*, 27(7), 21-28.
- Dekkers, H. (1993). Determinanten van seksespecifieke vakkenkeuzen in HAVO en VWO. *Pedagogische Studiën*, 70(6), 445-459.
- Dekkers, H. (1996). Onderwijs(beleid) voor allochtone meisjes. De stand van zaken. *Migrantenstudies*, 12(3), 152-163.
- Dekkers, H., Bosker, R.J., & Driessen, G. (1999). Complex inequalities of educational opportunities: a large-scale longitudinal study on the relation between gender, social class, ethnicity, and school succes. *Educational Research and Evaluation* (accepted).
- Driessen, G., Dekkers, H., & Bosker, R.J. (1997). *Sociaal-etnische en sekseverschillen in loopbanen in het voortgezet onderwijs*. Nijmegen: Instituut voor Toegepaste Sociale wetenschappen.
- Fraiture, J.F.M. de, Dronkers, J. & Erp, M. van. (1997). Etnische diversiteit in afstand tussen taal- en rekenscores tussen jongens en meisjes. *Tijdschrift*

voor nascholing en onderzoek van het rekenwiskundeonderwijs, 17(1), 31-37.

- Hof, L. Van 't, & Dronkers, J. (1993). Onderwijsachterstanden van allochtonen: klasse, gezin of etnische cultuur? *Migrantenstudies*, 9(1), 2-25.
- Johnson, S. (1996). The contribution of large-scale assessment programmes to research on gender differences. *Educational Research and Evaluation*, 2(1), 25-49.
- Maple, S., & Stage, F. (1991). Influences on the choice of math/science major by gender and ethnicity. *American Educational research Journal*, 28(1), 37-60.
- McNab, C., & Murray, A. (1985). Family composition and mathematics achievement. *Scandinavian Journal of Educational Research*, 29, 89-102.
- Sammons, P. (1995). Gender, ethnic and socio-economic differences in attainment and progress: a longitudinal analysis of student achievement over 9 years. *British Educational Research Journal*, 21(4), 465-485.
- Suhre, C., Wit, W. de, & Mulder, L. (1996). *Onderwijsachterstanden in het voortgezet onderwijs*. Nijmegen: Instituut voor Toegepaste Sociale wetenschappen; Ubbergen: Tandem Felix.
- Wolbers, M., & Driessen, G. (1996). Social class or ethnic background? Determinants of secondary school careers of ethnic minority pupils. *The Netherlands' Journal of Social Sciences*, 32(2), 109-126.

Auteurs

D. Uerz is als beleidsonderzoeker in opleiding werkzaam bij het ITS Nijmegen.

H. Dekkers is hoogleraar Onderwijskunde en werkzaam bij de sectie Onderwijs en Educatie en het ITS (beide Katholieke Universiteit Nijmegen).

J. Dronkers is hoogleraar Empirische Sociologie bij de Universiteit van Amsterdam en werkzaam bij het SCO-Kohnstamm Instituut.

Adres: ITS wetenschap voor beleid en samenleving, Postbus 9048, 6500 KJ Nijmegen.

Abstract

Mathematical and language skills and the choice of science subjects in secondary education

D. Uerz, H. Dekkers & J. Dronkers. *Pedagogische Studiën*, 1999, 76, 170-182.

A growing shortage of technical and science graduates means that the issue of determinants of differences in performance and choices in that field based upon, among other things, social background characteristics (gender, sex, ethnic origin, family composition) continues to be a topical one. In this study, secondary analyses of the 1989 VOCL-cohort study are used to determine the extent to which the gap between the mathematical and language achievements of pupils in secondary education, i.e. the extent to which they are relatively ahead in mathematics compared to language, plays a role in linking the above-mentioned background characteristics to the decision to opt for science and technical subjects. The analyses show that in addition to the most important explanatory variable, namely gender, the gap between mathematical and language skills makes a relevant contribution when it comes to predicting the decision to opt for science and technical subjects in secondary education. This gap between mathematical and language skills can, in turn, be predicted on the basis of the educational level of the parents, gender and, for pupils from the Antilles, Moluccas or Surinam, on the basis of family composition.

Realistisch modelleren van problematische wiskundige toepassingsopgaven op de basisschool

Een review van het recent onderzoek

L. Verschaffel, E. De Corte en S. Lasure

Samenvatting

Enkele jaren geleden hebben zowel Greer als Verschaffel, De Corte en Lasure aangetoond dat leerlingen uit de bovenbouw van de basisschool een sterke neiging hebben om hun ervaringskennis omtrent de concrete, reële situatie waarop een rekenvraagstuk betrekking heeft, te negeren tijdens het oplossingsproces. Daarbij maakten zij gebruik van een serie van speciaal ontwikkelde opgaven waarbij er zich vanuit realistisch oogpunt een probleem stelt bij het omzetten van de probleemsituatie in een passend wiskundig model. In dit artikel wordt een aantal vervolgstudies hieromtrent besproken. Eerst wordt ingegaan op enkele onderzoekingen waarin de resultaten uit de bovenvermelde studies gerepliceerd zijn onder vrijwel identieke omstandigheden. Dan worden twee lijnen van onderzoek besproken naar de effectiviteit van bepaalde ingrepen in de experimentele setting die erop gericht zijn bij de leerlingen een meer realistische aanpak en oplossingswijze uit te lokken. De eerste lijn betreft het toevoegen van 'scaffolds' gericht op het verhogen van de alertheid van leerlingen tijdens het oplossen van dergelijke problematische opgaven; de tweede het verhogen van de authenticiteit van de experimentele setting. Daarna komt onderzoek aan bod waaruit gebleken is dat ook (aspirant-)leerkrachten over het algemeen niet geneigd zijn om schoolvraagstukken op een realistische wijze te benaderen. Vervolgens wordt verslag uitgebracht van een onderwijsexperiment waarin men erin geslaagd is om de neiging van leerlingen tot niet-realistisch wiskundig modelleren via aangepaste instructie te doorbreken. Aan het einde van dit artikel wordt kort ingegaan op de theoretische, methodologische en onderwijspraktische implicaties van dit onderzoek.

1 Onrealistisch modelleren aangetoond: aanvankelijke studies

Een veel gehoorde kritiek aan het adres van het traditionele vraagstukkenonderwijs is dat leerlingen op het einde van de basisschool een sterke neiging ontwikkeld hebben om hun ervaringskennis omtrent de concrete situatie waarop een rekenvraagstuk betrekking heeft, compleet te negeren tijdens de verschillende fasen van het oplossingsproces. Dit zou een van de voornaamste verklaringen zijn waarom deze vraagstukkendidactiek er onvoldoende in slaagt om bij leerlingen de bekwaamheid te ontwikkelen in het functioneel gebruik van de geleerde wiskundige begrippen en technieken in reële toepassingssituaties (Burkhardt, 1994; De Corte & Verschaffel, 1985; Nesher, 1980; Treffers & De Moor, 1990). Greer (1993) en Verschaffel, De Corte en Lasure (1994) verzamelden overtuigende empirische steun voor deze neiging van leerlingen om vraagstukken op een oppervlakkige, routinematige manier aan te pakken, zonder rekening te houden met de concrete situatie of context waarop het vraagstuk betrekking heeft. In het onderzoek van Verschaffel e.a. (1994) werd aan een groep van 75 10-11-jarige leerlingen een toets gegeven bestaande uit tien itemparen. Het ene vraagstuk uit elk itempaar was een standaardopgave (S-item) die 'probleemloos' gemodelleerd en opgelost kan worden door middel van één of twee rekenkundige basisbewerkingen met de getallen uit de opgave (bijv. 'Karel heeft 6 planken van 2 meter gekocht. Hoeveel planken van 1 meter kan Karel hiervan zagen?'). Bij de andere opgave uit elk paar vormt de omzetting in een passend wiskundig model wél een probleem (P-item), tenminste als men de context waarrond deze opgave is

opgebouwd 'au sérieux' neemt (bijv. 'Steven heeft 4 planken van 2,5 meter gekocht. Hoeveel planken van 1 meter kan Steven hiervan zagen?'). Tabel 1 geeft een overzicht van de tien P-items uit deze toets. De leerlingen werd niet enkel gevraagd om hun antwoord op te schrijven, maar ook hoe ze aan dit antwoord gekomen waren en of ze daar nog commentaar op wensten te geven. Het aantal antwoorden en/of commentaren op de P-items waaruit bleek dat de leerling rekening gehouden had met de reële opgavensituatie, was alarmerend laag: slechts 17% van alle leerlingreacties op deze P-items was als realistische reacties (RRs) te beschouwen. Bij het zo-even vermelde plankenitem bijvoorbeeld gaf slechts 10% van de leerlingen te kennen dat men in werkelijkheid slechts 2 planken van 1 m kan zagen uit één plank van 2,5 m, en dat derhalve 4 van deze planken slechts 8 planken van 1 m kunnen opleveren (in plaats van 10). De overige leerlingen reageerden vrijwel allemaal met het verwachte, stereotiepe antwoord; $4 \times 2,5 = 10$ zonder enige verdere commentaar. Wel was er een relatief groot verschil in het percentage RRs op de diverse P-items: terwijl het percentage RRs bij acht van de 10 opgaven minder dan 20% bedroeg, steeg dit bij de twee opgaven waarin het quotiënt van een niet-opgaande deling adequaat geïnterpreteerd moet worden (nl. het bus- en het ballon-item) tot ongeveer 50%.

Gelijkaardige resultaten werden bekomen in Greers (1993) onderzoek, waarin een grote groep leerlingen van de eerste graad van het secundair onderwijs in vrijwel identieke omstandigheden als in de studie van Verschaffel e.a. (1994) een analoge toets kreeg. De helft van de P-items uit deze toets waren dezelfde als die uit het onderzoek van Verschaffel e.a. (1994), terwijl de andere helft uit andersoortige P-items bestond. Ook Greer vond op de meeste P-items een bedroevend gering aantal realistische antwoorden of commentaren. Een uitzondering hierop vormden weerom het bus-item en het ballon-item, die beide beduidend meer RRs uitlokten dan de overige P-opgaven.

De alarmerende resultaten van deze studies hebben de voorbije jaren geleid tot nogal wat vervolgonderzoek, dat in onderhavig artikel voorgesteld en besproken wordt. Paragraaf 2 bevat een overzicht van de studies waarin het door Greer (1993) en Verschaffel e.a. (1994) beschreven fenomeen bij leerlingen van diverse nationaliteiten is gerepliceerd onder (vrijwel) identieke omstandigheden. In de twee daaropvolgende paragrafen komen twee reeksen vervolgstudies aan bod, waarin gepoogd is om het aantal RRs op deze P-items te verhogen via diverse ingrepen in de experimentele setting die erop gericht zijn bij de leerlingen een meer realistische aanpak- en oplossingswijze uit te lokken. In de eerste reeks studies, die in paragraaf 3 besproken wordt, gebeurde dit door

Tabel 1
De tien P-items uit de toets van Verschaffel e.a. (1994)

P1	Karel heeft 5 vrienden en Joris heeft 6 vrienden. Karel en Joris besluiten samen een feestje te geven en nodigen al hun vrienden uit. Alle vrienden komen. Hoeveel vrienden zijn er op het feestje? (VRIENDEN)
P2	Steven heeft 4 planken van 2,5 meter gekocht. Hoeveel planken van 1 meter kan Steven hiervan zagen? (PLANKEN)
P3	Iemand giet eerst een volle emmer water van 40 graden en dan een volle emmer water van 80 graden in een kuip. Hoeveel graden is het water in deze kuip? (WATER)
P4	450 soldaten moeten per bus naar het oefenterrein vervoerd worden. Hoeveel bussen zijn er nodig, als je weet dat er 36 soldaten in één bus kunnen? (BUS)
P5	De recordtijd van Jan op de 100 meter is 17 seconden. In hoeveel tijd loopt Jan 1 kilometer? (LOPER)
P6	Bruno en Saskia gaan naar dezelfde school. Bruno woont 17 kilometer van school en Saskia 8 kilometer. Hoever wonen Bruno en Saskia van elkaar? (SCHOOL)
P7	Veerle, Lies, Paul en Sven krijgen van opa een doos met 18 ballonnen. Zij verdelen die ballonnen eerlijk onder elkaar. Hoeveel ballonnen krijgt ieder? (BALLON)
P8	Rob is geboren in 1978. Het is nu 1993. Hoe oud is Rob? (LEEFTIJD)
P9	De afstand tussen twee palen bedraagt 12 meter. Hoeveel stukken touw van 1,5 meter lang moet men aan elkaar knopen om de beide palen te verbinden? (TOUWEN)
P10	Deze vaas wordt gevuld met water. Na 10 seconden staat het water 4 cm hoog. Hoe hoog zal het water staan na 30 seconden? (Deze opgave was vergezeld van een tekening van een gedeeltelijk gevulde, kegelvormige vaas) (VAAS)

de leerlingen alert te maken op het problematisch gehalte van sommige toetsitems of door hen er uitdrukkelijk op te wijzen dat andere dan 'conventionele' antwoorden toegelaten waren. In de tweede reeks onderzoeken, die in paragraaf 4 aan bod komen, trachtte men dit te realiseren door het authenticiteitsgehalte van de experimentele setting te verhogen. In paragraaf 5 volgt een bespreking van enkele constaterende studies waarin men op zoek is gegaan naar bepaalde aspecten van de instructie die mede verantwoordelijk gesteld kunnen worden voor het ontstaan en de ontwikkeling van de neiging van leerlingen tot niet-realistisch modelleren. In paragraaf 6 beschrijven we een construerend onderzoek waarin is nagegaan of het mogelijk is om de dispositie tot niet-realistisch modelleren en oplossen van leerlingen uit de bovenbouw van de basisschool te doorbreken via aangepaste instructie. Aan het einde van dit artikel plaatsen we enkele theoretische en methodologische kanttekeningen bij de gerapporteerde onderzoeksgegevens en bekijken we de implicaties ervan voor de praktijk van het vraagstukkenonderwijs.

2 De studies van Greer (1993) en Verschaffel e.a. (1994) gerepliceerd

De resultaten van Greer (1993) en Verschaffel e.a. (1994) werden gerepliceerd in verschillende landen, waaronder Japan, Nederland,

Noord-Ierland, Venezuela en Zwitserland. Meestal gebeurde dit in het kader van een meer uitgebreide studie naar het effect van bepaalde variaties in de aanbieder van de P-items of in de testsetting. Vanwege deze variaties komen we op deze studies in de volgende paragrafen (paragraaf 3 of 4) nog terug. Maar ter ondersteuning van de alomtegenwoordigheid en de robuustheid van het door Greer (1993) en Verschaffel e.a. (1994) beschreven fenomeen, rapporteren we thans de resultaten van vier studies waarin groepen van leerlingen een gedeelte of alle tien P-items uit het onderzoek van Verschaffel e.a. (1994) aangeboden kregen onder gelijkaardige of identieke testcondities:

- Reusser en Stebler (1997a) pasten de test van Verschaffel e.a. (1994) toe bij 439 leerlingen van het eerste jaar secundair onderwijs uit het Duitssprekend gedeelte van Zwitserland;
- Yoshida, Verschaffel en De Corte (1997) legden dezelfde test voor aan 91 leerlingen van het vijfde leerjaar uit drie willekeurig geselecteerde klassen van een Japanse basisschool;
- Hidalgo (1997) vroeg 119 leerlingen van het vijfde leerjaar uit vier verschillende scholen van Venezuela eveneens deze test op te lossen;
- Caldwell (1995) bood vijf items uit de test van Verschaffel e.a. (1994) aan samen met twee P-items uit de test van Greer (1993) aan 75 jongens en meisjes van het vijfde leerjaar uit drie verschillende scholen in Noord-Ierland.

Tabel 2
Percentages RRs op de P-items uit de studies van Greer (1993), Verschaffel e.a. (1994), Reusser en Stebler (1997a), Yoshida e.a. (1997), Hidalgo (1997) en Caldwell (1995)*

P-item	Greer	Verschaffel	Reusser	Yoshida	Hidalgo	Caldwell
Vrienden	-	20	11	13	23	5
Planken	-	13	14	0	16	-
Water	-	17	21	11	11	-
Bussen	55	49	49	62	11	65
Loper	6	3	5	7	0	0
School	-	3	5	2	1	-
Ballonnen	85	59	75	52	55	81
Verjaardag	-	3	2	0	0	-
Touw	8	0	6	2	0	1
Vaas	2	4	0	4	0	-
Totaal	-	17	18,5	15	12	-

* Voor de volledige tekst van de P-items in de eerste kolom van de tabel, zie Tabel 1. De gemiddelde percentages RRs werden alleen berekend voor die studies waarin de tien P-items werden aangeboden.

De resultaten van deze replicatieonderzoekingen, die in Tabel 2 zijn opgenomen, komen zeer sterk overeen met die van de twee oorspronkelijke studies: het overgrote deel van de leerlingen vertoonde een sterke neiging om hun realiteitsgebonden kennis te negeren bij het oplossen van de meeste P-items, en het relatief percentage RRs voor de afzonderlijke P-items was zeer gelijkaardig als in de twee initiële onderzoekingen. Een opmerkelijke terugkerende vaststelling is dat de twee vraagstukken rond een niet-opgaande deling (nl. het bus- en het ballon-item) telkens veel meer RRs uitlokten dan de overige P-items.

In de studies van Reusser en Stebler (1997a), Hidalgo (1997) en Caldwell (1995) werden leerlingen achteraf ook individueel of klassikaal bevraagd over het waarom van het grote aantal niet-realistische reacties (NRs). Sommige leerlingen beweerden dat ze het 'problematisch' gehalte van bepaalde P-items weliswaar hadden opgemerkt, maar daar in hun uiteindelijk antwoord bewust geen rekening mee hadden gehouden, omdat zij een onderscheid maakten tussen (het oplossen van) een wiskundevraagstuk op school enerzijds en (het oplossen van) een reëel buitenschools probleem anderzijds. Bij wijze van voorbeeld geven we hierna enkele uitspraken van leerlingen tijdens een klasgesprek dat Reusser en Stebler (1997a, p. 317) organiseerden kort na de afname van hun collectieve toets:

- 'I did think about the difficulty, but then just calculated it the usual way. (Why?) Because I just had to find some sort of solution to the problem. And that was the only way it worked. I've got to have a solution, haven't I?'
- 'I suspected that it wouldn't work but I solved it anyway. (Why?) Because otherwise the task wouldn't have been solved.'

Anderen gaven toe dat ze er geen moment aan gedacht hadden dat deze P-items moeilijk op te lossen waren of beter op een andere manier waren opgelost (bijv. 'Simply did not think of it' of 'It would never have crossed my mind to ask whether this task can be solved at all').

De resultaten van deze replicatiestudies bieden overtuigende empirische steun voor het bestaan van een sterke, en bij sommige leerlingen blijkbaar bewuste neiging om realiteitsgebon-

den kennis uit te sluiten bij het modelleren en oplossen van toepassingsopgaven die in een schoolse context worden aangeboden. Men kan zich hierbij echter de vraag stellen of deze alarmerende resultaten niet mede in de hand gewerkt zijn door bepaalde aspecten van de onderzoekssituatie waarin ze werden verkregen. In de volgende twee paragrafen rapporteren we twee verschillende reeksen van onderzoekingen waarin pogingen ondernomen zijn om het aantal RRs van de leerlingen op de P-items uit de onderzoekingen van Greer (1993) en Verschaffel e.a. (1994) te verhogen via bepaalde ingrepen in de experimentele setting.

3 De relatieve ineffectiviteit van 'scaffolds' gericht op het verhogen van de alertheid van leerlingen tijdens het oplossen van P-items

In deze paragraaf bespreken we de opzet en de belangrijkste resultaten van een eerste reeks vervolgonderzoekingen die verder gaan dan het puur constateren van de neiging van leerlingen tot niet-realistisch wiskundig modelleren. Kenmerkend voor deze studies is het doorvoeren van minimale ingrepen in de experimentele setting die leerlingen ervoor waarschuwen dat de probleemopgaven speciale aandacht vergen of die de leerlingen sterker uitnodigen tot het geven van een realistische reactie.

De studie van Yoshida, Verschaffel en De Corte (1997) was in eerste instantie een replicatie van het onderzoek van Verschaffel e.a. (1994) bij Japanse vijfdeklassers (zie paragraaf 2). Maar daarnaast had deze studie tot doel het effect na te gaan van een beperkte vorm van hulp: terwijl de ene helft van de Japanse leerlingen ($n = 46$) dezelfde neutrale instructie kreeg als in de studie van Verschaffel e.a. (1994), werd de andere helft ($n = 45$) expliciet gewaarschuwd dat de test meerdere problematische items bevatte en uitdrukkelijk uitgenodigd om deze problematische items aan te stippen en uit te leggen waarom ze niet of moeilijk (eenduidig) op te lossen waren. Terwijl in de neutrale conditie 15% van alle reacties als

realistisch te beschouwen waren (zie ook Tabel 2), werden in de waarschuwingsconditie 20% RRs teruggevonden. Uit de statistische analyse bleek dat het verschil tussen deze gemiddelde percentages niet significant was, hetgeen betekent dat de extra waarschuwing geen betekenisvol effect had op de neiging tot niet-realistisch antwoorden van deze Japanse leerlingen.

De experimentele setting uit de aanvankelijke studies werd ook gewijzigd in een niet-gepubliceerde studie van Greer en Caldwell (voor een kort verslag zie Greer, 1997) met 13-14 jarigen in Noord-Ierland, waarvoor drie toetsen ontwikkeld werden – respectievelijk ‘Mathematics test’, ‘Estimation test’ en ‘Mathematical puzzles’ genaamd – waarvan elke leerling er één kreeg. De toets met als titel ‘Mathematics test’ bestond uit verscheidene S- en P-items uit de aanvankelijke studie van Greer (1993), die aan de leerlingen werden aangeboden als een wiskundetest, zonder extra instructies. De ‘Estimation test’ bevatte dezelfde P-items samen met drie echte schatopgaven. In de toets getiteld ‘Mathematical puzzles’ tenslotte kwamen eveneens dezelfde P-items voor, maar ditmaal waren ze vergezeld van drie complexe wiskundige raadsels. De verwachting was dat de leerlingen die met de laatste twee toetsen geconfronteerd zouden worden, meer op hun tellen zouden passen en (dus) op de P-items meer RRs zouden produceren. Doch deze alternatieve aanbiederwijzen leidden niet tot een betekenisvolle toename in de proportie RRs. Uit de gesprekken die de onderzoekers achteraf met de leerlingen voerden kwam naar voren dat velen er (uitdrukkelijk) van uitgingen dat er een onderscheid bestaat tussen het soort van ‘conventionele’ antwoorden dat van leerlingen verwacht wordt tijdens de wiskundelessen op school, en de manier waarop levensechte probleemsituaties beantwoord worden. Zo zei een 10-jarige in antwoord op de vraag van de interviewer waarom hij geen realiteitsgebonden kennis had ingeschakeld bij het oplossen van de P-items uit de test getiteld ‘Mathematics test’: ‘I know all these things, but I would never think to include them in a maths problem. Math isn’t about things like that. It’s about getting sums right and you don’t need to know outside things to get sums right’.

Van Lieshout, Verdwaald en Van Herk (1997) vergeleken de antwoorden van leerlingen uit twee klassen van het gewoon basisonderwijs en drie klassen voor bijzonder onderwijs op een schriftelijke toets bestaande uit tien S-items en tien P-items, waarvan de ene helft afkomstig was uit de studie van Verschaffel e.a. (1994) (nl. P1, P4, P5, P6 en P9 uit Tabel 1) en de andere door de auteurs zelf geconstrueerd was. Voor het begin van de test kregen de leerlingen een mondelinge, klassikale instructie waarin twee voorbeeldopgaven (waaronder één P-item!) besproken werden; daarbij beklemtoonde de leerkracht dat bij sommige opgaven verschillende correcte antwoorden (zelfs antwoorden zoals ‘Dit is een onoplosbare opgave’) acceptabel waren. Hoewel de resultaten in deze studie wat hoger lagen dan in die van Verschaffel e.a. (1994) – wellicht omwille van de introductie – bleef het aantal NRs overweldigend: meer dan de helft van de P-items lokte minder dan 20% RRs uit; het school-item (P6 uit Tabel 1) en het vrienden-item (P1) bleken het moeilijkst. Merkwaardig ten slotte is dat de leerlingen van het buitengewoon onderwijs significant vaker de P-items met RRs beantwoordden (42%) dan de leerlingen uit het regulier onderwijs (29%).

Gelijkaardige resultaten werden ook nog verkregen in twee experimenten van Reusser en Stebler (1997b) bij Zwitserse leerlingen, waarvan hier enkel het tweede, meest omvangrijke en meest systematische besproken wordt. Reusser en Stebler (1997b) lieten de tien P-items uit Verschaffel e.a. (1994) oplossen door 439 leerlingen van het eerste jaar secundair onderwijs. De experimentele condities waaronder de opgaven werden opgelost, varieerden evenwel in de mate waarin de leerlingen gewaarschuwd werden voor het mogelijks problematisch karakter van de opgaven.

- In conditie 1 werd een identieke instructie gegeven als in het onderzoek van Verschaffel e.a. (1994).
- In conditie 1a werd aan vier P-items (nl. het looper-item, het school-item, het touw-item en het vaas-item) een zin toegevoegd bedoeld om de leerlingen alert te maken voor een eventuele moeilijkheid met deze opgaven (bijv. ‘Maak een tekening voor je de volgende opgave oplost’ bij het school-

item, of 'Bestudeer de figuur aandachtig' bij het vaas-item).

- In conditie 2 werd na elk item expliciet gevraagd om de kwaliteit van de betreffende opgave (duidelijkheid van het item, oplosbaarheid,...) te beoordelen.
- In conditie 3 werd, naast de vragen uit conditie 2, aan het begin van de toets ook de volgende algemene waarschuwing opgenomen: 'Wees alert. Sommige opgaven zijn niet zo gemakkelijk als ze er op het eerste gezicht uitzien. Er zijn er zelfs waarvoor het niet zeker is of ze wel oplosbaar zijn'.

Hoewel het totaal aantal RRs (44%) in deze studie meer dan dubbel zo hoog lag dan in de studie van Verschaffel e.a. (1994) en in hun eigen replicatieonderzoek (zie Tabel 2), bleven nog steeds meer dan 50% van de leerling-reacties ondubbelzinnig niet-realistisch. Een variantieanalyse toonde aan dat de percentages RRs van de testcondities 1, 2 en 3 niet significant van elkaar verschilden (respectievelijk 42,5, 38,6 en 41,1%). Er werd echter wel een significante toename van RRs geobserveerd in conditie 1a: bij de vier P-items waaraan in deze conditie een specifieke waarschuwende zin was toegevoegd, lag het aantal RRs een stuk hoger. In tegenstelling tot een expliciete algemene waarschuwing over het dubbelzinnig of zelfs onoplosbaar karakter van de opgaven, leidden itemspecifieke waarschuwingen dus wel tot een matige stijging in het aantal RRs.

Tenslotte onderzocht Lasure (1995; zie De Corte, Verschaffel, Lasure, Borghart & Yoshida, in press; Verschaffel, 1996) de hardnekkigheid van de neiging van leerlingen uit de bovenbouw van de basisschool tot routinematig, niet-realistisch modelleren in een studie die uit twee delen bestond. In een eerste deel werden zeven itemparen uit Verschaffel e.a. (1994) (met de volgende P-items: het vrienden-item, het planken-item, het bus-item, het looper-item, het school-item, het touw-item en het vaas-item uit Tabel 1) schriftelijk aangeboden aan 64 leerlingen van het vijfde leerjaar. Zestien procent van alle reacties van de leerlingen op deze toets waren RRs, wat vrijwel gelijk is aan het resultaat uit de eerste studie (Verschaffel e.a., 1994; zie paragraaf 1) en geheel in de lijn ligt van de gemiddelde percentages RRs uit de in paragraaf 2 besproken replicatiestudies. Op

basis van de resultaten op deze toets werden de 15 meest 'realistische' en de 15 minst 'realistische' probleemoplossers geselecteerd om aan het tweede deel van het onderzoek deel te nemen. Daarin werden bij deze leerlingen individueel één of twee 'scaffolds' toegepast gericht op het realistisch modelleren van de P-items die zij oorspronkelijk op niet-realistische wijze aangepakt en beantwoord hadden. Bij de eerste 'scaffold' werden de leerlingen die op een bepaald P-item met een NR gereageerd hadden, geconfronteerd met de RR op datzelfde item gegeven door een fictieve klasgenoot (bijv. aan een leerling die oorspronkelijk geantwoord had met $4 \times 2,5 = 10$ planken' op het planken-item zei de interviewer: 'Zoals je op dit blad kunt zien, antwoordde één van je klasgenoten met $4 \times 2 = 8$ planken. Wat is het beste antwoord? Waarom?'). Indien de leerling bij zijn aanvankelijke NR bleef, werd vervolgens een tweede 'scaffold' gegeven, die bestond uit een expliciete 'hint' vanwege de interviewer om de probleemsituatie te verpersoonlijken, te concretiseren of te tekenen (bijv. 'Kan je de planken tekenen? Kan je ook tekenen wat er met deze planken in het vraagstuk gebeurt? Kan je op je tekening zien hoeveel planken van 1 m Steven kan zagen uit deze vier planken?'). Hoewel deze 'scaffolds' in een significant aantal gevallen de leerlingen hun NR voor een RR deed inruilen (het percentage RRs steeg van 23% tijdens de schriftelijke test tot 40% na de eerste vorm van hulp en tot 57% na de tweede 'scaffold'), bleef het totaal aantal RRs op het eind van de individuele interviews nog altijd verrassend laag. Immers, in bijna de helft van de gevallen (nl. 43%) bleven de leerlingen uiteindelijk vasthouden aan hun initieel onrealistisch antwoord, zelfs na de tweede sterke 'hint' in de richting van het realistisch antwoord. Hieruit kan geconcludeerd worden dat de neiging van leerlingen uit de bovenbouw van de basisschool tot niet-realistisch modelleren en oplossen van schoolvraagstukken zeer sterk en persistent is.

Globaal genomen laat deze eerste reeks van studies naar het effect van ingrepen in de experimentele setting zien dat expliciete, algemene waarschuwingen aan de leerlingen omtrent het dubbelzinnig of zelfs onoplosbaar karakter van de P-opgaven, geen betekenisvolle stijging in

het aantal RRs teweegbrengen. Daaruit blijkt de persistentie van de neiging van veel leerlingen om contextgebonden wiskundige toepassingsopgaven op een stereotiepe, niet-realistische manier te modelleren en op te lossen. Alleen heel nadrukkelijke, itemspecifieke aanvullingen of waarschuwingen resulteerden in een behoorlijke stijging van het aantal RRs, maar zelfs dan was deze stijging nog aan de lage kant.

4 De relatieve effectiviteit van het verhogen van de authenticiteit van de experimentele setting

In deze paragraaf bespreken we onderzoek naar het effect van het wijzigen van een ander aspect van de experimentele setting waarmee gewerkt is in de aanvankelijke studies van Greer (1993) en Verschaffel e.a. (1994), nl. de authenticiteit van de probleemsituaties waarmee de leerlingen geconfronteerd worden.

DeFranco en Curcio (1997) gingen uit van de vaak geconstateerde moeilijkheid van leerlingen om de uitkomst van een niet-opgaande deling, zoals het bus-item uit Tabel 1, adequaat te interpreteren (zie o.a. Greer, 1993; Silver, Shapiro & Deutsch, 1993; Verschaffel e.a., 1994) (zie paragraaf 2). Zij onderzochten hoe leerlingen de rest van een deling interpreteren bij toepassingsopgaven die in twee verschillende experimentele condities waren ingebed, nl. (1) een restrictieve schoolse en (2) een realistische buitenschoolse setting. In deel 1 van deze studie werden 20 zesdeklassers geconfronteerd met een variant van het bus-item in een schoolse context; in een individueel interview kregen ze een aantal vraagstukken aangeboden evenals enkele vragen over wiskundig probleemoplossen op school. In deel 2 van het onderzoek werden dezelfde 20 leerlingen gevraagd om per telefoon minibussen te bestellen om de leerlingen van hun klas naar een schoolfeest te vervoeren. Terwijl het bus-item in de schoolse context slechts door twee van de 20 leerlingen juist beantwoord werd, bedroeg het aantal correcte antwoorden in de meer realistische setting 16. Hoewel dit onder-

zoek slechts slaat op één type van P-items uit de toets van Greer (1993) en Verschaffel e.a. (1994) en op een gering aantal subjecten, leveren de resultaten toch sterke steun voor de hypothese dat het inbedden van deze P-opgaven in een (meer) authentieke setting het aantal RRs betekenisvol doet toenemen.

Uitgaande van een socioculturele analyse van de specifieke kenmerken van de experimentele setting in de studies van Greer (1993) en Verschaffel e.a. (1994), meenden Wyndhamn en Säljö (1997) dat door leerlingen de opgaven te laten oplossen in het kader van betekenisvolle en communicatieve projecten met anderen, het aantal RRs sterk zou toenemen. Om deze hypothese te toetsen, zetten zij een empirische studie op rond het school-item uit de toets van Verschaffel e.a. (1994) en rond nog een tweede opgave omtrent afstanden:

- Anna en Berra gaan naar dezelfde school. Anna woont op 500 meter van de school en Berra woont op 300 meter van de school. Hoever wonen ze van elkaar? (= school-item)
- Wat is de afstand tussen Alstad en Broby volgens deze wegwijzers? (= wegwijzer-item).

Deze items werden evenwel niet als klassieke schoolvraagstukken aangeboden, maar werden geïntroduceerd in een gesprek tussen telkens drie leerlingen en een onderzoeker. Het gesprek waarin de eerste opgave ingebed zat, ging over de weg die de leerlingen aflegden van en naar de school; terwijl het probleem werd geformuleerd, werd een fictieve kaart van een school met haar omgeving getoond. De tweede taak maakte deel uit van een gesprek over wegwijzers: hoe ze eruit zien, waar je ze kan vinden, hoe ze geïnterpreteerd dienen te worden, enz... Bij deze opgave hoorde een figuur met twee wegwijzers waarvan de ene de afstand tot Broby (17 km) in de ene richting aangeeft en de andere de afstand tot Alstad (8 km) in de tegenovergestelde richting. De eerste taak werd aan acht homogene groepen van 10-11-jarigen voorgelegd; de tweede taak aan zes groepen van 12-jarigen. In de loop van het gesprek kwamen alle groepen – behalve één – tot de conclusie dat het onmogelijk is één antwoord te geven, maar dat 'it all depends'. Het enige groepje dat niet tot deze conclusie

kwam bestond uit drie sterke leerlingen die uiteindelijk beslisten om toch maar een getalsmatig antwoord te geven op het school-item, hoewel ze zich tijdens het groepsgesprek gerealiseerd hadden dat er in feite meer dan één correct antwoord mogelijk was. Deze bevindingen contrasteren scherp met de resultaten van de oorspronkelijke studie van Verschaffel e.a. (1994) en de replicatieonderzoekingen (zie paragraaf 1 en 2), waarin voor het school-item weinig of geen RRs aangetroffen werden. De studie van Wyndhamn en Säljö (1997) toont dus aan dat het maken van realistische overwegingen en het gebruik daarvan bij het modelleren en oplossen van wiskundige toepassingsopgaven binnen het bereik ligt van de meeste leerlingen. Het volstaat de problemen in te bedenken in een andere, meer betekenisvolle en communicatieve context om hen tot meer realistisch en doordacht probleemoplossen te brengen.

Overigens hebben deze auteurs ook nog in andere studies en met andersoortige opgaven empirische evidentie verzameld die steun biedt aan deze conclusie (Säljö & Wyndhamn, 1993; Wyndhamn en Säljö, 1995). Zo voerden zij reeds in 1993 een experiment uit met 211 leerlingen van 15 à 16 jaar die individueel en schriftelijk het probleem 'Hoeveel kost het om een brief van 120 gram te verzenden naar een binnenlands adres?' moesten oplossen in een wiskundeles of in een les sociale wetenschappen. In beide settings kregen de leerlingen een tabel met de officiële posttarieven voor binnenlandse zendingen met een maximumgewicht van 20, 100, 250, 500 en 1000 gram. Zoals verwacht, werd deze opgave in de twee situaties verschillend opgelost. In de les sociale wetenschappen kwam de overgrote meerderheid (nl. 71%) van de oplossingen tot stand via het 'gewoon' aflezen van de uitkomst in de tabel; slechts een minderheid (29%) bepaalde het tarief via allerlei rekenkundige bewerkingen (bijvoorbeeld: door de tarieven voor een brief van 20 gram en een brief van 100 gram op te tellen, of via het vermenigvuldigen van het tarief voor een brief van 20 gram met 6). In de wiskundeles daarentegen interpreteerde niet minder dan 57% van de leerlingen deze taak als een 'wiskundeopgave' en voerde dus een rekenkundige bewerking uit om tot het ant-

woord te komen, in plaats van het juiste tarief 'gewoon' uit de tabel af te lezen.

Tenslotte zetten Reusser en Stebler (1997b) een onderzoek op waarin ze de oplossingen van 34 leerlingen uit het zesde leerjaar of uit het eerste jaar secundair onderwijs op P-items analyseerden in twee verschillende contexten, nl. (1) dezelfde schoolse context als in de studies van Greer (1993) en Verschaffel e.a. (1994) en hun eigen replicatie- en vervolgonderzoek (Reusser & Stebler, 1997a) dat in paragraaf 2 en 3 voorgesteld werd, en (2) een context waarin dezelfde P-items werden aangeboden onder de vorm van zgn. performantietaken ('performance tasks'). De hypothese van de auteurs was dat de aanbestedingswijze van de taak en de experimentele context waarin de leerlingen deze taak dienden op te lossen, de belangrijkste oorzaken vormden van de vele NRs uit de studies van Greer (1993), Verschaffel e.a. (1994) en hun eigen vroeger werk (Reusser & Stebler, 1997a). Op grond daarvan voorspelden zij dat leerlingen veel minder NRs zouden geven in deze performantiecontext dan in de schoolse context. Om deze hypothese te toetsen, vormden Reusser en Stebler (1997b) vijf van de tien P-items uit de toets van Verschaffel e.a. (1994) om tot performantietaken: het planken-item, het loper-item, het school-item, het touwen-item en het vaas-item. Dit hield in dat deze opgaven in een meer authentieke context werden aangeboden door concreet materiaal te voorzien (zoals bijvoorbeeld planken, een zaag en een vouwmetre voor het planken-item); tevens werd een duidelijke op handelen gerichte instructie gegeven bestaande uit de volgende stappen: (1) bestudeer de opgave en het concreet materiaal zorgvuldig, (2) maak een voorspelling van het antwoord, (3) voer de taak effectief uit en geef intussen commentaar bij wat je aan het doen bent, (4) beslis of je al dan niet bij je oorspronkelijke voorspelling blijft, (5) schrijf je uiteindelijk antwoord neer. De eerste dag loste elke leerling de vijf P-items op in de schriftelijke, schoolse context, en de volgende dag in de performantiesetting. Zoals verwacht, vonden Reusser en Stebler (1997b) een significante stijging in het percentage RRs van de schoolse conditie naar de performantieconditie: voor het planken-item, het loper-item, het touwen-item, en het vaas-item steeg het percentage RRs respectievelijk van 25 naar 56%,

van 12 naar 47%, van 18 naar 62% en van 7 naar 40%; alleen bij het school-item was de stijging niet significant (van 21% naar 36%). De analyse van de probleemoplossingsprocessen van de leerlingen bracht aan het licht dat zowel de aanwezigheid van concreet materiaal als de effectieve manipulatie ervan tijdens het opvolgen van de performantie-instructies had bijgedragen tot de vastgestelde stijging in het aantal RRs. Vermeldenswaard is ook dat enkele leerlingen die tijdens de praktische uitvoeringsfase realistisch geantwoord hadden, toch nog op een niet-realistisch antwoord terugvielen wanneer ze hun uiteindelijke oplossing op het antwoordblad moesten neerschrijven. Hoe dan ook, de betere prestaties van de leerlingen in de performantiesetting vormden een overtuigende bevestiging van de hypothese.

Samenvattend kunnen we stellen dat de onderzoeken die in deze paragraaf besproken werden, aantonen dat het verhogen van de authenticiteit van de context waarbinnen de P-items worden aangeboden (door ze in te bedden in een groepsdiscussie over een maatschappelijk relevant onderwerp of door concreet materiaal en een bijhorende performantie-opdracht te voorzien) een veel hogere stijging van het aantal RRs tot gevolg heeft dan de ingrepen in de experimentele setting die in paragraaf 3 besproken werden.

5 De rol van instructievariabelen op het niet-realistisch modelleren van wiskundige toepassingsopgaven

De resultaten van de besproken studies – en met name van de onderzoeken besproken in paragraaf 4 – suggereren dat de neiging van leerlingen om schoolvraagstukken op een niet-realistische manier aan te pakken niet voortvloeit uit een of ander ‘cognitief deficit’ bij deze leerlingen, maar veeleer het gevolg zijn van het feit dat zij zich in hun oplossingsgedrag – bewust maar vaak ook onbewust – laten leiden door een ‘didactisch contract’ (Brousseau, 1990). In de context van het wiskundeonder-

wijs verwijst deze term naar een samenhangend stel sociale en socio-mathematische normen (Cobb, Yackel & Wood, 1992; Gravemeijer, 1994, 1997) die in de klas heersen en de interacties tussen de leerkracht en de leerling(en) reguleren (veelal zonder dat de betrokkenen zich daarvan bewust zijn). Dit ‘didactisch contract’ schrijft dus voor hoe de leerlingen zich moeten gedragen, hoe ze moeten denken en praten tijdens deze lessen, hoe ze op een taak, een vraag, of een tussenkomst van de leerkrachten dienen te reageren... Voor het curriculumonderdeel vraagstukken houdt dit ‘didactisch contract’ o.a. de volgende (impliciete) afspraken en spelregels in: (a) een vraagstuk dient met een exact getalsmatig antwoord beantwoord te worden; (b) een vraagstuk kan op één en slechts één manier correct opgelost worden; (c) de correcte oplossing is te vinden door één of meer formeel-rekenkundige bewerkingen uit te voeren op de getallen die expliciet in de tekst gegeven staan; (d) het is niet toegelaten om de gegevens of het gevraagde ter discussie te stellen, noch om bij het oplossen van de opgave rekening te houden met contextueel relevante informatie die niet uitdrukkelijk in de probleemformulering ‘gegeven’ is... (De Corte & Verschaffel, 1985; Gerofsky, 1996; Gravemeijer, 1994, 1997; Greer, 1997; Lave, 1992; Reusser, 1988; Schoenfeld, 1991; Treffers & De Moor, 1990; Wyndhamn & Säljö, 1997). De gerapporteerde onderzoeksresultaten suggereren dat dergelijke afspraken en spelregels het oplossingsgedrag van de leerlingen in de context van het wiskundeonderwijs in sterke mate bepalen. Door de P-items in een andere socioculturele context te plaatsen waarin deze afspraken en regels niet langer gelden – bijvoorbeeld door de items om te vormen tot uitdagende doetaken of door ze in te bedden in een betekenisvolle communicatieve context – kan men de meeste leerlingen er evenwel toe brengen om contextgebonden wiskunde problemen op een meer realistische manier te modelleren en op te lossen.

Dit brengt ons bij de vraag welke aspecten van het (vraagstukken)onderwijs er voor zorgen dat de leerlingen zich in hun aanpak- en oplossingsgedrag door deze afspraken en spelregels laten leiden. We mogen er immers van uitgaan dat ze – evenals andere onderdelen van het

'didactisch contract' (Brousseau, 1990) of als andere socio-mathematische normen (Cobb e.a., 1992; Gravemeijer, 1994; Yackel & Cobb, 1996) – niet *expliciet* op school worden onderwezen noch *bewust* in stand worden gehouden, maar eerder tot ontwikkeling komen en versterkt worden via allerlei impliciete, niet-bedoelde aspecten van de dagelijkse praktijk van het wiskundeonderwijs. Volgens diverse auteurs is het ontstaan en de ontwikkeling ervan toe te schrijven aan de volgende twee kenmerken van het huidige wiskundeonderwijs: (a) het levensvreemd en stereotiep karakter van de overgrote meerderheid van de vraagstukken waarmee leerlingen (onder meer via de rekenmethoden en tests) worden geconfronteerd, en (b) de manier waarop tijdens de vraagstuklessen met deze opgaven wordt omgesprongen, en vooral de geringe aandacht die leerkrachten besteden aan de kritische verkenning van de complexe, en vaak enigszins problematische, relatie tussen de probleemsituatie waarnaar de opgavetekst verwijst enerzijds en het voorgestelde wiskundig model daarvan anderzijds (Davis, 1989; De Corte & Verschaffel, 1985; Gerofsky, 1996; Gravemeijer, 1994, 1997; Greer, 1993, 1997; Lave, 1992; Reusser, 1988; Schoenfeld, 1991; Treffers & De Moor, 1990; Verschaffel e.a., 1994; Verschaffel, in press). Door deze aspecten van de klassieke vraagstukkendidactiek zou bij leerlingen langzaam maar zeker de dispositie groeien om de concrete situatie waarin een wiskundig toepassingsprobleem ingebed zit, niet al te ernstig te nemen en om de eigen ervaringskennis daaromtrent buitenspel te zetten. In de onderzoeksliteratuur is er evenwel weinig of geen overtuigende rechtstreekse empirische evidentie voorhanden voor het bestaan van zo'n didactisch contract over het uitsluiten van realistische overwegingen in het vigerend (vraagstukken)onderwijs, noch voor de band tussen de aard en de sterkte van deze socio-mathematische norm enerzijds en de neiging tot niet-realistisch modelleren bij de leerlingen anderzijds.

Een uitzondering hierop vormt een beperkte deelstudie die Reusser en Stebler (1997a) koppelde aan hun onderzoek dat in paragraaf 3 besproken werd. Op het einde van de toetsboekjes met S- en P-items voegden deze onder-

zoekers ook nog een beperkte vragenlijst toe, waarin de leerlingen gevraagd werd om aan te duiden hoe vaak ze in de loop van de voorbije jaren in de wiskundelessen geconfronteerd waren met (a) onoplosbare vraagstukken, (b) open vraagstukken, (c) vraagstukken met meerdere correcte oplossingen en (d) vraagstukken waarin er niet gevraagd wordt naar een precies maar naar een benaderend antwoord. Reusser en Stebler (1997a) vonden geen significant effect van de vertrouwdeheid van de leerlingen met opgaven van type (c) en (d) op de proportie RRs op de P-items uit de test, maar voor de eerste twee categorieën van problematische opgaven werd er wel een sterke, positieve samenhang geconstateerd tussen de gerapporteerde frequentie waarmee men tijdens de wiskundelessen met dergelijke opgaven geconfronteerd was en de mate waarin men de P-items op een realistische wijze wist te modelleren en op te lossen.

Indirecte aanwijzingen zijn er ook te vinden in een ouder onderzoek van Radatz (1983), waarin gevonden werd dat het aantal leerlingen dat absurde vraagstukken zoals het bekende 'captain's problem' ('Op een schip zijn er 26 schapen en 10 geiten. Hoe oud is de kapitein?') beantwoordde met de uitkomst van een bewering op de twee gegeven getallen (bijv. ' $26 + 10 = 36$ jaar'), toenam met het aantal jaren genoten rekenonderwijs.

Een derde studie die enig licht werpt op de rol van instructievariabelen bij het tot stand komen van de neiging tot niet-realistisch modelleren bij leerlingen, is het onderzoek van Verschaffel, De Corte en Borghart (1997) waarin de opvattingen en houdingen (tegen)over realistisch wiskundig modelleren zijn nagegaan bij aspirant-leerkrachten. Een test bestaande uit 14 vraagstukken – waarvan de helft S-items en de helft P-items afkomstig uit de studie van Verschaffel e.a. (1994) – werd tweemaal aangeboden aan 228 eerstejaarsstudenten en 104 derdejaarsstudenten uit drie Vlaamse opleidingsinstituten voor toekomstige leerkrachten basisonderwijs. Tijdens een eerste sessie werden de aspirant-leerkrachten gevraagd om de opgaven zelf op te lossen (= test 1); in een tweede sessie dienden ze vier verschillende antwoorden van (denkbeeldige) leerlingen op

De afstand tussen twee palen bedraagt 12 meter. Hoeveel stukken touw van 1,5 meter lang moet je aan elkaar knopen om de beide palen te verbinden?*

$12 : 1,5 = 6$ Je hebt 6 stukken van 1,5 m nodig.	A
$12 : 1,5 = 8$ Je hebt 8 stukken van 1,5 m nodig.	B
$12 \times 1,5 = 18$ Je hebt 18 stukken van 1,5 m nodig.	C
Dat kan je niet precies weten.	D

* De gegeven score (1, 1/2 of 0) diende telkens in het bijpassend hokje rechtsonder elke oplossing te worden genoteerd (A, B, C en D). In deze opgave geldt oplossing B als het niet-realistisch antwoord en oplossing D als het (meest) realistische.

Figuur 1. Voorbeeld van een opgave uit test 1 van Verschaffel e.a. (1997)

dezelfde opgaven (namelijk een realistisch antwoord, een niet-realistisch antwoord, een rekentechnische fout en een verkeerde operationele fout) te evalueren (= test 2). Gevraagd werd om elk van deze vier antwoordalternatieven met 1 punt te scoren indien ze het antwoord correct achtten, met 0 punten indien ze het antwoord helemaal fout achtten en met 1/2 punt indien ze het antwoord deels correct en deels fout inschatten. Een voorbeeld van een van de items uit test 2 wordt gegeven in Figuur 1.

De resultaten wezen op een sterke neiging bij aspirant-leerkrachten – zelfs bij de derdejaarsstudenten – tot het negeren van hun realiteitsgebonden kennis. Dit bleek in de eerste plaats uit hun eigen spontane oplossingen van de P-items (in test 1), waaronder slechts 48% RRs werden aangetroffen, maar ook uit hun beoordelingen van de antwoorden van (denkbeeldige) leerlingen (test 2). Immers, de NRs werden door de aspirant-leerkrachten over het algemeen beduidend beter gescoord dan de RRs. Het gemiddelde percentage 1- 1/2- en 0-scores dat op de zeven P-items aan het realistische antwoordalternatief werd gegeven, bedroeg respectievelijk 47%, 6% en 47%; voor het niet-realistische alternatief bedroeg het percentage 1-, 1/2- en 0-scores respectievelijk 56%, 26% en 18. Niet-realistische antwoorden kregen dus

in het algemeen een hogere score van de aspirant-leerkrachten dan realistische! Weliswaar werd er een significant verschil in het totaal aantal RRs op test 1 vastgesteld tussen de eerstejaars- (45% RRs) en de derdejaarsstudenten (54% RRs) in het voordeel van de laatste groep, maar – zoals de cijfers laten zien – bleef het percentage RRs in deze laatste groep aan de lage kant. Ook in test 2 gaven de derdejaarsstudenten iets vaker een hoge score voor de RR en een lage score voor de NR dan de eerstejaarsstudenten. Tenslotte vermelden we nog dat in de schriftelijke verantwoordingen die de aspirant-leerkrachten voor hun lage scores op de realistische antwoorden uit test 2 gaven, geregeld een negatieve appreciatie voor het activeren van ervaringskennis over de opgavencontext doorklonk. We erkennen dat deze gegevens nog niet overtuigend aantonen dat de vastgestelde vakgebonden opvattingen en houdingen bij (aspirant-)leerkrachten rechtstreeks verantwoordelijk zijn voor de neiging van leerlingen om hun ervaringskennis uit te sluiten tijdens het modelleren en oplossen van schoolvraagstukken. Maar wanneer we rekening houden met de resultaten van recent onderzoek over de invloed van vakkennis en vakgebonden 'beliefs' en houdingen van leerkrachten op het denken en leren van leerlingen (zie o.m. De Corte, Greer & Verschaffel, 1996), is het erg

aannemelijk dat deze cognities, overtuigingen en attitudes van leerkrachten omtrent het belang van realiteitsgebonden kennis bij het interpreteren en oplossen van contextproblemen, een beslissende invloed uitoefenen op het modellerings- en oplossingsgedrag van hun leerlingen. Verder onderzoek hieromtrent is evenwel vereist.

6 Effect van aangepaste instructie op (leren) realistisch modelleren en interpreteren

In een exploratief onderwijsexperiment gingen Verschaffel en De Corte (1996) na of het mogelijk is om in een klascontext betekenisvolle positieve leereffecten te bereiken op het vlak van het realistisch modelleren en interpreteren van wiskundige toepassingsproblemen. Daartoe werd een alternatieve onderwijsleeromgeving aangeboden waarin vraagstukken uitdrukkelijk opgevat worden als oefeningen in realistisch modelleren en interpreteren. De leergang werd gerealiseerd in een vijfde leerjaar van een Vlaamse basisschool en bestond uit vijf leereenheden van ongeveer 2,5 uur die stuk voor stuk inzoomden op één type van toepassingssituaties waarbij er vanuit realistisch oogpunt een moeilijkheid rijst bij het oplossen van de probleemsituatie via de voor-de-hand-liggende rekenkundige bewerking(en). Bijvoorbeeld: vraagstukken over verdeel- of uitdeelsituaties waarbij de uitkomst van de deling nog geïnterpreteerd dient te worden in functie van de context, vraagstukken waarbij de context een oplossing gebaseerd op een lineair-proportioneel model uitsluit, ... De didactische uitwerking van deze leereenheden was gesteund op de volgende drie pijlers: (1) werken met een set van uitdagende, realistische problemen; (2) gebruik maken van sterk interactieve instructiemethoden zoals groepswerk en klassikale discussies afgewisseld met individuele opdrachten; (3) creëren van een nieuwe klas-cultuur o.a. door met de leerlingen te onderhandelen over nieuwe normen en afspraken omtrent wat er verstaan wordt onder een goed vraagstuk, een goed antwoord, een goede oplossingsweg, enz. Terwijl de leerlingen van de experimentele klas dit programma door-

werkten, volgden twee controleklassen gewoon wiskundeonderwijs.

Van alle leerlingen uit de experimentele klas en de twee controleklassen werd een schriftelijke voortoets en een parallelle natoets afgenomen bestaande uit een tiental P-items waarvan de ene helft niet al te veel afweek van de opgaven waarmee in de leergang was gewerkt (= leer-items) terwijl de andere helft daar sterker van afweek (= nabije-transferitems). In een van de twee controleklassen werd de natoets voorafgegaan door een inleiding waarin de leerlingen aan de hand van twee voorbeelden van P-items expliciet getoond werd dat routinematige antwoorden soms inadequaat zijn vanuit realistisch oogpunt (zoals het geval was in sommige studies die in paragraaf 3 besproken werden). In de experimentele klas werd achteraf ook nog een retentietoets afgenomen, bestaande uit een aantal parallelitems van de opgaven uit de voor- en natoets, maar daarnaast ook uit een aantal items die nog sterker afweken van de opgaven uit de leergang (verre-transferitems).

Uit de analyse van de resultaten bleek een significant positief effect van het experimenteel programma op de neiging van de leerlingen tot realistisch modelleren: terwijl in de beide controleklassen slechts sprake was van een minimale, niet-significante stijging van het aantal RRs van voor- naar natoets, steeg dit percentage in de experimentele klas van 7 naar 51%. Hoewel de stijging groter was voor de leer-items dan voor de nabije-transferitems, was het effect in beide gevallen significant. Tenslotte suggereerden de resultaten op de retentietoets, die ook naar verre transfer van het geleerde peilde, dat het positief effect van het experimenteel programma specifiek noch tijdelijk was. Hoewel er een aantal methodologische tekorten aan deze studie kleven (bijv. het ontbreken van een retentietoets in de controleklassen) en het aantal RRs op de natoets en de retentietoets niet bij alle leerlingen in dezelfde mate toenam (de grootste leerwinst werd geboekt door de sterke leerlingen), suggereren de resultaten globaal genomen toch dat het mogelijk is om de neiging van basisschoolleerlingen tot niet-realistisch modeleren en interpreteren van rekenvraagstukken via aangepaste instructie op een betekenisvolle en blijvende wijze te doorbreken.

7 Samenvatting en discussie

In dit artikel hebben we een overzicht gegeven van een reeks nauw verwante recente studies op het gebied van het modelleren en oplossen van problematische wiskundige toepassingsproblemen bij 10-14-jarige leerlingen. Aanleiding voor de besproken onderzoeken vormden de studies van Greer (1993) en van Verschaffel e.a. (1994), waaruit bleek dat de overgrote meerderheid van de leerlingen een sterke neiging vertonen om wiskundige toepassingsopgaven die problematisch zijn vanuit realistisch oogpunt (zgn. P-items) op een stereotiepe, niet-realistische manier te modelleren en op te lossen. Meer concreet komt hun aanpak hierop neer dat zij 'gewoon' een of meerdere standaardbewerkingen uitvoeren met de getallen uit de opgave, zonder rekening te houden met hun ervaringskennis omtrent de reële, concrete situatie waarop het vraagstuk betrekking heeft. In een eerste reeks vervolgstudies, die in paragraaf 2 aan bod kwamen, werden de bevindingen van Greer (1993) en Verschaffel e.a. (1994) gerepliceerd bij verschillende andere leerlingpopulaties.

Om beter zicht te krijgen op de denkprocessen die ten grondslag liggen aan deze neiging van leerlingen tot het negeren van realiteitsgebonden kennis bij het oplossen van dit soort P-items, werden de laatste jaren verscheidene nieuwe onderzoeken opgezet, die we in paragraaf 3 en 4 besproken hebben. Een eerste reeks studies liet zien dat ingrepen in de experimentele setting in de vorm van verbale tussenkomsten die leerlingen er in algemene termen toe aan zetten om alert te zijn, om rekening te houden met aspecten van de realiteit, en/of om ook alternatieve antwoorden te overwegen, niet of nauwelijks leidden tot een daling van het aantal niet-realistische reacties. Enkel wanneer dergelijke algemene waarschuwingen aangevuld werden met of vervangen werden door meer itemspecifieke 'scaffolds', viel er een significante, maar nog steeds bedroevend matige, toename van het aantal realistische reacties te constateren. In een tweede reeks vervolgonderzoeken werden een of meerdere types van P-items aangeboden in een meer authentieke setting. Dit bleek een veel effectievere interventie te zijn want zij leidde

wél tot een betekenisvolle toename van het aantal realistische reacties.

Overigens troffen we in sommige studies ook uittreksels van interviews of van klasgesprekken aan waaruit exemplarisch bleek dat sommige leerlingen zich wel bewust zijn van het feit dat zij zich anders gedragen wanneer zij in de klascontext een rekenvraagstuk aangeboden krijgen dan wanneer ze in de buitenschoolse realiteit met een kwantitatief probleem worden geconfronteerd.

Volgens verscheidene auteurs dient de oorzaak van het ontstaan van deze neiging tot oppervlakkig, niet-realistisch modelleren bij leerlingen te worden toegeschreven aan de praktijk en de cultuur van het traditionele vraagstukkenonderwijs. Zoals uit paragraaf 5 gebleken is, is de empirische basis voor deze verklaring alsnog eerder gering. In zekere zin bieden de bemoedigende resultaten van het exploratief onderwijsexperiment van Verschaffel en De Corte (1996), dat in paragraaf 6 kort besproken werd, onrechtstreeks steun aan de stelling dat de oorzaak van de neiging van leerlingen om vraagstukken op een oppervlakkige, niet-realistische manier aan te pakken, te zoeken is in de vigerende vraagstukkendidactiek. Immers, dit onderwijsexperiment laat zien dat het mogelijk is om deze neiging te doorbreken door een alternatieve leeromgeving te creëren waarin vraagstukken opgevat worden als heuse oefeningen in realistisch modelleren en interpreteren, eerder dan als oefenstof voor het indrillen van stereotiepe oplossingsprocedures.

Tot slot van deze bijdrage plaatsen we nog enkele theoretische en methodologische kanttekeningen bij het onderzoek waarover in dit artikel gerapporteerd werd en staan we stil bij de implicaties ervan voor de praktijk van het vraagstukkenonderwijs op de basisschool.

Ten eerste, de oorzaak van de sterke neiging van leerlingen tot niet-realistisch antwoorden op de problematische items – maar ook van de grote verschillen in het effect van bepaalde ingrepen in de testsituatie of in het onderwijs – hebben we gezocht in het zgn. 'didactisch contract' (Brousseau, 1990) of, anders gezegd, in de heersende socio-mathematische normen

(Cobb e.a., 1992; Gravemeijer, 1994, 1997; Yackel & Cobb, 1996) in verband met 'het spel der schoolvraagstukken' (De Corte & Verschaffel, 1985), die het denken en handelen van leerlingen bepalen. Op een meer algemeen niveau wijzen de gerapporteerde onderzoeksresultaten op de sterke gesitueerdheid of contextafhankelijkheid van de menselijke cognitie, – de kerngedachte van de 'situated cognition'-benadering (Lave, 1992; Rogoff, 1984; Säljö & Wyndhamn, 1988). In tegenstelling tot de traditionele cognitieve psychologie, waarin de menselijke cognitie gezien werd als 'abstract, transcendental, culture-free, unemotional, universal, decontextualized, disembodied, and hence formal' (Lakoff & Nunez, 1997), wordt in het 'situated cognition'-paradigma precies gewezen op het feit dat denk- en leerprocessen zich voltrekken in nauwe interactie met de sociale en culturele context en via participatie aan situatie- en cultuurgebonden activiteiten. Of, zoals Rogoff (1984, p. 2) het uitdrukt: 'From this alternative epistemological stance, thinking is always intricately interwoven with the context of the problem to be solved, thus making context an integral aspect of cognitive events, not a nuisance variable'. De besproken onderzoeksresultaten laten inderdaad zien hoe (subtiele) aspecten van de context waarin leerlingen wiskundige problemen leren oplossen en/of van de setting waarin ze gevraagd worden om het geleerde te demonstreren, een doorslaggevende rol kunnen spelen in de wijze waarop deze opgaven worden opgevat, aangepakt en beantwoord. Tevens suggereren deze bevindingen dat de wiskundige kenniselementen en vaardigheden die leerlingen op school opdoen, net zoals de wiskunde die buitenschools verworven wordt, *gesitueerde* kennis is, of – zoals Lave (1992, p. 81) het uitdrukt – dat '(...) school math is situated practice. School is a site of specialized everyday activity – not a privileged site where universal knowledge is transmitted'. Bijkomend onderzoek is echter vereist om beter zicht te krijgen op de manier waarop diverse aspecten van de klas-cultuur rechtstreeks of onrechtstreeks bijdragen tot het ontstaan en de handhaving van de socio-mathematische normen in verband met het (leren) oplossen van schoolvraagstukken en met de rol die ervaringskennis en realistische overwegingen daarin spelen.

Ten tweede, hoewel we daar in dit literatuuroverzicht slechts sporadisch op ingegaan zijn, is het goed eraan te herinneren dat de gevonden proporties realistische reacties (RRs) sterk verschilden naargelang van het soort van P-item. Een rationele taakanalyse laat zien dat de P-items waarmee in deze studies gewerkt is, onmogelijk over een zelfde kam te scheren zijn. Allereerst zijn ze gekoppeld aan totaal verschillende wiskundige inhoud; zo komen bij de interpretatie van de uitkomst van een niet-opgaande deling (zie bijv. het bus- en het ballonprobleem) andere wiskundige begrippen en vaardigheden aan bod dan bij het nadenken over de aanwezigheid van een lineair-proportioneel verband (zoals in het loper- en het vaasprobleem). Verder verschilt ook de ervaringskennis waarop beroep wordt gedaan sterk van opgave tot opgave; zo lijkt het aannemelijk dat kinderen sneller uit ervaring weten dat er geen halve bussen bestaan (in het bus-item) dan dat ze beseffen dat het peil van een vloeistof sneller stijgt naarmate het vat smaller wordt (in het vaasprobleem). Een derde belangrijk verschilpunt betreft de aard van de verwachte realistische reactie: bij sommige P-opgaven, zoals het bus- en het planken-item, is dit een precies getal (respectievelijk de getallen 13 en 8); bij andere P-items daarentegen is het (meest) realistische antwoord geen exact getal, maar een benaderend antwoord (bijv. het loper-item), een aanduiding van een interval waarbinnen de uitkomst kan liggen (bijv. het school-item), of een opmerking vanwege de leerling dat de opgave op basis van de beschikbare gegevens niet op te lossen is (bijv. het vrienden-item). Het is evident dat het geven van een realistisch antwoord op deze laatste categorie van opgaven (nog) heel wat verder af staat van het soort antwoorden dat leerlingen gewoon zijn te geven in een traditionele vraagstukkenles dan bij het bus-item bijvoorbeeld. Anders gezegd, om bij de laatste categorie van P-items tot een realistische reactie te komen, dienen leerlingen meer elementen uit het 'didactisch contract' inzake schoolvraagstukken te overtreden dan bij het busprobleem (Hatano, 1997; Van Lieshout e.a., 1997; Verschaffel e.a., 1994; Wyndhamn & Säljö, 1997). Het is derhalve aangewezen om in voortgezet onderzoek te werken aan een theoretisch en empirisch gefundeerde typologie van P-items.

Ten derde is het gebruik van de termen 'standaardopgave' versus 'problematische opgave' (respectievelijk afgekort als S- en P-item) en 'realistische reactie' versus 'niet-realistische reactie' (respectievelijk afgekort als RR en NR) minder eenduidig dan we in het voorafgaande hebben laten uitschijnen. Immers, bij om het even welk wiskundig toepassingsprobleem worden er allerlei veronderstellingen gemaakt die, wanneer men daar zou van afzien, het opgebouwde wiskundig model op de helling zetten. Bij de zgn. standaardvariant van het planken-item (nl. 'Karel heeft 6 planken van 2 m gekocht. Hoeveel planken van 1 m kan hij hieruit zagen?') bijvoorbeeld kan met argumenteren dat wie deze opgave met $6 \times 2 = 12$ planken beantwoordt in feite ook een realistisch aspect over het hoofd ziet, namelijk het feit dat er door het zagen normaliter enkele millimeters hout verloren zullen gaan en dat de resulterende planken dus in feite iets korter dan een meter zullen uitvallen (Gravemeijer, 1997; Greer, 1997). Ook het door ons als realistisch getypeerde antwoord op het bus-item, nl. 13 bussen, is misschien minder realistisch dan op het eerste gezicht lijkt. Zo merkt Hatano (1997) op dat het in de realiteit heel goed mogelijk is dat wanneer 450 kinderen per bus vervoerd moeten worden, er uiteindelijk geen 13 maar toch slechts 12 bussen zullen worden ingezet. Misschien zullen er in elke bus enkele kinderen meer geplaatst worden, of misschien schakelt men naast de bussen ook nog enkele auto's in, enz. Ook Greer (1997) erkent het probleem van de 'fuziness of reality' en merkt in dit verband op dat de kinderen uit de gerapporteerde onderzoeken zeker niet de enigen zijn die tijdens hun wiskundige probleemoplossende activiteit bepaalde aspecten van de realiteit negeren; ook wiskundigen en andere wetenschappers doen dit voortdurend. Maar het belangrijk verschil tussen leerlingen uit die onderzoeken en de wiskunde-experts is, aldus Greer (1997, p. 304), dat 'mathematicians do it *knowingly* (and with regard to the context and goals, the likely levels of inaccuracy, and the technical knowledge and power available)'. Hij vindt het dan ook niet zinvol om oeverloze discussies op te zetten omtrent wat er in een concreet geval wel en niet als een RR en een NR moet worden beschouwd. Veel belangrijker is volgens hem 'the realisation that acts of modeling are taking

place, that different models are possible, and that these are open to debate' (p. 304).

Dit brengt ons tenslotte bij de vraag naar de onderwijspraktische implicaties van het besproken onderzoek. Het is duidelijk dat het geheel van deze studies een pleidooi inhoudt voor een wiskundendidactiek die sterk door-drongen is van het zgn. modelleringperspectief (Greer, 1997). Dit houdt eerst en vooral in dat de leerlingen geconfronteerd worden met een gevarieerder en authentieker opgavenaanbod, waardoor ze 'aan den lijve' ondervinden dat het oplossen van wiskundige toepassingsproblemen meer inhoudt dan het uitvoeren van een of enkele standaardoperaties op de gegeven getallen. Een tweede aspect van dit modelleringperspectief is dat de onderwijsleeractiviteiten die rond deze opgaven plaatsvinden erop gericht zijn leerlingen te laten nadenken en discussiëren over de voor- en nadelen van verschillende modellen, over de spanning tussen de eenvoud en de accuraatheid van een wiskundig model, en over de factoren die ertoe kunnen leiden dat de uitkomst van een wiskundig model niet (helemaal) strookt met de werkelijkheid. Ten derde houdt dit modelleringperspectief in dat de confrontatie van de leerlingen met dit rijke en levensnabije opgavenaanbod en met de bijhorende onderwijsleeractiviteiten reeds vroeg en geïntegreerd in het curriculum aan bod moet komen. Betekent dit alles nu dat er voor de eenvoudige, klassieke rekenvraagstukken helemaal geen plaats meer is weggelegd in het wiskundeonderwijs op de basisschool en dat leerlingen voortdurend gestimuleerd moeten worden om op zoek te gaan naar aspecten van de probleemsituatie die de toepasbaarheid van een onderliggend wiskundig model kunnen hypothekeer? Samen met de meeste auteurs die het modelleringperspectief onderschrijven (Gravemeijer, 1994, 1997; Greer, 1997; Hatano, 1997), menen we van niet. In het wiskundeonderwijs op de basisschool zullen er steeds momenten blijven bestaan waarin men leerlingen aan de hand van eenvoudige probleemsituaties bovenal inzicht wil laten verwerven in een bepaalde wiskundige bewerking of formule en/of enige vaardigheid wil laten ontwikkelen in het gebruik ervan. Wanneer contextopgaven vooral deze functie(s) vervullen, heeft het allicht weinig zin

om de aandacht van de leerlingen toe te spitsen op allerlei aspecten die de modelfunctie van de aangereikte situatie of de toepasbaarheid van de wiskundige bewerking of formule in het gedrag (kunnen) brengen. Immers, dat zou de zaak alleen maar compliceren, nodeloos instructietijd opslorpen, en de aandacht van de leerlingen afhouden van het eigenlijke lesdoel (zie ook Gravemeijer, 1997; Greer, 1997). Maar als we het erover eens zijn dat er een (beperkte) plaats voorbehouden moet blijven voor klassieke vraagstukken, is wel de belangrijke vraag aan de orde of, en zo ja hoe én wanneer, men leerlingen bewust moet maken van de verschillende functies die contextgebonden toepassingsopgaven in het wiskundeonderwijs kunnen vervullen en van de socio-mathematische normen die gelden bij die verschillende gebruikswijzen.

Literatuur

- Brousseau, G. (1990). Le contrat didactique: Le milieu. *Recherches en Didactique de Mathématiques*, 9(3), 308-336.
- Burkhardt, H. (1994). Mathematical applications in school curriculum. In T. Husén & T.N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed.) (pp. 3621-3624). Oxford, England: Pergamon Press.
- Caldwell, L. (1995). *Contextual considerations in the solution of children's multiplication and division word problems*. (Master's thesis). Belfast, Northern-Ireland: University of Belfast.
- Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education*, 23, 2-33.
- Davis, R.B. (1989). The culture of mathematics and the culture of schools. *Journal of Mathematical Behavior*, 8, 143-160.
- De Corte, E., Greer, B., & Verschaffel, L. (1996). Learning and teaching mathematics. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 491-549). New York: Macmillan.
- De Corte, E., & Verschaffel, L. (1985). Beginning first graders' initial representation of arithmetic word problems. *Journal of Mathematical Behavior*, 4, 3-21.
- De Corte, E., Verschaffel, L., Lasure, S., Borghart, I. & Yoshida, H. (in press). Real-world knowledge and mathematical problem-solving in upper primary school children. In J. Bliss, R. Säljö, & P. Light (Eds.), *Learning sites: Social and technological contexts for learning*. Oxford: Elsevier Science Ltd.
- DeFranco, T.C., & Curcio, F.R. (1997). A division problem with a remainder embedded across two contexts. Children's solutions in restrictive versus real-world settings. *Focus on Learning Problems in Mathematics*, 19(2), 58-72.
- Gerofsky, S. (1996). A linguistic and narrative view of word problems in mathematics education. *For the Learning of Mathematics*, 16(2), 36-45.
- Gravemeijer, K. (1994). *Developing realistic mathematics education*. Utrecht, The Netherlands: Freudenthal Institute, University of Utrecht.
- Gravemeijer, K. (1997). Solving word problems: A case of modelling? *Learning and Instruction*, 7, 389-397.
- Greer, B. (1993). The modeling perspective on word problems. *Journal of Mathematical Behavior*, 12, 239-250.
- Greer, B. (1997). Modelling reality in the mathematics classroom: The case of word problems. *Learning and Instruction*, 7, 293-307.
- Hatano, G. (1997). Commentary: Cost and benefit of modeling activity. *Learning and Instruction*, 7, 383-387.
- Hidalgo, M.C. (1997). *L'activation des connaissances à propos du monde réel dans la résolution de problèmes verbaux en arithmétique*. (Niet-gepubliceerd proefschrift). Quebec, Canada: Université Laval.
- Lakoff, G., & Nunez, R. (1997). Cognitive foundations for a mind-based mathematics. In L. English (Ed.), *Mathematical reasoning. Analogies, metaphors, and images* (pp. 21-92). Mahwah, NJ: Erlbaum.
- Lasure, S. (1995). *Realistisch modelleren en interpreteren van schoolvraagstukken. Twee constaterende studies bij 10-11-jarige basisschoolleerlingen*. (Niet-gepubliceerde licentiaatsverhandeling). Leuven: Afdeling Didactiek, K.U. Leuven.
- Lave, J. (1992). Word problems: A microcosm of theories of learning. P. Light & G. Butterworth (Eds.), *Context and cognition: Ways of learning and knowing* (pp. 74-92). New York: Harvester Wheatsheaf.
- Nesher, P. (1980). The stereotypical nature of school word problems. *For the Learning of Mathematics*, 1, 41-48.

- Radatz, H. (1983) Untersuchungen zum Lösen eingekleideter Aufgaben. *Journal für Mathematikdidaktik*, 3, 205-217.
- Reusser, K. (1988). Problem solving beyond the logic of things: Contextual effects on understanding and solving word problems. *Instructional Science*, 17, 309-338.
- Reusser, K., & Stebler, R. (1997a). Every word problem has a solution – the social rationality of mathematical modeling in schools. *Learning and Instruction*, 7, 309-327.
- Reusser, K., & Stebler, R. (1997b, August). *Realistic mathematical modeling through the solving of performance tasks*. Paper presented at the 7th European Conference on Learning and Instruction, Athens, Greece.
- Rogoff, B. (1984). Introduction: Thinking and learning in social context. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in a social context* (pp. 1-8). Cambridge, MA: Harvard University Press.
- Säljö, R., & Wyndhamn, J. (1988). Cognitive operations and educational framing of tasks. School as a context for arithmetic thought. *Scandinavian Journal of Educational Research*, 32, 61-71.
- Säljö, R., & Wyndhamn, J. (1993). Solving everyday problems in the formal setting. An empirical study of the school as context for thought. In S. Chaiklin, & J. Lave (Eds.), *Understanding practice: Perspectives on activity and context*. (pp. 327-342). Cambridge: Cambridge University Press.
- Schoenfeld, A.H. (1991). On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics. In J.F. Voss, D.N. Perkins & J.W. Segal (Eds.), *Informal reasoning and education* (pp. 311-343). Hillsdale, NJ: Erlbaum.
- Treffers, A., & De Moor, E. (1990). *Proeve van een nationaal programma voor het reken-wiskundeonderwijs op de basisschool. Deel 2. Basisvaardigheden en cijferen*. Tilburg: Zwijssen.
- Van Lieshout, E.C.D.M., Verdwaald, A., & Van Herk, J. (1997, August). *Suppression of real-world knowledge and demand characteristics in word problem solving*. Paper presented at the 7th European Conference on Learning and Instruction, Athens, Greece.
- Verschaffel, L. (1996, January). *Real-world knowledge and mathematical modeling of school word problems*. Lecture presented as part of the 'Mathematikdidaktisches Kolloquium, Wintersemester 95/96' of the Institut für Didaktik der Mathematik, Universität Dortmund, Deutschland.
- Verschaffel, L., & De Corte, E. (1996). *Leren realistisch modelleren en interpreteren van vraagstukken*. Een exploratief onderwijsexperiment bij leerlingen van de bovenbouw van de basisschool. *Pedagogische Studiën*, 73, 322-337.
- Verschaffel, L., De Corte, E., & Borghart, I. (1997). Pre-service teachers' conceptions and beliefs about the role of real-world knowledge in mathematical modelling of school word problems. *Learning and Instruction*, 4, 339-359.
- Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction*, 4, 273-294.
- Wyndhamn, J. & Säljö, R. (1995, August). *Quantifying time as a discursive practice. Arithmetics, calendars, fingers and group discussions as structuring resources*. Paper presented at the Sixth European Conference for Research on Learning and Instruction, Nijmegen, The Netherlands.
- Wyndhamn, J. & Säljö, R. (1997). Word problems and mathematical reasoning – a study of children's mastery of reference and meaning in textual realities. *Learning and Instruction*, 7, 361-382.
- Yoshida, H., Verschaffel, L., & De Corte, E. (1997). Realistic considerations in solving problematic word problems: Do Japanese and European children have the same difficulties? *Learning and Instruction*, 7, 329-338.

Auteurs

L. Verschaffel is Onderzoeksdirecteur bij het Fonds voor Wetenschappelijk Onderzoek, Vlaanderen, en Buitengewoon Hoogleraar aan de Faculteit der Psychologie en Pedagogische Wetenschappen van de K.U.Leuven.

E. De Corte is Gewoon Hoogleraar aan de Faculteit der Psychologie en Pedagogische Wetenschappen van de K.U.Leuven.

S. Lasure is wetenschappelijk medewerker van de Afdeling Didactiek van de Faculteit der Psychologie en Pedagogische Wetenschappen van de K.U.Leuven.

Adres: L. Verschaffel, Centrum voor Instructiepsychologie en -Technologie, K.U.Leuven, Vesaliusstraat 2, B-3000 Leuven, België, tel: 32/16/32 62 58, fax: 32/16/32 62 74;
E-mail: lieven.verschaffel@ped.kuleuven.ac.be

Abstract

Realistic modeling and solving of mathematical word problems in the elementary school: A review

L. Verschaffel, E. De Corte en S. Lasure. Pedagogische Studiën, 1999, 76, 183-200.

Some years ago Greer and Verschaffel, De Corte and Lasure provided evidence that after several years of traditional mathematics instruction children have developed a tendency to reduce mathematical modeling to selecting the correct formal-arithmetic operation with the numbers given in the problem, without seriously taking into account their common-sense knowledge and realistic considerations about the problem context. This evidence was obtained by means of a series of especially designed word problems with problematic modeling assumptions from a realistic point of view, administered in the context of a mathematical lesson. After having summarized these two initial studies, we briefly review a series of replication studies executed in different countries showing the omnipresence of this tendency among pupils. Then two related but different lines of follow-up studies are presented. While the first line of research investigated the effects of different forms of scaffolds added to the testing setting aimed at enhancing the mindfulness of students' approach when solving these problematic items, the second one looked at the effectiveness of attempts to increase the authenticity of the testing setting. Afterwards a study is reported which reveals that the strong tendency toward non-realistic mathematical modeling is found among (student-) teachers too. The last study involves a teaching experiment showing that it is possible to break this tendency by means of appropriate forms of instruction. We end with some theoretical, methodological and instructional implications of the work reviewed in this article.

Effecten en implementatie van een tutorprogramma voor risicoleerlingen

A.J.M. Vosse

Samenvatting

In dit artikel worden de cognitieve effecten van een interventieprogramma voor rekenen/wiskunde beschreven, waarbij leerlingen uit groep 7 fungeren als tutor voor risicoleerlingen uit groep 4.

Door leerlingen met elkaar te laten samenwerken kan worden tegemoet gekomen aan een toenemende noodzaak om leerlingen op een gedifferentieerde manier te benaderen. De leerkracht delegeert op een gecontroleerde manier specifieke instructietaken aan medeleerlingen, die daardoor verantwoordelijkheidsgevoel ontwikkelen ten opzichte van zowel het eigen leren als dat van medeleerlingen.

In het tutorprogramma dat in deze studie wordt beschreven, werkten op twee scholen in totaal zestien 10 à 11-jarige leerlingen uit groep 7 gedurende twaalf weken driemaal per week samen met een 7 à 8-jarige leerling. De opdrachten die tijdens deze tutorlessen worden uitgevoerd, zijn erop gericht om de rekenvaardigheden van de leerlingen van groep 4 te verbeteren. Nagegaan wordt, wat de kortetermijneffecten zijn van dit programma: verbeteren de rekenprestaties van de leerlingen uit groep 4? Verbeteren de rekenvaardigheden van de tutores op hun eigen niveau, d.w.z. op de toets voor groep 7?

Voor de deelnemende risicoleerlingen uit groep 4 blijken de rekenresultaten op de toets sterk te verbeteren. De resultaten van de tutores in groep 7 verbeteren wel enigszins op het onderdeel Getallen & Bewerkingen, maar nauwelijks op het onderdeel Meten, Tijd & Geld. Tevens blijkt tijdens de implementatie van het programma een aantal belemmeringen op te treden, die integratie van het tutorprogramma binnen de school bemoeilijken. Deze belemmeringen verschaffen waardevolle informatie over de condities waaronder tutorprogramma's in scholen geïmplementeerd kunnen worden, zodat optimaal gebruik kan worden gemaakt van deze innovatieve onderwijsstrategie.

Inleiding en theoretisch kader

Peer tutoring, in het Nederlands 'tutorprogramma' genoemd, is een specifieke vorm van coöperatief leren waarbij leerlingen elkaar begeleiden bij het verwerven van bepaalde vaardigheden. In Engelstalige landen zoals de Verenigde Staten, Groot-Brittannië, Schotland, Australië en ook in Israël ziet men regelmatig tutorprogramma's, maar in niet-Engelstalige Europese landen komen tutorprogramma's zelden voor. In Nederland echter zijn de laatste paar jaar enkele snel groeiende projecten gestart, zoals 'Stap Door' dat als leesproject in vervolg op 'Overstap' is ontwikkeld en een project voor het aanvankelijk rekenen/wiskunde dat is geïnitieerd door de Universiteit van Amsterdam. In beide programma's werken leerlingen uit groep 7(8) gedurende een half uur en drie keer per week in tweetallen samen met leerlingen uit groep 4(5). Uit internationaal onderzoek blijkt dat, indien tutorprogramma's onder de juiste condities worden geïmplementeerd, de leerprestaties van de deelnemende leerlingen kunnen verbeteren (Slavin, 1993; Slavin 1994; Slavin, Karweit & Madden, 1989; Cohen, 1992; Topping, 1988). Ook zijn er aanwijzingen dat tutorprogramma's kunnen leiden tot positieve veranderingen in attitudes en zelfconcept, hoewel de aanwijzingen voor deze sociaal-emotionele effecten wat minder sterk zijn dan voor de cognitieve effecten. Cohen, Kulik en Kulik (1982) voerden een meta-analyse uit met vijfenzeftig onderzoeken naar effecten van tutorprogramma's. Uit de tweeënvijftig onderzoeken die cognitief van aard waren, concludeerden zij dat in 20 van deze onderzoeken de geholpen leerlingen ('tutees') significante leerwinsten maakten ten opzichte van de controlegroep. Een andere conclusie uit deze meta-analyse is, dat de programma's met een korte duur van vijf tot zes weken de grootste Effect Sizes laten zien. De gemiddelde leerwinst van de tutees (in de tweeënvijftig cogni-

tief-gerichte onderzoeken) is een verschuiving van het 50ste naar het 66ste percentiel. Van de vijf onderzoeken naar attitudes wordt in vier onderzoeken een positievere attitude gevonden voor het vakgebied waarmee in de tutorles is gewerkt, hoewel geen van deze effecten significant is. In zestien van de vijftien zestig onderzoeken is het zelf-concept van de leerling nagegaan, al of niet in combinatie met cognitieve effecten. In twaalf van de zestien onderzoeken wordt een positief effect gevonden, waarvan vier significant.

Een andere meta-analyse is uitgevoerd door Cook, Scruggs, Mastropieri en Casto (1985/1986). In deze studie zijn onderzoeken geanalyseerd waarin de tutores leer- of gedragsproblemen of intellectuele handicaps hadden. In totaal werden vierenzeventig effect sizes vergeleken. Eén van de conclusies is, dat ook leerlingen met cognitieve of sociaal-emotionele problemen effectieve tutores kunnen zijn voor andere leerlingen, die al of niet kampen met leerproblemen. Tutores blijken in het algemeen dezelfde leerwinsten te verkrijgen als tutees. Het zelfbeeld van de betrokken leerlingen blijkt slechts licht te verbeteren, wat overeenkomt met de bevindingen van Cohen et al. (1982). Helaas wordt in de beschreven meta-analyses nergens iets genoemd over de condities waaronder de resultaten zijn bereikt. Ook de implementatieprocedures van de programma's met de grootste resultaten worden niet beschreven. Het is daardoor op grond van deze onderzoeksresultaten niet duidelijk welke condities van belang zijn bij de implementatie van een tutorprogramma.

Tutorprogramma's lijken een relatief goedkope methode te zijn om de leerprestaties te verbeteren, vergeleken met andere onderwijskundige vernieuwingen op dit gebied. Levin, Glass en Meister (1984) concluderen, dat de kosteneffectiviteit van tutorprogramma's, vergeleken met andere innovatieve onderwijskundige methoden in de Verenigde Staten, hoog is. Zo blijken tutorprogramma's twee tot vier keer zo kosteneffectief te zijn als CAI (Computer Assisted Instruction), het verkleinen van de klassen van 35 tot 30 leerlingen en het verlengen van de schooldag met één uur extra lesdij. Vanwege deze hoge kosteneffectiviteit is het de moeite waard om na te gaan welke condities van belang zijn bij de implementatie van tutorprogramma's.

Een belangrijk onderscheid binnen typen tutorprogramma's, is het verschil in leeftijd van de leerlingen die bij het programma zijn betrokken (Vosse, 1994). Een eerste vorm is de Classwide Peer Tutoring (CWPT), waarbij alle leerlingen binnen een klas in paren samenwerken (Greenwood, Delquadri & Hall, 1989). De beide leerlingen van het paar zijn hier van ongeveer dezelfde leeftijd. De paren kunnen al naar gelang het prestatieniveau van de leerlingen ofwel heterogeen ofwel homogeen zijn samengesteld. In heterogene paren heeft de beter presterende leerling de rol van tutor, in homogene paren kan de tutorrol wederkerig zijn (Palincsar & Brown, 1984). In een ander onderzoek beschrijft Ros (1994) hoe heterogene paren werken aan een taak over hefboomsonnen. De leerlingen van groep 7 die goed zijn in rekenen, geven een klasgenoot uitleg. Vóór en na het geven van de uitleg werd gemeten in hoeverre de medeleerlingen daarvan hadden geprofiteerd. De leerwinst van de uitlegontvangers bleek samen te hangen met de wijze waarop de uitleg was gegeven. Als de uitleggever effectieve instructie had gegeven, d.w.z. stimuleerde tot actief nadenken, structuur in de uitleg aanbracht en belangrijke (deel)stappen herhaalde, dan was de leerwinst bij de uitlegontvanger groter dan bij het ontbreken van effectieve instructie. Tevens bleken uitlegontvangers een grotere leerwinst te behalen, als zij uitleg kregen van leerlingen die hoog op de voortoets scoorden. Uitleggevers met een lage score op de voortoets bereikten een lage leerwinst bij hun uitlegontvanger. Kennelijk zijn leerlingen in heterogene paren met een goed inzicht in de stof, beter in staat om effectieve instructie te geven en bereiken zij ook betere leerresultaten bij de uitlegontvangende leerling.

In een 'cross-age' tutorprogramma is de tutor meestal afkomstig uit een hogere groep. In dit geval is de tutor ouder dan de tutee en daardoor verder in de leerstof, zodat hier altijd sprake is van heterogeniteit.

De laatste tientallen jaren is, internationaal gezien, relatief veel onderzoek gedaan naar peer tutoring, hoewel dit niet is gericht op specifieke minderheidsgroepen zoals risicoleerlingen en allochtonen. Eén van de gevolgen van de groeiende groep allochtonen, is het toenemende verschil in prestaties van leerlingen,

waardoor grote verschillen in leer- en instructiebehoefte ontstaan. Leerkrachten ervaren problemen bij het differentiëren van de instructie. Echter, door (mede)leerlingen in de klas in te schakelen en verantwoordelijkheid te geven voor het eigen leren en voor dat van anderen, kan de leerkracht gelegenheid creëren om specifieke hulp te bieden waar dat nodig is. Het werken met een tutorprogramma is één van de strategieën, waardoor differentiëren mogelijk is: elk kind kan die instructie en die lesinhoud krijgen die voor hem of haar nodig is. Bovendien voelen leerlingen zich onderling vrijer om hun mening te geven, vragen te stellen en niet goed-overwogen uitspraken te doen (Damon & Phelps, 1989). De persoonlijke band met de tutor geeft de samenwerking een ongedwongen en ontspannen karakter. De kwaliteit van de tutor- uitleg is mogelijk wat minder dan wanneer de leerkracht iets uitlegt (Ros, 1994), maar de leerlingen geven elkaar veel meer feedback en individuele aandacht dan de leerkracht ooit kan geven.

Samenwerkingsrelaties leiden niet automatisch tot betere leerprestaties, omdat eerst aan een aantal voorwaarden moet worden voldaan (Van der Linden, 1987; Kagan, 1994; Ebbens, Ettekoven & Van Rooijen, 1997). Een eerste voorwaarde is het nastreven van een gemeenschappelijk doel, dat via positieve onderlinge afhankelijkheid wordt bereikt, dat wil zeggen dat de leerlingen elkaar nodig moeten hebben om tot het eindresultaat te komen. Ook dienen de leerlingen zich ieder persoonlijk verantwoordelijk te voelen voor hun eigen bijdrage aan de taak. De leerkracht heeft hiervoor specifieke kwaliteiten nodig, zoals het voortdurend feedback geven aan leerlingen en bijsturen waar dat nodig is. Ook de samenwerkende leerlingen hebben communicatieve en sociale vaardigheden nodig, die de leerkracht in trainingslessen kan laten oefenen. Hoe meer de samenwerkingsrelaties voldoen aan de genoemde voorwaarden, hoe positiever de effecten zijn. Jenkins en Jenkins (1987) beschrijven de componenten van een succesvol tutorprogramma. De leerkrachten structureren de lessen zorgvuldig vooraf, stemmen inhoudelijk af op het niveau van de leerling en controleren of de leerlingen de gestelde doelen behalen. Actieve supervisie van de leerkrachten is daarmee een belangrijke component: dagelijks

wordt gevolgd of de leerlingen de stof beheersen. Bij het hieronder beschreven onderzoek zijn de bovengenoemde componenten zoveel mogelijk verwerkt in de opzet van het programma.

Onderzoeksvragen

Het bovenstaande resulteert in de volgende onderzoeksvragen:

1. Leidt een cross-age tutorprogramma voor aanvankelijk rekenen/wiskunde bij risico-leerlingen uit groep 4 tot betere rekenprestaties?
2. Leidt het functioneren als tutor voor een leerling in groep 4 tot betere rekenprestaties op het eigen rekenniveau van de tutores in groep 7?

1 Methode

1.1 Experimentele en controlescholen

Vier scholen participeerden in dit onderzoek: op twee scholen werd een tutorprogramma ingevoerd, twee andere scholen met een vergelijkbare populatie dienden als controlegroep. Voor alle leerlingen in de groepen 4 en 7 werd met de pretest een individuele vaardigheidsscore bepaald. De experimentele groep bestond uit één reguliere boven-modale school en één school met een percentage van 80% allochtone leerlingen van Turkse of Marokkaanse afkomst (onderwijsvoorrangsschool). De controlegroep bestond eveneens uit een reguliere boven-modale school en een onderwijsvoorrangsschool. De controlescholen hadden samen 33 leerlingen in groep 4 en 41 leerlingen in groep 7.

In totaal deden 16 koppels mee aan het experiment: op de beide onderzoeksscholen steeds acht leerlingen uit groep 4 en acht tutores uit groep 7. De scholen zijn niet-random geselecteerd en deden op basis van vrijwilligheid mee, nadat zij door de onderzoeker waren benaderd en geïnformeerd over het doel en de inhoud van het programma.

1.2 Leerkrachten en studenten

Vier studenten van de Universiteit van Amsterdam (zonder onderwijsbevoegdheid) voerden

in het kader van hun Pedagogiek-stage het tutorprogramma op de twee scholen in. Op elke school waren twee student-stagiaires. De leerkrachten van groep 4 en 7 in de betrokken scholen moesten lesgeven tijdens de tutorlessen en waren daardoor slechts indirect betrokken bij de uitvoering van het programma. Wel werd in een tweewekelijks overleg met de leerkrachten het verloop en de inhoud van het programma besproken.

1.3 Training en voorbereiding van de studenten

De studenten werden twee- of driewekelijks op de Universiteit begeleid, om de voorbereidingen voor het tutorprogramma te treffen en om inhoudelijk op het gebied van rekenen/wiskunde geschoold te worden. Na de tweede pretest in januari startte het tutorprogramma.

1.4 Selectie van tutees

Na de tweede pretest werden uit de groepen 4 van de beide onderzoeksscholen elk acht leerlingen gekozen, die lage scores hadden op de beide voormetingen met de Cito-toets voor Rekenen & Wiskunde. Op de beide scholen werden hiermee zestien tutees van 7 of 8 jaar geselecteerd.

1.5 Selectie van tutoren

De zestien tutoren waren 10 of 11 jaar oud en werden geselecteerd uit de groepen 7 van de beide onderzoeksscholen. Alle leerlingen van groep 7 konden na een introductiebijeenkomst over het tutorprogramma schriftelijk aangeven of zij tutor wilden worden. Slechts enkele leerlingen gaven aan dit niet te willen. De groepsleerkracht selecteerde vervolgens de tutoren op grond van twee criteria: ofwel een lage score op de Cito-toets, of vanuit een behoefte aan sociale ondersteuning door bijvoorbeeld gebrek aan sociale contacten in de klas. In veel gevallen voldeden de tutoren aan deze beide criteria.

1.6 Het matchen van de koppels

De tutor/tutee-koppels werden door de eigen groepsleerkrachten in overleg samengesteld, waarbij zaken als temperament of bepaalde overeenkomsten in hobbies of vrijetijdsbesteding een reden vormde om kinderen aan elkaar te koppelen.

1.7 Training van de tutoren

De studenten trainden de tutoren in zes lessen gedurende twee weken. De eerste les, een introductieles, werd door heel groep 7 gevolgd en bevatte informatie over het tutorprogramma. Na deze eerste introductieles en na de selectie van de acht tutoren, werd buiten de klas de training in sociale, pedagogische en didactische vaardigheden vervolgd (Myrick & Bowman, 1991; Foster, 1992).

Na de start van de tutorlessen volgden de tutoren wekelijks een begeleidingsles, waarin de inhoud van de tutorlessen werd toegelicht en zonodig geoefend. Ook werd feedback gegeven op de tutorlessen van de afgelopen week en werd zonodig nog teruggekomen op de vaardigheden die in de begintraining waren geleerd. Tevens oefenden de tutoren in het stellen van de juiste vragen bij de verschillende opdrachten, met als doel om de tutee zoveel mogelijk te stimuleren om zelf na te denken en verschillende oplossingen te bedenken. De tutoren oefenden drie soorten vragen:

- a. Oriënteringsvragen, zoals 'Wat zie je? Wat denk je dat je hier moet doen?'
- b. Hulp- en steun-vragen, zoals 'Wat ga je doen? Wat moet er nu gebeuren?'
- c. Reflectieve vragen, zoals 'Weet je nog een manier om dit op te lossen? Kun je het nog handiger aanpakken?'

Na afloop van de tutorles volgde een korte nabespreking met de tutoren.

2 Procedurele aspecten van de implementatie

2.1 Algemeen raamwerk

Het tutorprogramma is ingevoerd tussen februari en juni, uitgezonderd de schoolvakanties. In totaal zijn ± 35 tutorlessen (drie per week) gegeven. De tutorlessen vormden een aanvulling op de gewone reken- en wiskundelessen in de eigen groep.

2.2 Inroostering

Op de beide scholen werden de tutorlessen op verschillende tijdstippen gegeven, afhankelijk van het binnen de school gehanteerde lesrooster. De tijdstippen waren zo gekozen, dat de koppels geen belangrijke instructies in hun eigen groep zouden missen. Na verloop van

een aantal weken bleek, dat de tijdstippen soms door de tutores als ongunstig werden ervaren. Op één school misten zij aan het begin van de ochtend het kringgesprek met de hele groep, waarover zij hun ongenoegen duidelijk kenbaar maakten, maar een wijziging in de inroostering bleek niet mogelijk.

2.3 Bepalen van het rekenniveau van de tutee

Na de pretests die in de hele groep werden afgenomen, werd door de studenten met elke tutee afzonderlijk een diagnostisch gesprek gevoerd, om te achterhalen welke oplossingsstrategieën door hen werden gebruikt. Alle tutees bleken een zwak inzicht in de getallenrij tot 100 te hebben en de meeste opgaven tellend uit te rekenen. Met deze informatie werd de inhoud van het tutorprogramma voor alle leerlingen vastgesteld op hetzelfde beginniveau.

2.4 Activiteiten en materialen

In elke tutorles werden steeds drie of vier gevarieerde activiteiten gedaan, wat werd genoteerd op een opdrachtkaart. Er werd gewerkt met de ketting van 100 kralen, met getallenlijnen, dobbelstenen en afbeeldingen van dobbelsteenfiguren (getalbeelden), oefenbladen uit Remelka, rekenspelletjes uit diverse rekenmethodes, rekenraadsels en opgaven uit het MORE-onderzoek van Van den Heuvel-Panhuizen en Gravemeijer (1989).

2.5 Betrokkenheid van het team

De leerkrachten werden door de studenten in de tweewekelijkse vergaderingen geïnformeerd over het verloop van en de vorderingen met het programma. Op uitnodigingen om een (deel van een) tutorles of begeleidingsles bij te wonen konden de leerkrachten zelden ingaan, omdat zij op die momenten lesgaven. Integratie van het tutorprogramma in de school werd hierdoor bemoeilijkt.

2.6 Motivatie en uitval

De tutores waren aanvankelijk zeer enthousiast over hun nieuwe taak. Mede door de inroostering op de ongunstige momenten nam de motivatie echter af, wat leidde tot enkele wisselingen van leerlingen. Dit speelde vooral op de onderwijsvoorrangsschool. Bovendien was het voor de studenten moeilijk om de orde te hand-

haven in de groep, waarbij het gebrek aan onderwijservaring wellicht een rol speelde.

3 Design en data-analyse

3.1 Design

Het quasi-experimentele onderzoek heeft een pretest-posttest controlgroep design. Bij alle leerlingen op de vier scholen zijn twee pretests afgenomen, in september en in januari, vlak vóór de start van het tutorprogramma. De posttest vond in juni plaats, direct na afloop van het programma. Voor zowel groep 4 als voor groep 7 is de Cito-toets voor Rekenen en Wiskunde gebruikt.

3.2 Data-analyse

De beide onderzoeksvragen worden beantwoord door analyse van de Cito-toetsresultaten met behulp van multivariate analysetechnieken in SPSS (MANOVA). In dit design met herhaalde metingen worden de interactie-effecten tussen het verloop van de beide voormetingen en de nameting voor de beide groepen geanalyseerd.

4 Resultaten

Beantwoording van de onderzoeksvragen:

Onderzoeksvraag 1:

1. Leidt een cross-age tutorprogramma voor aanvankelijk rekenen/wiskunde bij risicoleerlingen uit groep 4 tot betere rekenprestaties?

Eén tutee heeft door vroege uitval slechts kort aan het programma deelgenomen, zodat vijftien van de zestien tutees in de analyses zijn opgenomen. Het aantal controleleerlingen bedraagt drieëndertig. In Tabel 1 worden de gemiddelde toetsscores weergegeven van de experimentele en de controlegroep. Bij de eerste pretest heeft de controlegroep 56 punten als gemiddelde vaardigheidsscore. Dat is 6 punten beter dan de experimentele groep met de 15 zwakke rekenaars (de tutees). Na de tweede pretest is dit verschil opgelopen tot 10 punten. Bij de posttest behalen beide groepen gemiddeld 70 punten, er is geen verschil meer tussen de beide groepen leerlingen.

Tabel 1

Gemiddelden en standaarddeviaties voor groep 4 op de beide pretests en de posttest van de Cito-toets voor Rekenen & Wiskunde

	Experimentele groep (N=15)	Controle groep (N=33)
	Mean (SD)	Mean (SD)
pretest 1	48 (10.1)	56 (11.0)
pretest 2	52 (9.1)	62 (10.5)
posttest	70 (6.8)	70 (9.0)

In Figuur 1 zien we deze ontwikkeling gevisualiseerd.

Om verschillen in het verloop van de toetsscores tussen de experimentele en de controle-groep statistisch te toetsen, is gebruik gemaakt van multivariate-analyse. Voor dit doel zijn de interactie-effecten tussen de verschillende meetmomenten van het al of niet deelnemen aan de experimentele groep getoetst. Het verloop van de leercurves van de beide onderzoeksgroepen blijkt significant verschillend te zijn ($F(2,92)=158.4$; $p=.000$). Het verschil in het verloop van de leercurves tussen de beide groepen is significant.

Onderzoeksvraag 2:

Leidt het functioneren als tutor voor een leerling in groep 4 tot betere rekenprestaties op het eigen rekenniveau van de tutores in groep 7?

- a. voor Getallen & Bewerkingen
- b. voor Meten, Tijd & Geld

Resultaten:

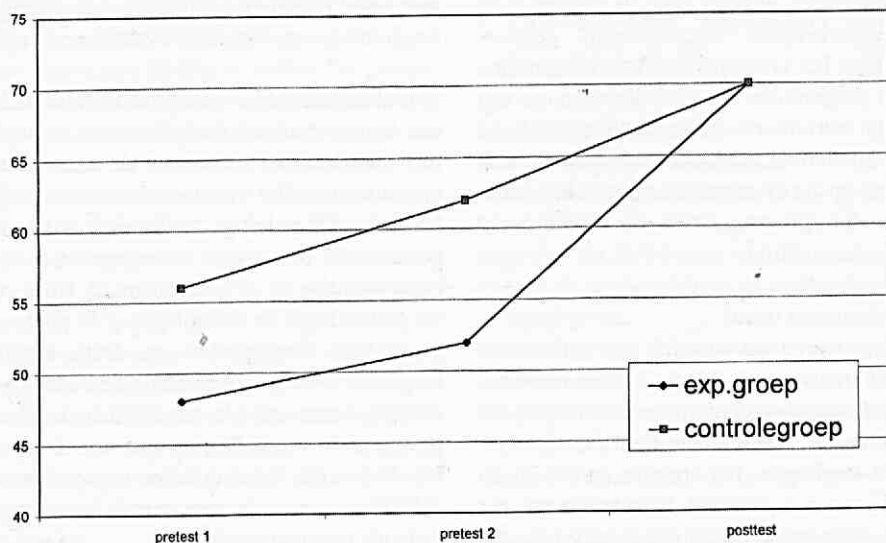
Twee tutores hebben het programma voortijdig verlaten, zodat uiteindelijk 14 tutores en 41 controleleerlingen uit groep 7 in de analyse zijn betrokken.

Tabel 2

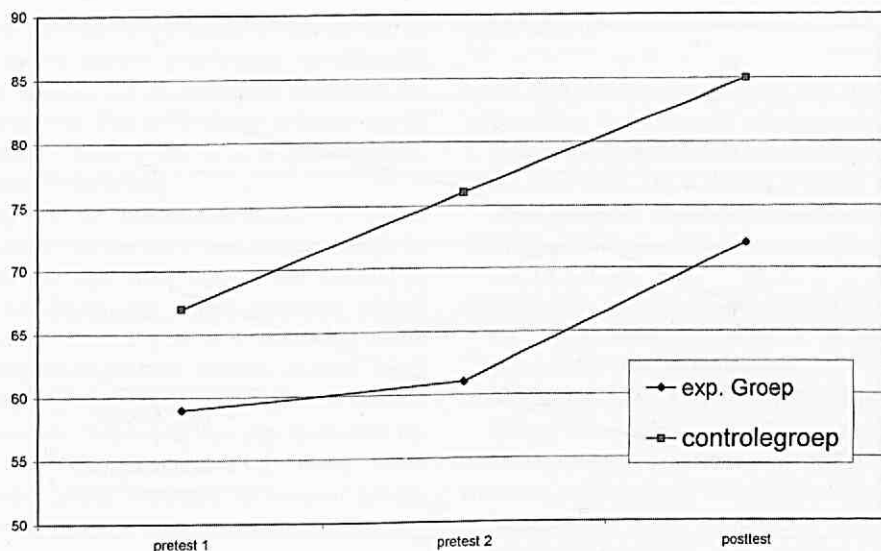
Gemiddelden en standaarddeviaties voor groep 7 op de beide pretests en de posttest, gemeten met de Cito-toets voor Rekenen & Wiskunde, onderdeel Getallen & Bewerkingen

	Experimentele groep (N=14)	Controle groep (N=41)
	Mean (SD)	Mean (SD)
pretest 1	59 (10.2)	67 (8.8)
pretest 2	61 (15.6)	76 (14.9)
posttest	72 (10.0)	85 (10.6)

Tabel 2 toont de gemiddelde scores van de tutores op het onderdeel 'Getallen & Bewerkingen'. Hier zien we niet zo'n duidelijke groei van de gemiddelde scores als bij de tutees. De tutores (gemiddelde score: 59) scoren bij de eerste pretest 8 punten minder dan de controle-groep, bij pretest 2 is dit 15 punten minder. Bij de posttest scoren zij 13 punten lager. In Figuur 2 is deze ontwikkeling te zien.



Figuur 1. Gemiddelden op de Cito-toets Rekenen & Wiskunde groep 4



Figuur 2. Gemiddelden op de Cito-toets R & W Getallen & Bewerkingen groep 7

Ook hier is gebruik gemaakt van multivariate-analyse om de verschillen in het verloop van de leercurves te toetsen. De leercurves van de beide onderzoeksgroepen blijken significant van elkaar te verschillen ($F(2,106)=72.9$; $p=.019$). Al is de vooruitgang visueel in Figuur 2 niet goed waarneembaar, kennelijk hebben de tutoren toch geprofiteerd van het oefenen in de basisvaardigheden die voor het onderdeel Getallen & Bewerkingen van de Cito-toets nodig zijn: het verschil in het verloop van de leercurves tussen de tutoren en de controleleerlingen is significant.

Tabel 3
Gemiddelden en standaarddeviaties voor groep 7 op de beide pretests en de posttest, gemeten met de Cito-toets voor Rekenen & Wiskunde, onderdeel Meten, Tijd & Geld

	Experimentele groep (N=14)	Controle groep (N=41)
	Mean (SD)	Mean (SD)
pretest 1	55 (12.0)	66 (8.4)
pretest 2	60 (12.1)	74 (11.5)
posttest	67 (13.0)	79 (12.6)

In Tabel 3 zien we de toetsgemiddelden voor de tutoren op de subtest 'Meten, Tijd & Geld'. Bij de eerste pretest hebben de tutoren een achterstand ten opzichte van de controlegroep van

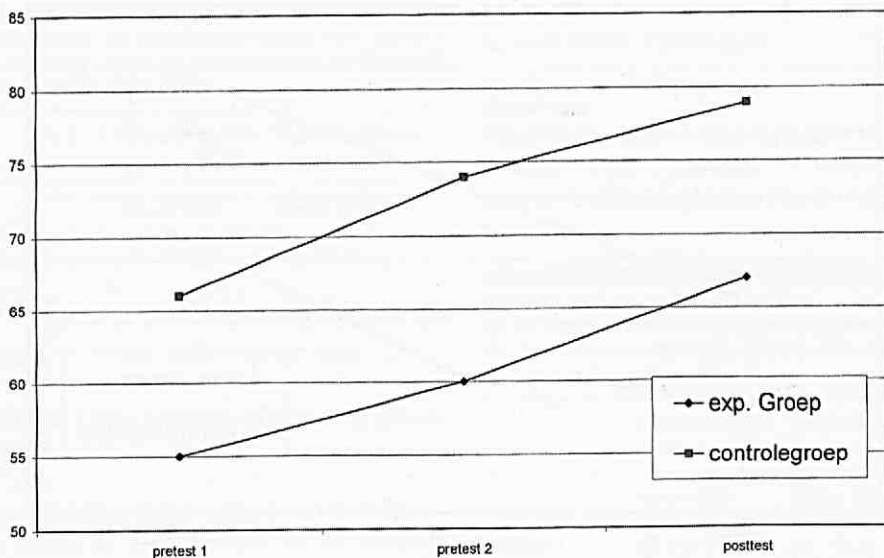
11 punten, bij de tweede pretest is dit verschil 14 punten, bij de posttest nog 12 punten.

De resultaten van tabel 3 zijn weergegeven in Figuur 3.

Bij toetsing van de interactie-effecten tussen de drie meetmomenten, wordt geen significant verschil voor de experimentele groep gevonden ($F(2,106)=42.62$; $p=.366$). Voor het gebied Meten, Tijd & Geld blijkt het uitvoeren van het tutorprogramma niet van invloed te zijn op de rekenresultaten van de tutoren, zodat de onderzoeksvraag hier negatief moet worden beantwoord.

Kort samengevat blijken de tuteurs die deelnemen aan het tutorprogramma een aanzienlijke vooruitgang te vertonen in rekenprestaties op de Cito-toets. Bij de tutoren wordt op de subtest 'Getallen en Bewerkingen' wel een significant effect gevonden, bij Meten, Tijd & Geld niet. De achterstand ten opzichte van de controlegroep wordt echter niet ingelopen, zoals is te zien in Figuren 2 en 3.

Vanwaar komt dit verschil in resultaat tussen tuteurs en tutoren? De tuteurs hebben tijdens de tutorlessen gewerkt met leerstof die op hun feitelijke rekenniveau ligt en die tevens in de Cito-toets wordt getoetst. De verbetering in de basisvaardigheden waarin is geoefend komt direct tot uiting in een hogere toetscore. De



Figuur 3. Gemiddelden op de Cito-toets R & W Meten, Tijd & Geld groep 7

tutoren hebben echter gewerkt met leerstof die ver afstaat van de rekenstof waar in de eigen groep mee wordt gewerkt. De Cito-toets bevat opgaven op het niveau van groep 7. Dat er toch een effect wordt gevonden bij Getallen & Bewerkingen komt waarschijnlijk, doordat bepaalde basisvaardigheden voor het rekenen onder de 100 nodig zijn voor de hoofdrekenopgaven in dit toetsonderdeel. Het hoofdrekenen is namelijk wel tijdens de tutorlessen geoefend, hoewel op een lager niveau dan in de Cito-toets, zodat hier een licht effect bij de tutoren optreedt. Voor wat betreft het toetsonderdeel Meten, Tijd & Geld is het verklaarbaar dat geen verbetering optreedt, daar met deze onderdelen tijdens de tutorlessen nauwelijks is gewerkt.

5 Conclusies en discussie

Naast de beschreven, overwegend positieve effecten, bleek de implementatie en integratie van het tutorprogramma binnen de scholen niet als vanzelfsprekend te verlopen. Waarschijnlijk heeft dit te maken met het feit dat in deze studie niet met leerkrachten vanuit de school is gewerkt, maar met studenten die noch hun onderwijsbevoegdheid bezaten, noch deel uitmaakten van het schoolteam. Bovendien werden de tutorlessen buiten de klas gegeven, waardoor integratie binnen de school werd

bemoeilijkt. De leerkrachten in de groep bleken niet actief bij het programma betrokken te zijn, hoewel zij vooraf bewust voor het programma hadden gekozen. Zelfs de tweeweekelijkse bijeenkomsten, die tot doel hadden om de actieve steun van de leerkrachten te krijgen, bereikten dit doel niet. Mogelijk moeten leerkrachten zelf actief participeren, bijvoorbeeld door (delen van) de tutortraining te geven, of door de tutorlessen in de groepen te laten plaatsvinden, teneinde integratie binnen de school te kunnen bewerkstelligen.

Een ander aspect is het inroosteren van het tutorprogramma. Het missen van essentiële groepsactiviteiten blijkt een negatieve invloed op de motivatie uit te oefenen. Momenten waarop leerlingen zelfstandig aan het werk zijn, lijken daarom een gunstiger tijdstip voor de tutorlessen te zijn, omdat dan geen groepsactiviteiten worden gemist. Verder hebben de tutoren af en toe nieuwe prikkels nodig om voldoende gemotiveerd te blijven. Daarbij is het zorgen voor voldoende variatie in de opdrachten die tijdens de tutorlessen worden uitgevoerd van belang. Ook kan eraan worden gedacht om na verloop van tijd de koppels opnieuw samen te stellen. Het gebruikmaken van spel- of wedstrijdvormen, zodat een positieve vorm van competitie ontstaat, is een andere mogelijkheid om de motivatie vast te houden (Slavin, 1996).

Wat de inhoud van het programma betreft,

bleken sommige werkbladen, vooral als daarop veel moest worden geschreven, onvoldoende uit te nodigen tot de intensieve interactie die bedoeld was. Een zorgvuldige selectie van de opdrachten waarbij interactie is gegarandeerd, is daarom van belang.

Tijdens de begeleidingslessen is steeds benadrukt, dat de tutor veel vragen diende te stellen aan zijn tutee. Opdrachten die niet in een begeleidingsles waren geoefend bleken vaak niet helemaal goed te worden uitgevoerd. De begeleidingslessen hebben daarom zeker een bijdrage geleverd aan de effecten van het programma. Supervisie van een leerkracht tijdens de tutorlessen is tevens van belang, om de tutores continu feedback te kunnen geven, zodat zij reflectie op het eigen handelen ontwikkelen.

De cognitieve resultaten van met name de tutees zijn in het hier beschreven experiment bemoedigend. De tutees boekten grote leerwinsten en ook bij de tutores wordt op de subtest Getallen & Bewerkingen een effect gevonden. Het ontbreken van een effect op de subtest Meten, Tijd & Geld, ligt mogelijk aan het feit dat de toetsen zijn afgenomen op het klassenniveau van de tutores en niet op het niveau waarmee zij in het tutorprogramma hebben gewerkt. Overigens is enige voorzichtigheid met betrekking tot de generaliseerbaarheid van de gevonden effecten geboden, daar slechts twee scholen aan het experiment deelnamen. Het beschreven onderzoek vormt een pilot-studie voor een grootschaliger onderzoek waarin meer scholen zijn betrokken, zodat generalisatie daarin wel mogelijk is. In dit grootschaliger onderzoek dat op dit moment wordt uitgevoerd, is het implementatieproces verbeterd. Zo wordt bijvoorbeeld niet langer met studenten gewerkt, maar met bevoegde leerkrachten. Ook in veel andere opzichten is het programma aangepast. Over enige tijd zullen de resultaten van dit onderzoek beschikbaar komen.

Met het hier beschreven onderzoek en de problemen die tijdens de implementatie naar voren zijn gekomen, is kennis omtrent de invoering van tutorprogramma's verkregen, die in geval van toekomstige implementaties van nut kan zijn. Hiermee is tevens een bijdrage geleverd aan de omvangrijke literatuur die over het onderwerp peer tutoring al beschikbaar is.

Literatuur

- Cohen, E.G. (1992). *Restructuring the classroom: conditions for productive small groups*. Paper presented at the IASCE-conference, Utrecht.
- Cohen, P.A., Kulik, J.A., & Kulik, C.C. (1982). Educational outcomes of tutoring: a meta-analysis of findings. *American Educational Research Journal*, 19, 237-248.
- Cook, S.B., Scruggs, T.E., Mastropieri, M.A., & Casto, G.C., (1985/1986). *Handicapped students as tutors*. Utah: Utah State University.
- Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *Journal of Educational Research*, 13(1), 9-19.
- Ebbens, S., Etteken, S., & Rooijen, J. van (1997). *Samenwerken leren: praktijkboek*. Groningen: Wolters-Noordhoff.
- Foster, E.S. (1992). *Tutoring: learning by helping; a student handbook for training peer and cross-age tutors*. Minneapolis: Educational Media Corporation.
- Greenwood, C.R., Delquadri, J.C., & Hall, R.V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology*, 81, 371-383.
- Heuvel-Panhuizen, M. van den, & Gravemeijer, K. (1989). De eerste uitkomsten. De eerste MORE-gegevens over het rekenonderwijs in groep 3 van de basisschool. In E. de Moor (red.), *Panama cursusboek 7. Rekenen-wiskunde, periodieke peiling onderwijsniveau, beredeneerde eindtermen, proeve van een nationaal programma* (pp.59-68). Utrecht: HMN / SOL en OW&OC.
- Jenkins, J.R., & Jenkins, L.M. (1987). Making peer tutoring work. *Educational Leadership*, 44(6), 64-68.
- Kagan, S. (1994). *Cooperative learning*. San Clemente: Kagan Cooperative Learning.
- Levin, H., Glass, G.V., & Meister, G.R. (1984). *A cost-effectiveness analysis of four educational interventions*. (Project Report No. 84-A11). Stanford: Stanford University, Institute for Research on Educational Finance and Governance.
- Linden, J. van der (1987). Samenwerken en leren in groepsverband. In P. Span, J.M.C. Nelissen, H.F. Pijning, C. Dietvorst, (red.) *Onderwijzen en leren: een onderwijspsychologische basis voor het handelen op school* (pp. 201-222). Groningen: Wolters-Noordhoff.
- Myrick, R.D., & Bowman, R.P. (1991). *Children helping children: teaching students to become friendly helpers*. Minneapolis: Educational Media Corporation.

- Palincsar, A.S., & Brown, A.L. (1984). Reciprocal teaching of comprehension monitoring activities. *Cognition and Instruction*, 2, 117-175.
- Ros, A.A., (1994). *Analyse en effecten van uitleg door leerlingen en leerkrachten*. Groningen: GION.
- Slavin, (1993). *Cooperative learning and achievement: an empirically-based theory*. AERA-conference, Atlanta.
- Slavin, R.E. (1994). *Preventing early school failure: research, policy, and practice*. Massachusetts: Allyn and Bacon.
- Slavin, R.E. (1996). *Education for all: contexts of learning*. Lisse: Swets & Zeitlinger.
- Slavin, R.E., Karweit, N.L., & Madden, N.A. (1989). *Effective programs for students at risk*. Boston, London, Sydney, Toronto: Allyn and Bacon.
- Topping, K. (1988). *The peer tutoring handbook: promoting cooperative learning*. London, Sydney: Croom Helm; Cambridge, MA: Brookline Books.
- Vosse, A.J.M. (1994). *Kinderen begeleiden kinderen bij achterstanden: een tutorbegeleidingsprogramma voor kinderen in het basis- en speciaal onderwijs*. Groningen: Wolters-Noordhoff.

Met dank aan Jos van der Linden voor zijn commentaar op een vorige versie van dit artikel.

Auteur

A.J.M. Vosse is als onderzoeker verbonden aan de faculteit Pedagogische Wetenschappen van de Universiteit van Amsterdam en is tevens werkzaam bij de Onderwijsbegeleidingsdienst Zuid-Kennemerland en bij het Seminarium voor Orthopedagogiek. Adres: Faculteit Pedagogische Wetenschappen, Universiteit van Amsterdam, Wibautstraat 4, 1091 GM Amsterdam.

Abstract

Effects and Implementation of a peer tutoring programme for children at risk

A.J.M. Vosse. *Pedagogische Studiën*, 1999, 76, 201-210.

This article describes the cognitive effects of an intervention strategy for maths, in which pupils from year 5 tutor 'at risk' year 2 children. Adaptive instruction is a growing necessity in education, and Cooperative Learning is one of the methods which can achieve this. The teacher delegates specific instructional tasks to peers in a controlled way. These peers (tutors) develop a sense of responsibility for their own learning as well as for the learning of their tutees.

In this study sixteen 10- or 11 year old tutors work with their 7- or 8 year old tutees for three times a week over a 12-week period in two schools. The pairs work together during the tutoring sessions in order to raise the mathematical attainment of the tutees, who are all pupils from year 2. The short-term effects of the program for both tutees and tutors are examined: do the tutees improve their maths skills as measured by the national test?; is there any benefit for the year 5 tutors?; do they improve their academic results at their own academic level? Results show that the children at risk from year 4, the tutees, greatly improve their mathematical attainment. The achievements of the tutors improve slightly on the sub-test 'Numbers and Operations', but scarcely at all on the sub-test 'Measuring, Time & Money'.

During the implementation process a number of difficulties arose, which interfered with the full integration of the peer tutoring programs in schools. Thus knowledge is developed about the essential conditions for implementing peer tutoring programs in schools, so that optimal effects can be attained by this innovative educational strategy.

Boekbesprekingen

M. Depaepe

De pedagogisering achterna. Een aanzet tot een genealogie van de pedagogische mentaliteit in de voorbije 250 jaar.

Leuven/Amersfoort: Acco, 1998, 285 pagina's, f 73,50, ISBN 90 334 3992 1

Omvattende, wetenschappelijk interessante, historische boeken over de pedagogiek in de Nederlandse taal, zijn dun gezaaid. En dat zal nog verergeren: in de hedendaagse academische wereld kan men geen waardering meer oogsten met het schrijven van boeken (in plaats van artikelen), al helemaal niet als die in een lokale taal als het Nederlands geschreven zijn (in plaats van in de Engelse wereldtaal), en zeker ook niet als het gaat om op synthese gerichte beschouwingen (in plaats van op empirische data gerichte analyses). Het getuigt dan ook van moed en onafhankelijkheid als een belangwekkend en gerespecteerd geleerde als Marc Depaepe – onderzoeksdirecteur van het Fonds voor Wetenschappelijk onderzoek Vlaanderen en buitengewoon hoogleraar in de historische pedagogiek aan de KU Leuven – zich van een dergelijk klimaat niets aantrekt en een diepgaand, breed georiënteerd en zorgvuldig geconcentreerd boek aflevert over 250 jaar pedagogiek.

Het boek bestaat, blijkens het Voorwoord, uit een 'herbewerking van de colleges historische pedagogiek' (blz. 9), en daarvan het tweede deel ('van de Aufklärung tot heden'), dat overigens in de vorm van een college-dictaat bij menigeen al bekend was. Het is een rijk boek, waarin de ganse geschiedenis van de pedagogiek vanaf Rousseau tot vandaag de revue passeert, ingebed in de context van de cultuurgeschiedenis, waarvan de pedagogiek deel uitmaakt. Het onderwerp van het boek is overigens niet de pedagogiek, zoals uit de titel blijkt, maar: de pedagogisering. Daaronder verstaat Depaepe '...de toegenomen (en nog steeds toenemende) invloed van het pedagogische op het dagelijks leven van kinderen en jongeren' (blz. 10), een proces dat volgens de schrijver ten nauwste samenhangt met de 'modernisering van de (westerse) samenleving en dat vooral vanaf de tweede helft van de 18de eeuw zijn beslag heeft gekregen' (t.a.p.).

Met de term 'genealogie' heeft de auteur willen 'knipogen' naar de Franse filosoof Michel Foucault, en hij heeft daarmee willen aangeven dat 'het om een 'ontmaskerende', bijwijlen zelfs 'carnavaleske' vorm van geschiedschrijving gaat' (blz. 10). Ook meldt de schrijver – bescheiden?, veiligheidshalve? – dat 'het gebodene slechts een inleiding met een erg 'onaf' karakter betreft' (blz. 11), maar zo voegt hij er, wederom met een post-moderne knipog, aan toe: 'dat kan – op straffe van zelf weer aan de geschiedenis te willen ontsnappen – moeilijk anders' (blz. 11).

Het betoog van Depaepe is overigens vervolgens tamelijk traditioneel: hij begint bij de welbekende visie van Ariès, en bekritiseert en nuanceert deze in het licht van de belangrijkste publicaties van de laatste decennia. Het aardige is dat Depaepe zich niet wil laten verleiden tot het kiezen van een extern (historisch of sociologisch), 'geleend' interpretatiekader (zoals bijvoorbeeld dat van Elias, of van Ariès) om de pedagogiek te plaatsen, maar hij wil de pedagogisering vanuit 'immanente kaders' begrijpen (blz. 27). Daartoe wil hij de verwaarlozing van de 'pedagogische werkvloer (het microniveau)' (blz. 31) opheffen, en 'het reële opvoedingsgebeuren binnen en buiten het gezin, in de schoolruimte zowel als in het klaslokaal' (blz. 31), in het centrum van de pedagogische historiografie plaatsen.

Na het inleidende hoofdstuk, waaruit het voorgaande geciteerd werd, volgt een groot en informatief hoofdstuk over de modernisering (Aufklärung en Romantiek) als drijfkracht achter de pedagogisering: Rousseau en de filantropijnen, de ontwikkeling van het onderwijs via de concepties van Condorcet versus Destutt de Tracy, Pestalozzi en Fröbel, en dan: een uiteenzetting over de verwetenschappelijking van de pedagogiek (Trapp, Kant, academisering van onderwijzersopleidingen) om de pedagogisering nader kracht bij te zetten. Het hoofdstuk sluit af met een specifiek overzicht van de pedagogisering in deze periode in België (zoals steeds in dit boek: met de nadruk op Vlaanderen).

Hoofdstuk 3 behandelt de 'wording van de gepedagogiseerde samenleving' in de periode 1840-1914. Veel aandacht wordt hier besteed aan de systematisering van de pedagogisch-didactische theorie via het Herbartianisme, dat

uiteindelijk leidde tot een passieve, intellectua-
listische luisterschool, en dat de reactie van
de reformpedagogiek opriep. De reformpeda-
gogiek en de pedologische 'Child Study
Movement' worden systematisch behandeld.
Overigens is Depaepe nogal ironisch, zo niet
wat cynisch, over de betekenis van al die
wetenschappelijke opvoedkunde. Hij meent,
voor de derde maal 'met een knipoog naar wat
in de marge van het postmodernisme-debat
aan het licht is gekomen' (blz. 143), dat de
experimentele aanpak, evenals de gehele
reformpedagogiek (wier vertegenwoordigers -
Kerschensteiner, Dewey, Reddie, Lietz, Ligt-
hart, Ellen Key, Montessori en Decroly in dit
hoofdstuk uitvoerig besproken worden) veel-
eer tot 'nieuwe en subtielere vormen van
machtsuitoefening heeft geleid' (t.a.p.).
Depaepe is van mening dat experimenteel
wetenschappelijke pedagogiek '...eerder een
bedrijf is geweest dat het beroepsethos van
zijn professionelen heeft gevoed, dan wel de
transfer van theorie naar praktijk en omge-
keerd' (blz. 143). Kortom: de experimentele
wetenschap bood (en biedt?) weinig meer dan
ijdelheid, en schijn-legitimatie. Overigens zij
aangetekend dat Depaepe's gebruik van de
term 'experimenteel-wetenschappelijk' nogal
slordig is: hij verwijst er gewoonlijk niet mee
naar de experimentele psychologie (en/of
pedagogiek) maar naar de vragenlijst-traditie
van Stanley Hall en naar de differentieel psy-
chologische test-traditie (Binet, Rice). Overi-
gens is Depaepe ook nogal relativerend ten
aanzien van de reformpedagogiek: aansluitend
bij het werk van Oelkers meent hij dat de
reformpedagogen veel meer continuïteit met
de pedagogisering uit het verleden, dan dis-
continuïteit aan de dag leggen. Dat ze 'iets
nieuws' uitprobeerden berust wat hem betreft
meer op legendevorming en mythologisering,
en dat was functioneel: wie de suggestie weet
te wekken dat men met iets werkelijk nieuws
 bezig is kan in het maatschappelijk debat
gemakkelijker 'aanzien' en status verwerven'
(blz. 133). De gebruikelijke geschiedschrij-
ving krijgt dan ook van Depaepe een veeg uit
de pan: '...duidelijk is alvast dat de klassieke
canon over de reformpedagogiek mede was
ingegeven door de traditionele opvatting over
de geschiedenis van de opvoedkunde als
dienstmaagd van de pedagogiek'. (blz. 135).

In dit derde hoofdstuk wordt tevens inge-
gaan op de uitbreiding van de wetenschap van
het kind naar kind en jeugdige. Zo is er ook de
nodige aandacht voor het ontstaan van de
jeugdbeweging en de jeugdbescherming. Ook
dit hoofdstuk sluit af met lange en gedetail-
leerde, en interessante, uitweidingen over de
specifieke ontwikkelingen (cultureel, maat-
schappelijk, politiek) in Vlaanderen.

In het laatste hoofdstuk behandelt Depaepe
'Van pedagogisering van de massa naar post-
modernistische massacultuur: opvoeding en
onderwijs in het teken van de opkomst en de
ondergang van de grote ideologieën (1914-
heden)'. Allereerst worden hier de normatief
pedagogische stromingen van communistische
origine (Blonski, Krupskaja, Makarenko),
katholieke origine (in Vlaanderen: Dehovre,
Decoene), en fascistische origine (Gentile) aan
de orde gesteld. Pijnlijk (voor o.m. de katho-
lieke landgenoten van Depaepe) is dat Depae-
pe een zekere 'paradigmatische eenheid'
(blz. 108) van deze stromingen aannemelijk
weet te maken. Van de 'experimentele bena-
dering' wordt vervolgens beweerd dat die tot
een einde kwam. Het treurige beeld dat De-
paepe hier ophangt van de testtraditie (Binet,
Terman), door deze te verbinden met de euge-
netica, en door ook nog in één adem het op
'uitwendige gedragscontrole' gerichte
behaviorisme te ridiculiseren, moet het voor
de lezer aannemelijk maken dat de inbreng
van de zogenaamde 'experimentele benade-
ring' in Europa van verwaarloosbare betekenis
is geweest, behalve voor het aanzien van de
professionaliteit van de pedagogen zelf.
Depaepe concludeert hierover: 'Wetenschaps-
historisch onderzoek heeft voldoende aange-
toond dat het streven naar academische status
en prestige de pedagogen mee in de richting
van het empirisch-analytisch theorietype heeft
gestuurd.' (blz. 212) Overigens meent De-
paepe dat ook de 'geesteswetenschappelijke
pedagogiek' geen alternatief is. Met Flitner is
hij van mening dat de geesteswetenschappe-
lijke aanpak in de jaren '60 'am Ausgang ihrer
Epoche' was gekomen (blz. 213). In dit hoofd-
stuk vraagt de auteur vooral wat de pedagogi-
sering nu eigenlijk impliceert: beheersing of
bevrijding? Dit dilemma wordt behandeld aan
de hand van de emancipatorische pedagogiek,
geïnspireerd op de Frankfurter Schule, en aan

de hand van de 'koloniale' opvoedings-ervaring (in Belgisch-Kongo), die naar het oordeel van Depaepe 'scherper nog dan de westerse opvoedingsgeschiedenis, de systeemfouten en de paradoxen blootlegt van de westerse modernisering en pedagogisering.'

Depaepe komt overigens niet veel verder dan het schetsen van dit dilemma: emancipatie versus koloniseren, 'wachsen lassen versus führen', ontwikkelen versus leren. Het hoofdstuk, en daarmee het boek, loopt af met een bespreking van de hedendaagse toestand in België (met zeer informatieve paragrafen over: eigentijdse gezinsopvoeding, eigentijds onderwijs, eigentijds jeugd- en vormingswerk, eigentijdse hulpverlening en eigentijdse pedagogiek). En al informerende over Belgische toestanden, gaat het boek uit als de spreekwoordelijke nachtkaaars.

Van het voornemen om de pedagogische werkvloer in het centrum te zetten, van het voornemen om het 'microniveau' te laten spreken is al met al niet veel terecht gekomen. Het grootste deel van het boek is gewijd aan bredere kaders, op meso- en macroniveau. Het aantal bladzijden dat daadwerkelijk besteed is aan de pedagogische omgang op microniveau, in plaats van aan culturele, maatschappelijke, politieke kaders, is uiteindelijk te verwaarlozen gering. Ook van de interessante en uitdagerende beloften uit het Voorwoord en uit het eerste hoofdstuk – over 'ontmaskering' en 'carnavaleske' geschiedschrijving – is per saldo niet zo heel veel terecht gekomen. Van de ontmaskering is er in zoverre iets te vinden dat de auteur de lezer duidelijk heeft gemaakt dat er niet veel is om in te geloven: niet in de normatieve pedagogieken, niet in de experimentele pedagogiek, niet in de geestes-wetenschappelijke pedagogiek. Waarin wel?? In 'intellectuele ascese' en in 'een besef van menselijke waardigheid' schrijft de auteur in de slotregels (blz. 272). Maar deze slotregels vallen een beetje uit de lucht, ze zijn niet verbonden met de voorgaande 271 bladzijden. En dat is spijtig.

Ondanks de expliciete en impliciete bedenkingen die in het voorgaande doorklinken, wil ik besluiten met waardering en bewondering. Depaepe heeft een eigentijdse geschiedenis van de pedagogisering geschreven, die een fraai, onbevooroordeeld, raamwerk biedt om

de pedagogische literatuur mee te lijf te gaan. De erudiete en kritische wijsheid van de auteur doet naar méér verlangen. Aangezien er ongetwijfeld een volgende druk komt, en de auteur dynamisch genoeg is om dan ook de tekst aan te vullen en te verbeteren, zou ik graag besluiten met de volgende wensen mijnerzijds: voeg een afsluitend samenvattend / concluderend hoofdstuk toe, waarin zo mogelijk ook nog iets over de toekomst (van de pedagogisering) gezegd wordt, en neem s.v.p. in het boek een naam- en zaakregister op.

W. Koops

L. Verschaffel, E. de Corte, G. van Vaerenbergh, S. Lasure, H. Bogaerts & E. Ratincx

Leren oplossen van wiskundige problemen in de bovenbouw van de basisschool.

Universitaire Pers Leuven, 1998, 187 pagina's.

ISBN 90 6186 897 1.

Wanneer leerlingen onverwacht worden geconfronteerd met een niet-standaard wiskundeprobleem, is de kans groot dat ze er standaard op reageren. Neem bijvoorbeeld het volgende probleem: 'Een groep van 22 kinderen wordt met bootjes naar een eiland gebracht. In één bootje kunnen vier kinderen. Hoeveel bootjes zijn er nodig?' Keurig volgens de regels rekenend, zullen er leerlingen zijn die 'vijf rest twee' als oplossing geven, maar het antwoord 'vijf en een half bootje' getuigt van nog meer precisie. De zeewaardigheid van een halve boot is geen probleem, het gaat immers om een correcte en standaard uitgevoerde berekening. In een aantal – in verschillende landen uitgevoerde – onderzoeken werd nagegaan hoe de leerlingen op dit type realistische 'word problems' reageerden en de bevindingen waren over het algemeen weinig hoopgevend (voor een overzicht van de onderzoeksverslagen verwijs ik naar *Teaching en Instruction* 1997, 7(4)). De onderzoekers dragen verschillende verklaringen aan voor deze tegenvallende resultaten, bijvoorbeeld dat de leerlingen zich richten op de 'syntax' in plaats van de betekenis van een opgave. Maar wordt – op basis van zulke problemen – in wezen niet gewoon onderzocht of leerlingen in een valkuil kunnen lopen. Maar dat kunnen ze, zoiets hoeft dus niet onderzocht te worden.

De kritische vraag is, met andere woorden, of de mate van waakzaamheid (alertheid) van een leerling wordt onderzocht of het niveau van de rekenvaardigheid. In een onderzoek dat onlangs werd afgesloten, is aan deze kritiek tegemoet gekomen en werden de leerlingen geleerd wiskundige problemen op te lossen en daarbij rekening te houden met de realiteit van de context. Een gewaarschuwde leerling telt vervolgens voor twee.

De onderzoekers menen dat de problemen van de leerlingen voortkomen uit een drietal factoren. Een stereotype en levensvreemd opgavenaanbod, een tekort schietende instructie (met name te weinig oog voor 'modellen') en een negatieve klascultuur (wiskunde is zowel saai als moeilijk).

Hoe komt het dat kinderen blind bewerkingen uitvoeren en wiskundige problemen oplossen zonder rekening te houden met de realiteit van de context? Op basis van deze centrale vraag werden enkele 'construerende onderzoeken' geanalyseerd waarin 'leeromgevingen' werden ontworpen met het doel de probleemoplossende vaardigheid van de leerlingen te vergroten door onder meer eigen inbreng te bevorderen, door interactie en door nadruk op heuristieken. De resultaten bleken weliswaar bemoedigend maar niet ten volle overtuigend zodat uitgaande van de opgedane ervaringen een nieuw onderzoek werd opgezet. Dat is met de nodige grondigheid aangepakt.

Het doel van dit onderzoek is 'leerlingen van de bovenbouw van de basisschool te stimuleren om meer actieve, planmatige en bewuste oplossters van wiskundige toepassingsproblemen te worden' (p. 41). Bovendien werd beoogd: 'de foutieve overtuigingen en de negatieve attitudes die bij vele leerlingen rond het vak wiskunde en het oplossen van vraagstukken bestaan, in positieve zin te beïnvloeden'. Dat zijn erg interessante maar geen geringe ambities, met name het beïnvloeden van attitudes is, zoals we weten, geen sinecure. Welke leerervaringen moeten de leerlingen volgens de onderzoekers opdoen om 'wiskundige toepassingsproblemen' te kunnen oplossen? Daarvoor moeten ze zich heuristieken eigen maken, zoals het gebruiken van ervaringskennis, het vormen van juiste representaties, een stappenplan leren opstellen en

uitvoeren en de uitkomst die werd verkregen interpreteren en controleren. De opgaven waarmee de leerlingen worden geconfronteerd moeten uitdagend zijn en een hoog probleemgehalte bevatten. De didactiek waarvoor wordt gekozen is sterk interactief en in de lessen wordt veel aandacht besteed aan het uitlokken van een positieve instelling ten opzichte van de wiskunde (dit wordt genoemd 'de klascultuur').

Het programma waarin dit allemaal verwezenlijkt moet worden bestaat uit 20 lessen bedoeld voor de 5^e groep (7^e in Nederland). Na de eerste introductieles volgt een vijftiental lessen waarin de leerlingen kennismaken met de oplossingsstrategie, bestaande uit vijf stappen: 1. Ik stel me het probleem voor (tekening of tabel maken) 2. Ik beslis hoe ik het probleem zal aanpakken (boomdiagram maken) 3. Ik reken uit 4. Ik interpreteer mijn uitkomst 5. Ik controleer.

Een vraag die zich hier opdringt luidt of alle leerlingen dit heuristisch stappenplan op de voet moeten volgen. Zo ja, wordt dit door vlotte en begaafde rekenaars dan niet snel ervaren als een keurslijf? Misschien zou dit creatieve oplossingsproces iets meer in liberale zin kunnen worden geïnterpreteerd. De auteurs wijzen er overigens in dit verband wel op dat sterke leerlingen de heuristieken *spontaan* zullen benutten. Maar juist de sterke leerlingen zullen heuristieken ook spontaan *negeren*, omdat ze voor hen overbodig zijn. Deze leerlingen zien 'in één klap' de oplossingsweg.

Tijdens de derde stap 'ik reken uit' komen verschillende manieren van uitrekenen onder de aandacht zoals cijferen en schatten. De leerlingen worden vervolgens geconfronteerd met antwoorden op opgaven en moeten beoordelen of die antwoorden zinvol zijn en – indien gewenst – correcties aanbrengen. In de zogenoemde projectlessen worden de leerlingen geconfronteerd met problemen die complex, levensecht en uitdagend zijn. Er is veel aandacht voor groepswork en de leraar heeft daarin duidelijk omschreven taken.

Al met al besteden de onderzoekers veel aandacht aan het ontwikkelen van een adequate didactiek als basis voor het leren oplossen van levensechte wiskundeproblemen. Van Parreren typeerde zo'n onderzoek als

formerend onderzoek (in tegenstelling tot constaterend onderzoek) en in de realistische vakdidactiek spreekt men 'ontwikkelingsonderzoek'. Verschaffel c.s. kenmerken hun onderzoek als 'construerend' onderzoek en met al deze termen is hetzelfde bedoeld namelijk dat de onderzoekers niet tevreden zijn met vaststellen wat kinderen kunnen, maar waar- toe ze onder gunstige condities in staat zijn.

Bij het effectonderzoek zijn vier experimentele en zeven controleklassen betrokken. Er werden voortoetsen, natoetsen en retentietoetsen afgenomen. Door middel van een vragenlijst gingen de onderzoekers na of de leerlingen de geleerde heuristieken hanteerden en hoe hun houding ten opzichte van wiskunde en het oplossen van wiskundige vraagstukken zich had ontwikkeld. Het initiële wiskundniveau van de kinderen in beide groepen werd gecontroleerd met behulp van een 'standaardvorderingentoets'. Met drie leerlingen uit de experimentele én de controlegroep werden mondeling nog enkele vraagstukken besproken, op de band opgenomen en vervolgens geanalyseerd. Er werd daarbij vooral gelet op het gebruik van de geleerde heuristieken en op het hanteren van metacognitieve vaardigheden. Een interessante vraag is in dit verband hoe de onderzoekers de relatie zien tussen heuristieken en metacognitieve activiteit. Mensen passen niet zomaar (experts uitgezonderd) een heuristiek toe. Dat is het resultaat van overwegingen gericht op de aard van het probleem in relatie tot wenselijke werkwijzen ('Indien dit wordt gevraagd, moet ik zus of zo handelen'). Met andere woorden, er wordt gereflecteerd op het handelen – vaak heuristisch van aard – ten dienste van het oplossen van problemen. In al die gevallen is metacognitie (reflectie) en heuristiekgebruik nauw verweven en moeilijk van elkaar te onderscheiden. Het gebruik van heuristieken kán dus heel goed het gevolg zijn van reflectie. En omgekeerd kan het gebruik van een heuristiek – met name als zich onvoorziene situaties voordoen – reflectie oproepen ('Waarom lukt dit nou niet, wat doe ik verkeerd?'). Het is niet zo dat de onderzoekers dit allemaal geheel over het hoofd hebben gezien. Als een kenmerk van 'procesbewaking' wordt immers verwezen naar het veranderen van heuristieken. Echter, op welke wijze en in welke mate

in het onderzoek met deze verwevenheid rekening is gehouden, werd mij uit de rapportage niet duidelijk. Vooral nog echter zijn de metingen niet geheel onafhankelijk van elkaar maar aan die complicatie als zodanig is geen aandacht geschonken. Een interessante vraag is wellicht ook of hoogwaardige interacties tot een toename van reflectie hebben geleid, maar ik kon daar geen gegevens in de rapportage over vinden. Er zijn echter genoeg interessante hypothesen en vragen waar de onderzoekers wél uitvoerig op ingaan.

De algemene en belangrijkste onderzoekshypothese luidt dat de krachtige leeromgeving een significante, positieve invloed zou hebben op de vaardigheid van de leerlingen in het oplossen van wiskundige toepassingsproblemen. De tweede hypothese luidt dat hetzelfde effect verwacht wordt wat betreft de houdingen en overtuigingen tegenover dit soort van problemen. De volgende hypothese luidt dat de leerlingen meer gebruik maken van wiskundige heuristieken en metacognitieve vaardigheden. Bovendien werd voorspeld dat ook zwakke leerlingen van de krachtige leeromgeving zouden profiteren. Een volgende reeks hypothesen – die ik vanwege de beperkte ruimte niet weergeef – had betrekking op de mate van implementatie van de leeromgeving door de leraren.

Wat de eerste hypothese betreft werd er inderdaad een significant verschil gevonden tussen de proef- en de controlegroep. Eveneens vond men een verschil wat betreft 'plezier en volharding' en 'visie op vraagstukken' telkens in het voordeel van de proefgroep. Hieruit kan een positieve invloed van de leeromgeving worden afgeleid. Overigens werden er telkens ook *binnen* de beide groepen aanzienlijke verschillen vastgesteld. Een middelgroot effect werd geconstateerd bij nameting op de schoolvorderingentoets (een toets die een betrouwbare indicatie geeft van het niveau van de algemene rekenvaardigheid). Dat betekent dat het volgen van het reguliere programma niet had geleden onder de arbeidsintensieve experimentele interventie.

Het effect van de leeromgeving bleek – op basis van de nameting en de retentiemeting – wat betreft het heuristiekgebruik zeer significant. Sommige heuristieken zijn echter duidelijk meer favoriet bij vooral de leerlingen uit

de proefgroep dan andere, bijvoorbeeld: 'maak een tekening', 'onderscheid noodzakelijke en overbodige gegevens' en 'maak een patroon'. Sommige heuristieken worden geheel genegeerd in de controlegroep (bijvoorbeeld: 'maak een tabel') maar ook in de proefgroep komen die slechts sporadisch voor. 'Werk met eenvoudiger getallen' is een heuristiek die door geen enkele leerling wordt toegepast. Dit grote verschil in het gebruik van heuristieken is voor de onderzoekers aanleiding om het trekken van algemene conclusies niet zinvol te achten. Wel menen ze uit kwalitatieve analyse te mogen afleiden dat de invloed op de vaardigheid problemen op te lossen groot is. Over het gebruik van metacognitieve vaardigheden kan het volgende vastgesteld. Gelet op de oriëntering is er weinig en wat betreft planning, procesbewaking en controle is er iets meer vooruitgang. In de reflectie achteraf werd weer weinig verbetering geconstateerd. Zoals we zien, is al met al de groei van de reflectievaardigheid niet indrukwekkend. Dit zou wel eens kunnen samenhangen met de, mijns inziens, niet zeer uitgesproken wijze waarop het reflecteren bij de leerlingen is gestimuleerd tijdens het experiment.

Belangrijk is dat de effecten in het algemeen werden behaald door sterke, middelmatige en (zelfs met name) de zwakkere rekenaars. De vrees dat van onderwijsvernieuwingen – en in dit geval mogen we spreken van ingrijpende interventies – steeds alleen maar de goede leerlingen profiteren, vindt dus geen empirische ondersteuning.

De leeromgeving is door alle leraren in bevredigende mate geïmplementeerd, al waren er wel onderlinge verschillen. Een duidelijk verband met de leerwinst, met het heuristiekgebruik en gebruik van reflectievaardigheden en een verband met de gewijzigde opvattingen van de leerlingen (over wiskunde) werd er niet gevonden. Wel bleek dat de leraren zich doorgaans goed hadden kunnen vinden in de opzet en het verloop van het onderwijsexperiment zoals de aard van de vraagstukken, het groepswork, het gebruik van de handleiding. De betrokken leraren – bescheiden in aantal: vier – bleken over het algemeen tevreden maar niet onkritisch terug te blikken op hun deelname aan het onderzoek. Zij meenden in didactisch

opzicht ervan geleerd te hebben, de nieuwe ideeën hadden duidelijk betekenis voor hun praktijk maar het groepswork hadden ze ervaren als moeilijk en als een grote belasting. Hierbij moet benadrukt worden dat dit onderzoek niet door de onderzoekers zelf, maar door leraren is uitgevoerd. Dat verhoogt de ecologische validiteit in hoge mate, maar tegelijkertijd brengt deze opzet extra problemen met zich mee (begeleiding van de leraren, overdracht en adoptie van ideeën enzovoort). Het verslag van het onderzoek dat hier is besproken, geeft de lezer nooit het gevoel de draad kwijt te raken en munt uit door helderheid. Kwantitatieve analyses geven betekenis aan kwalitatieve en omgekeerd. De onderzoekers maken steeds duidelijk waar het hen om gaat en ook aan tegenvallende resultaten wordt ruim aandacht besteed en dat komt de geloofwaardigheid van deze rapportage ten goede. Onderwijsonderzoeken waarin is nagedacht over een 'integrale aanpak' en waarin aan (in de woorden van deze onderzoekers) 'een gehele leeromgeving' is gesleuteld, dragen niet alleen bij aan een betere praktijk maar ook aan de onderwijskundige theorievorming. Er is niet volstaan met een enkele interventie gevolgd door een voor- en nameting, integendeel, er is veel aandacht besteed aan bijvoorbeeld de opgaventypen, aan de instructievormen, aan de vorming van heuristieken, aan groepswork en de taak van de leraar hierbij, aan de vorming van een nieuwe 'klascultuur' en aan de begeleiding van de leraren. Omdat leraren zelf het experiment uitvoerden, is dit wellicht een maximum aan variabelen, waarmee in de praktijk nog kan worden gewerkt en dat in onderzoek controleerbaar is en geen wonder dus dat niet alle verwachtingen zijn uitgekomen. Zo vraagt ook de relatie tussen heuristisch en reflectief handelen om verheldering en nadere theoretische uitwerking. De belangrijkste onderzoeksvragen kunnen echter in positieve zin worden beantwoord en dat juist ook zwakkere leerlingen van de interventie ten volle hebben geprofiteerd, vind ik nog wel het méést hoopgevend.

J.M.C. Nelissen

Mededelingen

Oraties

Mw. prof. dr. H.P.J.M. Dekkers

Op vrijdag 12 maart sprak prof. Dr. H. Dekkers haar intrede *Toegankelijkheid en effectief onderwijs* uit. Daarmee aanvaardde zij het hoogleraarschap aan de Katholieke Universiteit Nijmegen (Faculteit der Sociale Wetenschappen, Vakgroep Onderwijskunde).

Prof. dr. R.J. Bosker

Op donderdag 25 maart sprak prof. Dr. Roel Bosker zijn intrede *Ontbrekende schakels* uit. Daarmee aanvaardde hij het hoogleraarschap aan de Universiteit Twente (Faculteit der Toegepaste Onderwijskunde).

Prof. dr. P.J.C. Sleegers

Op vrijdag 25 juni sprak dr. Peter Sleegers zijn intrede *Leiding geven aan leren* uit. Daarmee aanvaardde hij het hoogleraarschap aan de Katholieke Universiteit Nijmegen (Faculteit der Sociale Wetenschappen).

Prof. dr. J.J.G. van Merriënboer

Op donderdag 1 juli sprak Jeroen van Merriënboer zijn intrede *Cognitie en multimedia-ontwerp voor complex leren* uit. Daarmee aanvaardde hij het hoogleraarschap aan de Open Universiteit Nederland.

Prof. dr. A.J.M. de Jong

Op donderdag 30 september spreekt Ton de Jong zijn intrede uit. Daarmee aanvaardt hij het hoogleraarschap aan de Universiteit Twente (Faculteit Toegepaste Onderwijskunde).

Promoties

Universiteit Twente

Promovendus: S. (Simone) Doolaard

Titel proefschrift: *Schools in change or schools in chains*

Promotoren: Prof.dr. J. Scheerens en prof.dr. R.J. Bosker

Datum: 18 februari 1999

Promovendus: M.R. (Marcel) van der Klink
Titel proefschrift: *Effectiviteit van werkplekopleidingen*

Promotor: Prof.dr. W.J. Nijhof

Datum: 26 maart 1999

Promovendus: A.A.M. (Ida) Wognum

Titel proefschrift: *Strategische afstemming en de effectiviteit van bedrijfsopleidingen*

Promotoren: Prof.dr. R.J. Bosker en

prof.dr. J. Scheerens

Datum: 9 april 1999

Promovendus: J.W.G. Geerligs

Titel proefschrift: *Design of responsive vocational education and training, a reconstruction of a systems change in agricultural education*

Promotor: Prof.dr. W. J. Nijhof

Co-promotor: dr. A. F. M.

Nieuwenhuis (Stoas)

Promovendus: C.A. van Vilsteren

Titel proefschrift: *Onderwijskundig leiderschap van schoolleiders in het voortgezet onderwijs: Schoolleiding en Vaksecties*

Promotor: Prof.dr. J. Scheerens

Katholieke Universiteit Nijmegen

Promovendus: C.H.E. (Kitty) Kwakman

Titel proefschrift: *Leren van docenten tijdens de beroepsloopbaan. Studies naar professionaliteit op de werkplek in het voortgezet onderwijs*

Promotoren: Prof.dr. J.H.G.I. Giesbers en

prof.dr. P.R.J. Simons

Co-promotor: Dr. M.J.M. Voeten

Datum: 16 maart 1999

Universiteit Utrecht

Promovendus: P.E. Bimmel

Titel proefschrift: *Training en transfer van leesstrategieën. Training in de moedertaal en transfer naar een vreemde taal, een effectstudie bij leerlingen uit het voortgezet onderwijs*

Promotoren: Prof.dr. G.J. Westhoff en

prof.dr. C.M. de Gloppe (UvA)

Co-promotor: Dr. R.J. Oostdam (UvA)

Promovendus: E.W.A. de Moor

Titel proefschrift: *Van vormleer naar realistische meetkunde. Een historisch-didactisch onderzoek van het meetkundeonderwijs aan*

kinderen van vier tot veertien jaar in Nederland gedurende de negentiende en twintigste eeuw

Promotoren: Prof.dr. A. Treffers en
prof.dr. K. van Berkel (RUG)
Co-promotor: Dr. R.A. de Jong

Rijksuniversiteit Groningen

Promovendus: J. Wolthuis

Titel proefschrift: Lower technical education in the Netherlands 1798-1993. The rise and fall of a subsystem.

Promotoren: Prof. dr. J.J.H. Dekker en
prof. dr. M.S. Archer

Promovendus: S. Waslander

Titel proefschrift: Koopmanschap en burgerschap; marktwerking in het onderwijs

Promotor: Prof.dr. J.L. Peschar

Datum: juni 1999

Afscheid

Op 15 april heeft de Universiteit Maastricht officieel afscheid genomen van prof.dr. WH.F.W. Wijnen. Omdat Wynand Wijnen een grote stempel heeft gedrukt op het onderwijs, organiseerde de capaciteitsgroep Onderwijsontwikkeling en -research ter ere van zijn afscheid een symposium over studeerbaarheid, selectie en het studiehuis. Aan de discussie, in de vorm van een onderwijsarena, namen verschillende personen uit diverse geledingen deel. Tevens werd de eerste Wynand Wijnen Prijs uitgereikt aan een instelling of een persoon die een bijzondere bijdrage heeft geleverd aan de innovatie van het onderwijs. Aan de deelnemers van het symposium is een themanummer van het *Tijdschrift voor Hoger Onderwijs* (april 1999) uitgedeeld. Dit themanummer bevat artikelen over ontwikkelingen in het onderwijs waarin Wijnen op een of andere wijze de hand heeft gehad: beoordeling en toetsing, studeerbaarheid, kwaliteitsbewaking, het studiehuis en probleemgestuurd onderwijs.

Inhoud andere tijdschriften

Tijdschrift voor Orthopedagogiek

Jaargang 38 nummer 1, 1999

Oog voor kwaliteit. Ontwikkeling van een model voor kwaliteitsevaluatie in de residentiële zorg voor personen met een handicap door B. Moes, L. Geeraert & B. van den Bruel
De nieuwe poortwachters van het speciaal onderwijs door C.M. van Rijswijk
Opvoedingsvragen en -noden van ouders met een tweeling door K. Nys, H. Colpin, L. Vandemeulebroecke & A. de Munter

Tijdschrift voor Orthopedagogiek

Jaargang 38 nummer 2, 1999

Stresstheoretische aspecten van de orthopedagogiek door J. C. van der Wolf
Commentaar op oratie J.C. van der Wolf door L. Stevens
Onderwijs aan kinderen die zijn behandeld vanwege een hersentumor door B. W. Duijzer-van Dijk & J.E.H. van Luit

Tijdschrift voor Orthopedagogiek

Jaargang 38 nummer 3, 1999

De samenwerking onderwijs – jeugdhulpverlening: de problematiek van de leerlingen door F W Louwersheimer
Alleenstaande minderjarige asielzoekers over hun opvang en begeleiding door M. Smit & M. Thomeer-Boswens
Kinderen met infantiele encephalopathie in het speciaal onderwijs door A. Kieve, L.M.E. Smit & E.J. Karel

Tijdschrift voor Orthopedagogiek

Jaargang 38 nummer 4, 1999

Honderd jaar kinderrevalidatie door J.M.H. de Moor
Kinderrevalidatie: Geneeskunde of Opvoeding? door A. Vermeer
Vroeghulp, onderzocht bij Sociaal Pedagogische Diensten door W G Sipma, B. F. van der Meulen & E.M Geerlings
Hoe spel jij gespelt? door M. van Diepen & A.M.T. Bosman

Postmoderne orthopedagogiek. door A.T.G. van Gennep
Kanttekeningen bij een rede door J. van Weelden
Taalvaardigheid van moeilijk lerende autochtone en allochtone kinderen door L. Verhoeven & A. Vermeer
Populaire kinderen hebben een Betere Werkhouding dan Verworpen Kinderen door V.A.F. Vogels, L.J. Roerink, S.E.H. de Graaf & A.M.T. Bosman

Ontvangen boeken

Baartman, H., Leij, A. van der, & Stolk, J. *Het perspectief van de orthopedagoog*. Bohn Stafleu Van Loghum, Houten/Diegem, 1999, f 79,00
Berghs, H. Kunst, *Opvoeding en ontwikkeling*. Nelissen, Baarn, 1999, f 27,50.
Biemans, H.J.A., Jongmans, C.T., Jong, F.P.C.M. de. *Het implementeren van 'leren leren' in de agrarische onderwijspraktijk*. Leerstoelgroep Agrarische Onderwijskunde, Wageningen, 1999.
Boendermaker, L. *Justitiële behandelinrichtingen voor jongeren. Populatie en werkwijze*. Garant, Leuven-Apeldoorn, 1999.
Dewey, J. (vertaald en ingeleid door Biesta, G. & Miedema, S.). *Ervaring en opvoeding*. Bohn Stafleu Van Loghum, Houten, 1999, f 34,50.
Doolaard, S. *Schools in change or schools in chains* (dissertatie). Universiteit Twente, Faculty of educational science and technology, Enschede, 1999.
Dumez, C., Kesters, M. & Ham, P. van. *Relatie(v)aar-dig*. Garant, Leuven-Apeldoorn, 1998, f 27,00.
Gent, B. van & Zee, H. van der. *Competentie en arbeidsmarkt. Een multidisciplinaire visie op ontwikkelingen rond mens en werk*. Elsevier, 's-Gravenhage, 1998, f 39,75.
Hamel, R., Elshout-Mohr, M., & Milikowski, M. (red.). *Meesterschap. Zestien stukken over intelligentie, leren, denken en probleemoplossen voor Jan J. Elshout*. Vossiuspers AUP, Amsterdam, 1999.
Heuvel, van den, -Panhuizen, M. Vermeer, H.J. *Verschillen tussen meisjes en jongens bij het vak rekenen-wiskunde op de basisschool*. Freudenthalinstituut, Utrecht, 1999.

Imelman, J.D. *De afgebroken jaargang '41-'42. Pedagogische Studiën als praktisch wetenschappelijk tijdschrift in het nauw*. Universiteit Utrecht, Utrecht, 1999.
Jumulet, H. & Teunis, C. *Kwaliteit in uitvoering. Kwaliteit in de jeugdzorg*. SWP, Utrecht, 1999.
Kalmijn, M., Bernasco, W., Weesie, J. *Huwelijks- en samenwoonrelaties in Nederland. De organisatie van afhankelijkheid*. Van Gorcum, Assen, 1999, f 45,00.
Knorth, E.J. & Smit, M. (red.). *Planmatig handelen in de jeugdhulpverlening*. Garant, Leuven-Apeldoorn, 1999, f 64,50.
Kools, Q. *Gelijk of verschillend. Motieven, ervaringen en toekomstplannen van vrouwelijke en mannelijke leerlingen in het middelbaar agrarisch onderwijs*. Landbouwniversiteit, Wageningen, 1999.
Levering, B. & Smeyers, P. (red.). *Opvoeding en onderwijs leren zien. Een inleiding in interpretatief onderzoek*. Boom, Amsterdam, 1999.
Moor, E.W.A. de. *Van vormleer naar realistische meetkunde*. CD beta Press, Utrecht, 1999, f 70,00.
Nijse, M. *Schoolkeuze en schoolloopbaan na de basisschool*. Vakgroep Psychologie, sectie persoonlijkheids- en onderwijspsychologie, Groningen, 1998.
Penninx, R. & Slijper, B. *Voor elkaar? Integratie, vrijwilligerswerk en organisaties van immigranten*. IMES (Instituut voor Migratie en Etnische Studies), Amsterdam, 1999, f 15,00.
Vries, V.M.C. de, & Jong, F.P.C.M. de. *Kwaliteit van het hoger onderwijs vanuit het perspectief van studenten. Een onderzoek naar de beleving van onderwijskwaliteit door studenten in het hoger agrarisch onderwijs*. Leerstoelgroep Agrarische Onderwijskunde, Wageningen, 1999, f 30,00.

AB-Lop: TS OCT 6441 AB-IC

Onderwijs en leeropbrengsten in de Basisvorming: een tussenbalans

Inleiding op het themanummer

A.A.M. Houtveen, E.C. Roelofs, J.N. Streumer

De Basisvorming als vernieuwing

In augustus 1993 is in de Nederlandse scholen voor voortgezet onderwijs de Basisvorming van start gegaan. Het betrof een grootschalige vernieuwing met gevolgen voor zowel de schoolorganisatie als de wijze van lesgeven. Met de Basisvorming werd een drietal algemene doelstellingen beoogd. Ten eerste uitstel van verplichte studie- en beroepskeuze, met name om ongelijkheid van kansen (naar sociaal milieu en sekse) tegen te gaan. Een gemeenschappelijk onderwijsprogramma zou hierin een sleutelfunctie moeten vervullen. Ten tweede een algehele verhoging van het peil van het jeugdonderwijs. De verhoging van het peil heeft te maken met de door de leerlingen gerealiseerde leerprestaties. De overheid heeft getracht de 'prestatielat' hoger te leggen door de inhoud van het onderwijs zo aan te passen, dat deze meer voldoet aan de eisen die de huidige samenleving stelt. Een probleem hierbij is overigens wel dat veranderingen in de leerinhoud het meten van de toe- of afname van het peil bemoeilijken. Een derde doel was de versterking van het beroepsvoorbereidend karakter van het LBO (Tweede Kamer, 1991), een doel dat kennelijk niet is gerealiseerd, getuige de invoering van het voorbereidend middelbaar beroepsonderwijs (VMBO) per 1 augustus 1999.

Ter realisering van de Basisvorming zijn destijds voor 15 vakken kerndoelen geformuleerd. Concreet betekende dit een verbreding van het vakkenpakket tijdens de periode van Basisvorming. Daarnaast geldt dat er behalve voor ontwikkeling van kennis, ook volop aandacht zou moeten zijn voor vaardigheden van verschillende aard. Gevolg zou zijn minder uitval en betere aansluiting op vervolgonderwijs.

In aanvulling op de kerndoelen heeft het ProcesManagement Basisvorming (PMB) als

een van de belangrijkste innovatie-ondersteunende activiteiten de zogeheten TVS-karakteristieken geformuleerd: toepassen, vaardigheid en samenhang (PMB, 1993a, 1993b). Deze geven de gewenste richting aan met betrekking tot het didactisch handelen. Toepassingsgericht onderwijs betekent volgens het PMB dat de te verwerven kennis levensecht en authentiek moet zijn, betekenis moet hebben voor het persoonlijk leven en voor de maatschappij. Leren door doen zou een belangrijke plaats moeten gaan innemen. Wil onderwijs toepassingsgericht zijn, dan zal een aantal vaardigheden ontwikkeld dienen te worden. Een vijftal algemene vaardigheden van zeer uiteenlopende aard wordt expliciet genoemd: een eenvoudig onderzoek verrichten, een eigen standpunt verwoorden, samenwerken aan een opdracht en de resultaten gezamenlijk presenteren, relaties leggen met de praktijk van verschillende beroepen, criteria hanteren bij het zelf beoordelen van werkstukken. Met de kernkarakteristiek 'samenhang' wordt bedoeld op het belang van een duidelijke relatie tussen de kerndoelen van verschillende vakken, door middel van vakken-integratie en het gebruik van vakoverstijgende thema's. Daarnaast wordt bedoeld op de versterking van de samenhang tussen onderwijs en begeleiding, door versterking van de relatie tussen vaklessen en mentorlessen, en tussen binnen- en buitenles-activiteiten.

Onderzoek naar de invoering van de Basisvorming

In dit themanummer wordt een tussenbalans opgemaakt als het gaat om de implementatie van de Basisvorming en de (eerste) leerresultaten. Nadrukkelijk kunnen we spreken van een tussenbalans, omdat bekend is dat vernieuwin-

gen niet in enkele jaren worden geïmplementeerd.

Veel van de onderzoeksresultaten hadden tot nu toe voornamelijk betrekking op voorwaarden voor implementatie van de Basisvorming en konden bovendien worden getypeerd als 'kleinschalig'. In dit themanummer wordt verslag gedaan van studies naar organisatorische voorwaarden en factoren op school- en klasniveau, het didactisch handelen en leeruitkomsten. Het verkregen beeld kan als representatief voor de Nederlandse situatie worden beschouwd, als we kijken naar de gehanteerde steekproeven.

De dataverzameling omtrent organisatorische voorwaarden en factoren en didactisch handelen heeft plaatsgevonden in de periode vlak na de invoering (1993-1994) tot drie jaar na de invoering van de Basisvorming (1996-1997). De studie naar leeruitkomsten is gebaseerd op data afkomstig uit VOCL-cohortstudies, waarbij leerresultaten van voor (cohort gestart in 1989) en na de Basisvorming (cohort gestart in 1993) worden vergeleken.

Van den Bergh, Peters-Sips en Zwarts doen verslag van vier deelevaluaties zoals uitgevoerd door de Inspectie op ruim 300 vestigingen voor voortgezet onderwijs. In wezen is sprake van vier deelonderzoeken die in het artikel samenvattend worden gerapporteerd. Een vrij breed scala aan onderwijskundige en schoolorganisatorische voorwaarden is in kaart gebracht die van belang zijn voor de invoering van de Basisvorming. In het artikel wordt verslag gedaan van de realisatie van een positief pedagogisch klimaat, de zorgbreedte in de Basisvorming, voorwaarden voor voorbereiding op studie en beroep en de realisatie van schoolbeleid gericht op implementatie van de Basisvorming in het jaar 1996. De Inspectie heeft voor de vier genoemde aspecten een aantal standaarden ontwikkeld, op basis van de eisen die de overheid aan de Basisvorming stelt. In belangrijke mate is daarbij gebruik gemaakt van een beleidsanalyse (Karstanje, 1996). De standaarden zijn uitgewerkt in deelaspecten, waarvoor op verschillende wijzen data zijn verzameld: documentenanalyses, observaties, interviews, en vragenlijsten. Gerapporteerd worden de uiteindelijke oorde-

len van de inspectie op de onderscheiden deelaspecten.

In de onderzoeken van respectievelijk Roelofs en Houtveen en Roelofs, Franssen, Houtveen en Lagerweij is gepoogd een operationalisering te maken van (gewenst) didactisch handelen, zoals dat voortvloeit uit de TVS-karakteristiek. De auteurs onderscheiden kenmerken van een didactiek van authentiek leren, die afgeleid kunnen worden uit recente leer- en instructietheoretische inzichten. Het geschetste raamwerk vormt de kern van moderne opvattingen over onderwijzen en leren en voor onderwijs in de Basisvorming. Vier kenmerken worden onderscheiden: 1) constructie van kennis in complete taaksituaties, 2) gerichtheid op de leefwereld van leerlingen, 3) waarde van leeractiviteiten voor situaties buiten school, 4) communicatie en samenwerking.

In een grootschalige interview- en surveystudie onder een representatieve groep docenten Nederlands en wiskunde afkomstig uit 90 scholen voor voortgezet onderwijs zijn Roelofs en Houtveen nagegaan in welke mate docenten anno 1996-1997 een didactiek van authentiek leren realiseren. Daarbij is tevens gelet op de invulling van condities op docent- en schoolniveau, die samenhangen met de realisatie van de gewenste didactiek in de basisvorming, zoals rolopvatting, deskundigheid, coördinatie en afstemming. Voorts is nagegaan in hoeverre de succesvolle realisatie van condities gepaard gaat met een frequentere realisatie van authentiek leren.

Roelofs, Franssen, Houtveen en Lagerweij onderzochten in een longitudinale dieptestudie op drie brede scholengemeenschappen de realisatie van authentiek leren door verschillende databronnen aan te boren. Zelfrapportages van docenten over hun eigen lesgedrag (via vragenlijsten) zijn aangevuld door percepties van leerlingen te bevragen. Daarnaast zijn interviews gehouden met docenten over de specifieke vormgeving van methodegebruik bij de vakken Engels en wiskunde. Authentiek leren betekent bij moderne vreemde talen een nadruk op de communicatieve functie van taal. Bij wiskunde gaat het om contextrijk leren. Ook in deze studie is ingegaan op de vervulling van condities voor authentiek leren op klas- en schoolniveau.

In de vierde bijdrage van dit themanummer gaan Van der Werf, Lubbers en Kuyper in op de vraag of de onderwijsopbrengsten van het voortgezet onderwijs zijn verbeterd sinds de invoering van de Basisvorming. Ook wordt nagegaan of sinds de Basisvorming de groepsgebonden ongelijkheid van kansen is toe- of afgenomen. Gedoeld wordt op verschillen in prestatie en bereikte onderwijspositie samenhangend met sekse, sociaal-economische status en etniciteit. Daarmee wordt ingegaan op de vraag of de Basisvorming tot een algehele verhoging van het peil van jeugdonderwijs komt in die zin dat alle leerlingen er gelijkelijk van profiteren. Van der Werf et al. baseren zich in hun uitspraken op een analyse van data afkomstig uit twee VOCL-cohorten, met beide een omvang van aanvankelijk bijna 20.000 leerlingen. Vergeleken worden leerlingen die in 1989 startten in het voortgezet onderwijs en leerlingen die dat in 1993 deden, ten tijde van de invoering van de Basisvorming.

Literatuur

- Karstanje, P.N. (1996). *Beleids Theorie basisvorming. Een proeve van reconstructie*. Amsterdam: SCO-Kohnstamm Instituut.
- Procesmanagement Basisvorming (1993a), *Van kerndoel naar praktijk, deel A*. Almere: Procesmanagement Basisvorming.
- Procesmanagement Basisvorming (1993b), *TVS-karakteristieken, deel B*. Almere: Procesmanagement Basisvorming.
- Tweede Kamer (1991). *Wijziging van onder meer de Wet op het voortgezet onderwijs en de Wet op het basisonderwijs in verband met de invoering van basisvorming in het voortgezet onderwijs, de invoering van kerndoelen in het basisonderwijs en voortgezet onderwijs en de invoering van het voorbereidend beroepsonderwijs* (kamerstuk 20381, nr. 141). 's Gravenhage: Staatsuitgeverij.

Deelstudies in het kader van de evaluatie van de Basisvorming

H. van den Bergh, M. Peters-Sips en M. Zwarts

Samenvatting

De Tweede Kamer krijgt najaar 1999 van de Staatssecretaris van Onderwijs een evaluatierapport over de invoering van de basisvorming. De Inspectie schrijft het rapport op basis van eigen onderzoek en onderzoek van anderen.

In 1996 voerde de inspectie de eerste deelevaluaties uit op 330 scholen. Deze evaluaties betroffen: 'pedagogisch klimaat', 'schoolbeleid', 'voorbereiding op studie en beroep' en 'zorgbreedte'. De inspectie bestudeerde schooldocumenten, voerde gesprekken met schoolleiders, middenmanagers en leerlingen, bezocht lessen en bevroeg leerlingen en leraren schriftelijk.

De belangrijkste conclusies uit deze eerste deelevaluaties zijn:

- Op de meeste scholen is sprake van een goed school- en klassenklimaat.
- Een ruime meerderheid van de scholen kent voldoende zorgbreedte.
- De determinatie van de leerlingen verloopt meestal zorgvuldig, maar slechts de helft van de scholen biedt goede voorwaarden voor de oriëntatie op studie en beroep.
- Het schoolbeleid gericht op de basisvorming is maar op de helft van de scholen voldoende bevonden.

Inleiding

Van meet af aan is de behoefte aan evaluatie van de invoering van de basisvorming kenbaar gemaakt. Reeds in het wetsvoorstel Basisvorming, dat in het zittingsjaar 1987-1988 naar de Tweede Kamer werd gezonden, was sprake van een evaluatieparagraaf. In 1993 heeft het Ministerie van Onderwijs en Wetenschappen het Platform Evaluatie Basisvorming ingesteld, waarin de verschillende betrokkenen bij de evaluatie (SVO, Cito, PMB en inspectie) participeerden. In vervolg daarop wordt het onderzoek 'Beleidsanalyse basisvorming' aanbe-

steed via SVO (Karstanje, 1996). In 1995 maakt de staatssecretaris Mw. T. Netelenbos de inspectie verantwoordelijk voor de inhoudelijke coördinatie en de voortgangsbewaking van de uitvoering van het evaluatieprogramma basisvorming. Het door de inspectie opgestelde 'Plan voor de evaluatie van de basisvorming' wordt in februari 1997 door de staatssecretaris aangeboden aan het Onderwijsoverleg Primair en Voortgezet Onderwijs (POVO). Op 24 maart 1997 wordt het evaluatieplan tijdens het POVO-overleg besproken. Vervolgens wordt het plan op 12 mei 1997 ter informatie naar de Tweede Kamer gezonden. Op 1 juli verzoekt de Vaste Kamer Commissie (VKC) de staatssecretaris om een aantal vragen te beantwoorden. Op de schriftelijke beantwoording volgt geen mondeling overleg meer. Met de acceptatie van het plan wordt opnieuw vastgelegd dat de evaluatie van de basisvorming in het najaar van 1999 in de Tweede Kamer aan de orde zal zijn.

Terwijl deze formele procedures liepen, is de inspectie los daarvan al in 1993 begonnen met systematisch aandacht te geven aan de invoering van de basisvorming tijdens de schoolbezoeken. Zij rapporteerde daarover jaarlijks in het Onderwijsverslag. In 1995 is een plan opgesteld voor de door de inspectie zelf uit te voeren evaluatie-onderzoeken. Er werden onderzoeken geprogrammeerd en in gang gezet op het gebied van schoolorganisatorische aspecten van de basisvorming, op het gebied van het onderwijsleerproces en op het gebied van de opbrengsten. Vanwege de omvang van de vernieuwing van het onderwijs in de onderbouw is de evaluatie over een langere periode gespreid, namelijk van 1996 tot het najaar van 1999.

De eerste deelevaluaties door de inspectie betroffen vier organisatorische aspecten van de basisvorming: schoolbeleid, pedagogisch klimaat, voorbereiding op studie en beroep en zorgbreedte. De dataverzameling vond als

volgt plaats: iedere inspecteur bezoekt met een centraal ontwikkeld instrumentarium via een aselect getrokken steekproef aangewezen scholen uit het eigen ambtsgebied.

In het schooljaar 1997-1998 werden de deelevaluaties met betrekking tot het onderwijsleerproces uitgevoerd, waaronder: het onderwijsaanbod (de mate waarin kerndoelen en algemene vaardigheidsdoelen onderwezen worden), het (vak)didactisch en pedagogisch handelen van de docent, de wijze van instructie en differentiatie, de taken en het beleid van vaksecties, etc.

Tenslotte wordt in 1999 de evaluatie van de opbrengsten van de basisvorming (de leerlin- genresultaten) afgerond.

In dit artikel wordt verslag gedaan van de vier evaluatie-onderzoeken m.b.t. de genoemde organisatorische aspecten, die de inspectie in 1996 uitvoerde. In vier inspectierapporten (Inspectie van het onderwijs, 1997a, 1997b, 1997c, 1997d) is verslag gedaan van resultaten van deze evaluaties. Een samenvattende publieksversie van deze rapporten werd in september 1997 naar de scholen voor voortgezet onderwijs gezonden (Inspectie van het onderwijs, 1997e).

De volgende vragen worden achtereenvolgens beantwoord:

- In welke mate zijn de aspecten: pedagogisch klimaat, schoolbeleid, voorbereiding op studie en beroep, en zorgbreedte in de basisvorming gerealiseerd? (2.1 en 2.2)
- In welke mate verschillen de scholen op deze aspecten? (2.3)
- Is er samenhang tussen deze aspecten? (2.4)
- Is er een relatie tussen eigenschappen van de scholen of de schoolomgeving en de realisatie van de genoemde aspecten van de basisvorming (3)?

1 Opzet

1.1 Constructie van toetsingskader en instrumenten

Voor deze vier onderzoeken heeft de inspectie per deelevaluatie een toetsingskader ontwikkeld: een referentiekader voor de beoordeling van de kwaliteit van de betreffende aspecten in de basisvorming. Een toetsingskader bestaat uit

standaarden, indicatoren en normen. In de standaarden zijn de kwaliteitseisen geformuleerd, de indicatoren vormen de operationalisatie van de standaarden en de normen bevatten de richtlijnen voor de beoordeling van de (mate van) realisatie van de standaarden. Elk toetsingskader is uitgewerkt in instrumenten die inspecteurs hanteren tijdens de voorbereiding en uitvoering van het schoolbezoek. (Zie voor een beschrijving van deze inspectiewerkwijze: Janssens, 1997).

Voor de selectie van de standaarden is in eerste instantie uitgegaan van de eisen die overheid aan de basisvorming stelt. Daarvoor is onder meer gebruik gemaakt van een beleidsanalyse van Karstanje (1996). Voorts is gebruik gemaakt van kennis die de onderwijswetenschappen hebben opgeleverd en van kennis van de mogelijkheden van de onderwijspraktijk. De standaarden zijn opgesteld door teams van inspecteurs, voorgelegd aan externe deskundigen (uit onderwijswetenschap en -onderzoek, onderwijspraktijk, onderwijs-ondersteuning en -begeleiding en procesmanagement) en gecheckt op een aantal formele kenmerken als volledigheid, consistentie en onafhankelijk. De standaarden die uiteindelijk zijn gehanteerd, zijn opgenomen in Tabel 1.

De standaarden zijn, zoals gezegd, geoperationaliseerd in indicatoren: variabelen die een indicatie vormen voor het realiseren van de standaarden. Omdat gewerkt wordt met inspecteurs van het onderwijs, zijn de indicatoren op een hoog inferentieniveau geformuleerd. Een voorbeeld van een indicatorenset is weergegeven in Tabel 2.

De sets met indicatoren zijn opgenomen in instrumenten (richtlijnen voor documentanalyse, observatie-, interview- en schriftelijke vragenlijsten) waarmee het mogelijk is de aanwezigheid van de indicatoren te scoren.

1.2 Uitvoering

Bij alle deelevaluaties heeft documentanalyse plaatsgevonden en zijn gesprekken met schoolleiders en andere leden van het schoolmanagement gevoerd. Verder is gesproken met leerlingen (zorgbreedte, voorbereiding op studie en beroep, pedagogisch klimaat), zijn klassen bezocht (pedagogisch klimaat) en zijn leerlingen en leraren schriftelijk bevraagd (pedago-

Tabel 1*Een overzicht van de standaarden voor de vier deelevaluaties**Pedagogisch klimaat*

- P1 Er zijn condities voor een goed schoolklimaat
- P2 De mening van docenten over het pedagogisch klimaat
- P3 De mening van leerlingen over het pedagogisch klimaat
- P4 Het pedagogisch handelen van de leraren bevordert een goed klassenklimaat
- P5 De leraren bevorderen in hun lessen een actief leerproces bij de leerlingen
- P6 De leerlingen nemen actief aan het leerproces deel

Schoolbeleid

- S1 De schoolleiding heeft haar visie op het onderwijskundig beleid ten aanzien van de basisvorming geëxpliciteerd
- S2 De schoolleiding coördineert de organisatie van het onderwijskundig beleid ten aanzien van de basisvorming
- S3 De schoolleiding coördineert de inhoud van het onderwijskundig beleid ten aanzien van de basisvorming
- S4 De schoolleiding evalueert de organisatie van het onderwijskundig beleid ten aanzien van de basisvorming
- S5 De schoolleiding heeft haar personeelsbeleid en haar financieel en materieel beleid afgestemd op haar onderwijskundig beleid ten aanzien van de basisvorming

Voorbereiding op studie en beroep

- V1 Door planning, evaluatie en bijsturing van de Oriëntatie op Studie en Beroep (OSB) wordt de doeltreffendheid vergroot
- V2 In de schoolorganisatie zijn de voorwaarden voor een adequate inrichting van OSB gerealiseerd
- V3 De leraren besteden in de vakken voldoende aandacht aan de relatie tussen vakinhouden en de praktijk van de verschillende beroepen
- V4 De leerlingen worden voldoende voorbereid om adequaat keuzes te kunnen maken
- V5 De keuzes voor de determinatie van een schoolsoort zijn goed geregeld en vinden zorgvuldig plaats
- V6 De keuzes voor de determinatie van een afdeling zijn goed geregeld en vinden zorgvuldig plaats
- V7 De adviezen die de school aan het einde van het tweede leerjaar over het door de leerling verder te volgen onderwijs verstrekt, komen zorgvuldig tot stand
- V8 Selectie en determinatie vormen geen belemmering voor een optimale doorstroming

Zorgbreedte

- Z1 De school laat leerlingen zorgvuldig toe
- Z2 Om te kunnen voldoen aan extra onderwijsbehoeften van leerlingen maakt de school gebruik van adequate groeperingsvormen
- Z3 Het onderwijs voor leerlingen met extra onderwijsbehoeften kenmerkt zich door differentiatie in inhoud, omvang, methodiek en tijd
- Z4 De school heeft een adequaat leerlingbegeleidingssysteem
- Z5 De condities op school zijn gunstig voor onderwijs-op-maat
- Z6 Het zorgbreedte beleid van de school is gericht op het bereiken van aantoonbare effecten van onderwijs-op-maat

Tabel 2*Standaard V7 met vijf gehanteerde indicatoren*

De adviezen die de school aan het einde van het tweede leerjaar over het door de leerlingen verder te volgen onderwijs verstrekt, komen zorgvuldig tot stand.

- 1 Alle leerlingen ontvangen een schriftelijk advies
- 2 Bij alle leerlingen is dit advies onderbouwd
- 3 Het advies wordt gebaseerd op systematisch verzamelde gegevens over prestaties, houdingen en interesses
- 4 Het advies is in samenspraak met leraren, mentor en/of decaan tot stand gekomen
- 5 Voorafgaand aan het advies is gesproken met leerlingen en ouders

gen en leraren schriftelijk bevraagd (pedagogisch klimaat).

Het evalueren van één aspect aan één school door één inspecteur nam één dagdeel in beslag, met uitzondering van pedagogisch klimaat. Aan deze evaluatie werd per school door één inspecteur één gehele dag besteed.

Op basis van de scores op de indicatoren is door de inspecteurs de realisatie van de standaarden beoordeeld. Daarvoor is gebruik gemaakt van een vierpuntsschaal die ontleend is aan de Schotse inspectie van het onderwijs (HM Inspectors of Schools, 1995). De kwalifi-

caties op de schaal zijn: 'vooral zwak', 'zwakke overheerst', 'sterkte overheerst' en 'vooral sterk'. Voor het onderscheid tussen 'zwakke overheerst' en 'sterkte overheerst' is een normeringsvoorschrift opgesteld. Daarvoor is gebruik gemaakt van een variant van de Ang-off-procedure; een procedure voor 'standard setting' waarbij per indicator aangegeven wordt waar scholen minimaal aan moeten voldoen (Van der Linden, Meijer & Vos, 1997). Inspecteurs en in beperkte mate ook leraren, leden van het schoolmanagement en deskundigen van buiten de scholen hebben in het kader van deze procedure normeringsvoorstellen gedaan. Bij de uitvoering van de evaluatie gold voor de inspecteurs de opgestelde normen als uitgangspunt; een beargumenteerde afwijking hiervan was echter mogelijk.

Om met de instrumenten te leren werken, kregen de inspecteurs een theoretische en een praktische scholing. De theoretische scholing vond plaats via uiteenzettingen door deskundigen gevolgd door discussie en zelfstudie. De praktische scholing werd opgedaan tijdens schoolbezoek in tweetallen (zogenoemde duo-bezoeken). Ook in de uitvoeringsfase zijn duo-bezoeken afgelegd; in dat geval met name vanwege de wenselijkheid zicht te krijgen op de interbeoordelaarsbetrouwbaarheid.

1.3 Steekproef

Voor de vier evaluaties is een gestratificeerde steekproef van 330 scholen of erkende nevenvestigingen getrokken. (Een erkende nevenvestiging is een nevenvestiging zoals bedoeld in de artikelen 75 en 75a van de WVO. In het ver-

volg worden, tenzij anders vermeld, met scholen zowel zelfstandige scholen als erkende nevenvestigingen bedoeld.) De strata zijn ingedeeld naar schoolsamenstelling (smal met VBO, smal zonder VBO en breed) en naar zelfstandigheid (zelfstandig of erkende nevenvestiging). Om de betrokken scholen niet overmatig te belasten zijn op elke school twee random aangewezen aspecten geëvalueerd. Dit leverde 55 scholen per combinatie van aspecten op en 165 scholen per aspect. Uit Tabel 3 is af te lezen dat deze opzet grotendeels is gerealiseerd. Slechts twee scholen zijn geheel uitgevallen, bij elf scholen ontbreken de waarnemingen op één of meer aspecten. Voor twee combinaties zijn alle beoogde 55 scholen geëvalueerd. Voor de andere vier combinaties ontbreken twee tot vier scholen.

De gekozen steekproefopzet heeft ontegenzeggelijk voordelen waar het gaat om de afnameprocedure, in het bijzonder de belasting voor scholen en inspecteurs. Echter, data-analytisch zijn er ook nadelen aan verbonden: de correlatiematrix tussen de oordelen op de standaarden van de vier deelprojecten als geheel is minder makkelijk te geven of te maken. Zeker, wanneer blijkt dat de realisatie van de steekproef minder systematisch oogt dan de planning dient eerst nagegaan te worden of de verschillende substeekproeven opgevat mogen worden als één grote steekproef uit één populatie.

De verschillen tussen steekproeven kunnen op verschillende manieren getoetst worden. Er is gekozen voor een relatief eenvoudige variantie-analytische wijze (Van den Bergh, Peters-Sips & Zwarts, 1998). Hierbij worden de ver-

Tabel 3

Een overzicht van de gerealiseerde aantallen oordelen en scholen per combinatie van aspecten

Combinatie van aspecten	Aantal scholen
Zorgbreedte en Voorbereiding op studie en beroep	55
Zorgbreedte en Pedagogisch klimaat	55
Zorgbreedte en Schoolbeleid	53
Voorbereiding op studie en beroep en Pedagogisch klimaat	52
Voorbereiding op studie en beroep en Schoolbeleid	51
Schoolbeleid en Pedagogisch klimaat	51
Schoolbeleid	2
Pedagogisch klimaat	4
Voorbereiding op studie en beroep	4
Zorgbreedte	1
Totaal	328

Tabel 4

Een overzicht van de standaarden per aspect

Standaarden	aantal indicatoren	IB	B	V
<i>Pedagogisch klimaat</i>				
P1 condities voor goed schoolklimaat	7	.78	.52	.73
P2 mening van leraren over pedagogisch klimaat	28 ¹	—	.89	—
P3 mening van leerlingen over pedagogisch klimaat	30 ¹	—	.90	—
P4 leraren bevorderen goed klassenklimaat	8	.91	.38	.55
P5 leraren bevorderen actief leren	9	.78	.62	.74
P6 leerlingen nemen actief aan leerproces deel	8	.84	.62	.66
<i>Schoolbeleid</i>				
S1 explicitering van beleidsvisie op het onderwijs	7	.94	.59	.79
S2 coördinatie van beleidsuitvoering	5	.94	.65	.86
S3 coördinatie inhoud onderwijskundig beleid	7	.94	.76	.85
S4 evaluatie van onderwijskundig beleid	7	.87	.81	.88
S5 personeels-, financieel en materieel beleid	5	.87	.62	.85
<i>Vorbereiding op studie en beroep</i>				
V1 planning, evaluatie en bijsturing van OSB	4	.95	.64	.90
V2 voorwaarden voor OSB gerealiseerd	6	.93	.53	.85
V3 relatie vakinhoud en praktijk gelegd	4	.90	.62	.86
V4 voorbereiding op adequaat kiezen	5	.98	.48	.80
V5 keuze van schoolsoort zorgvuldig	5	.98	.71	.68
V6 keuze van afdeling zorgvuldig	4	.98	.71	.65
V7 adviezen einde tweede leerjaar zorgvuldig	5	.95	.67	.90
V8 geen belemmering doorstroming	8	.95	— ²	.53 ³
<i>Zorgbreedte</i>				
Z1 zorgvuldige toelating van leerlingen	8	1.00	.40	.70
Z2 adequate groeperingsvormen	5	.83	0	.61
Z3 interne differentiatie	5	.92	.13	.86
Z4 adequaat leerlingbegeleidingssysteem	7	.94	.50	.83
Z5 condities voor onderwijs-op-maat	4	.92	.24	.81
Z6 op effecten gericht beleid	4	.94	.37	.87

IB beoordelaarsovereenstemming

B coëfficiënt α

V correlatie tussen somscore op items en oordeel

¹ vragenlijsten waarvoor geen expliciet oordeel is gevraagd² bij deze standaard zijn geen indicatoren gebruikt, maar een karakterisering van de doorstroming³ correlatie tussen schaaftscore en oordeel

schillen tussen (een combinatie van) aspecten opgevat als afwijking van een algemeen gemiddelde. Vervolgens is per combinatie nagegaan of het gemiddelde en de variantie afwijken van het algemeen gemiddelde van de oordelen over alle deelprojecten tesamen en de variantie hiervan.

De verschillende combinaties van aspecten hebben in hoge mate vergelijkbare oordelen opgeleverd, zodat de verschillende deelsteekproeven als uit één populatie afkomstig beschouwd kunnen worden. Waar zinvol zijn derhalve de evaluatieresultaten die in dit artikel worden gepresenteerd gebaseerd op de schattingen voor de 330 scholen in plaats van op de 165 scholen, die per aspect geëvalueerd zijn.

1.4 Kwaliteit van het instrumentarium

Om de kwaliteit van de instrumenten na te gaan zijn de overeenstemming tussen de beoordeelaars, de betrouwbaarheid en de validiteit vastgesteld. Gedetailleerde informatie staat in een aantal technische rapporten (Van der Grint & Plomp, 1997; Van den Donk & Visser, 1997; van de Wal & Zwarts, 1997a, 1997b; Van den Bergh & Zwarts, 1998). De belangrijkste gegevens zijn opgenomen in Tabel 4.

De overeenstemming tussen de beoordeelaars is vastgesteld door de Gower-coëfficiënt (Heuvelmans & Sanders, 1993). Voor dichotome data komt deze coëfficiënt overeen met de proportie overeenkomende oordelen, bij polytome data wordt rekening gehouden met

het aantal scoringscategorieën. De overeenstemming tussen de oordelen die zijn opgenomen in Tabel 4 varieert van .78 tot 1.00 met een gemiddelde van .92. Over het geheel is dit goed te noemen.

De betrouwbaarheid is vastgesteld met behulp van Cronbach's α . De betrouwbaarheid van de somscores is over het algemeen laag, met name bij het aspect zorgbreedte waar zelfs een uitschieter met de waarde 0 voorkomt. Daarbij moet echter met twee factoren rekening worden gehouden. In de eerste plaats zijn de indicatoren niet bedoeld als parallelle metingen van een zelfde construct. Daardoor is de heterogeniteit vrij hoog en vormt coëfficiënt α een onderschatting van de betrouwbaarheid. In de tweede plaats is het aantal indicatoren over het algemeen laag. Te laag voor het bepalen van een schaalwaarde, maar voldoende voor het indiceren van de realisatie van een standaard door een ervaren inspecteur. Dit neemt

niet weg dat de resultaten op de standaarden P4, Z2, Z3 en Z5 met de nodige voorzichtigheid gehanteerd moeten worden.

Als indicatie voor de validiteit is de correlatie vastgesteld tussen de somscore op de indicatoren en het oordeel. Over het geheel genomen is deze bevredigend, ook in de situaties waarin Cronbach's α aan de lage kant is (zie Tabel 4). De hoogte van de correlaties wordt voor een deel verklaard door het verband dat gelegd is door middel van de normering. Ook als daarmee rekening wordt gehouden, blijkt er een duidelijk verband te zijn tussen de indicatoren en het oordeel.

2 Analyse en resultaten

2.1 Beoordeling van de vier aspecten

Na raadpleging van alle bronnen zijn door de inspecteurs per school, volgens de vooraf vast-

Tabel 5
Verdeling van de oordelen per standaard (percentages)

standaard	vooral zwak	zwakte overheerst	sterkte overheerst	vooral sterk
<i>Pedagogisch klimaat</i>				
P3 mening van leerlingen over pedagogisch klimaat ¹	-	-	-	-
P2 mening van leraren over pedagogisch klimaat ¹	-	-	-	-
P4 leraren bevorderen goed klassenklimaat	0	5	73	23
P1 condities voor goed schoolklimaat	1	14	69	16
P5 leraren bevorderen actief leren	4	37	56	3
P6 leerlingen nemen actief aan leerproces deel	22	51	27	0
<i>Schoolbeleid</i>				
S2 coördinatie van beleidsuitvoering	12	17	47	24
S1 explicitering van beleidsvisie op het onderwijs	6	30	43	21
S5 personeels-, financieel en materieel beleid	11	27	42	20
S3 coördinatie inhoud onderwijskundig beleid	13	33	41	14
S4 evaluatie van onderwijskundig beleid	34	34	24	9
<i>Voorbereiding op studie en beroep</i>				
V6 keuze van afdeling zorgvuldig	4	11	46	39
V5 keuze van schoolsoort zorgvuldig	5	12	50	33
V8 geen belemmering doorstroming	4	17	65	14
V7 adviezen einde tweede leerjaar zorgvuldig	10	15	46	29
V4 voorbereiding op adequaat kiezen	6	20	56	18
V2 voorwaarden OSB gerealiseerd	11	29	45	15
V3 relatie vakinhoud en praktijk gelegd	21	36	36	7
V1 planning, evaluatie en bijsturing van OSB	25	42	22	12
<i>Zorgbreedte</i>				
Z1 zorgvuldige toelating van leerlingen	0	4	55	42
Z5 condities voor onderwijs-op-maat	2	9	63	26
Z2 adequate groeperingsvormen	2	11	73	14
Z4 adequaat leerlingbegeleidingssysteem	2	13	54	32
Z6 op effecten gericht beleid	5	22	54	18
Z3 interne differentiatie	8	29	52	10

¹ vragenlijsten waarvoor geen expliciet oordeel is gevraagd

gelegde normering, oordelen uitgesproken over de realisatie van de standaarden. In Tabel 5 is de verdeling van de oordelen op de gehanteerde schaal weergegeven, zoals die voor de 165 bezochte scholen gegeven zijn. De volgorde van de aspecten in de tabel is willekeurig. De standaarden zijn per aspect op volgorde gezet, dat wil zeggen dat steeds de standaarden met het hoogst aantal 'voldoende' oordelen ('sterkte overheerst' plus 'vooral sterk') bovenaan staan.

De standaarden voor een goed pedagogisch klimaat zijn in het algemeen positief beoordeeld. Met name de condities voor een goed schoolklimaat en het pedagogisch handelen van leraren kregen een hoge waardering. Uitgesproken laag was de waardering voor de mate waarin leerlingen actief aan het leerproces deelnemen.

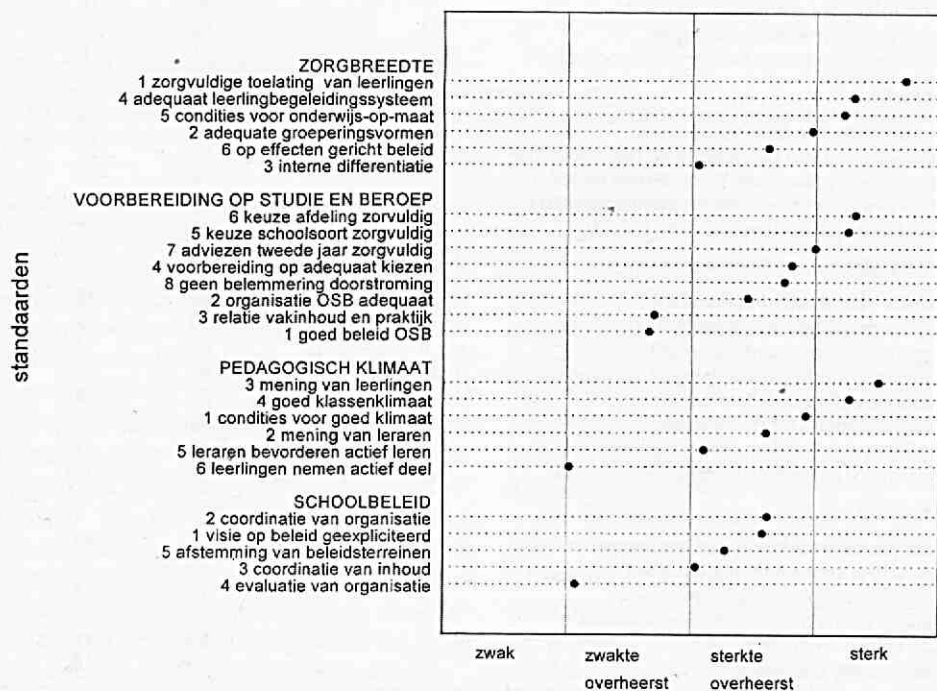
Het schoolbeleid is in het algemeen veel minder positief beoordeeld. De coördinatie van de organisatie van het onderwijskundig beleid werd op ongeveer zeventig procent van de scholen positief gewaardeerd en de explicitering van de visie op het onderwijskundig beleid op zo'n vijf procent minder. De evaluatie van de organisatie van het onderwijskundig beleid werd bij tweederde van de scholen onvoldoende gevonden.

De helft van de standaarden aan de hand waarvan een oordeel werd uitgesproken over de voorbereiding op studie en beroep heeft betrekking op het determinatieproces, de andere helft op de oriëntatie op studie en beroep (OSB). Op meer dan driekwart scholen werd het determinatieproces positief beoordeeld. OSB kreeg minder waardering van de inspecteurs, met name wordt slechts op éénderde van de scholen planning, evaluatie en bijstelling van OSB gebruikt om de doeltreffendheid van OSB te verhogen.

Zorgbreedte is het meest positief beoordeeld van de vier aspecten. De toelating van leerlingen, het leerlingbegeleidingssysteem, de condities voor onderwijs-op-maat en de gehanteerde groeperingsvormen kregen een positieve waardering op zo'n negentig procent van de scholen. De interne differentiatie kreeg slechts op iets meer dan de helft van de scholen een voldoende.

2.2 Vergelijking van de oordelen op de standaarden

Omdat de verschillende deelsteekproeven als uit één populatie afkomstig beschouwd kunnen worden, konden de gemiddelde oordelen op alle standaarden met elkaar vergeleken worden. In Figuur 1 is het gemiddelde oordeel per



Figuur 1. Gemiddeld oordeel per standaard

standaard grafisch weergegeven. Daarbij is van de volgende codering gebruik gemaakt:

- 1 vooral zwak
- 2 zwakte overheerst
- 3 sterkte overheerst
- 4 vooral sterk.

Een gemiddelde van 2 betekent dus dat gemiddeld genomen het oordeel 'zwakte overheerst' is gegeven; 3 betekent dat gemiddeld 'sterkte overheerst' is geoordeeld.

In feite worden de resultaten in Figuur 1 op een wat andere wijze gepresenteerd dan in Tabel 5; in de figuur zijn de standaarden zijn nl. gesorteerd op de gemiddelde score (van hoog naar laag). Zo blijkt uit Figuur 1, dat de meeste standaarden boven de 2.5 (de grens tussen 'zwakte overheerst' en 'sterkte overheerst') 'scoren' en dus aan de positieve kant van de schaal zitten.

Bij zorgbreedte geldt dat voor alle standaarden. Daarbij valt echter weer op dat de randvoorwaarden (toelating, leerlingbegeleidingssysteem en condities) beter beoordeeld zijn dan de uitvoering (groeperingsvormen, gerichtheid op effecten en interne differentiatie).

Bij voorbereiding op studie en beroep vallen twee standaarden (relatie van vakinhoud met de praktijk en een goed beleid voor OSB) onder de grens van 2.5. Gemiddeld is het oordeel daarop 'zwakte overheerst'. Beide standaarden hebben betrekking op OSB; een onderdeel dat

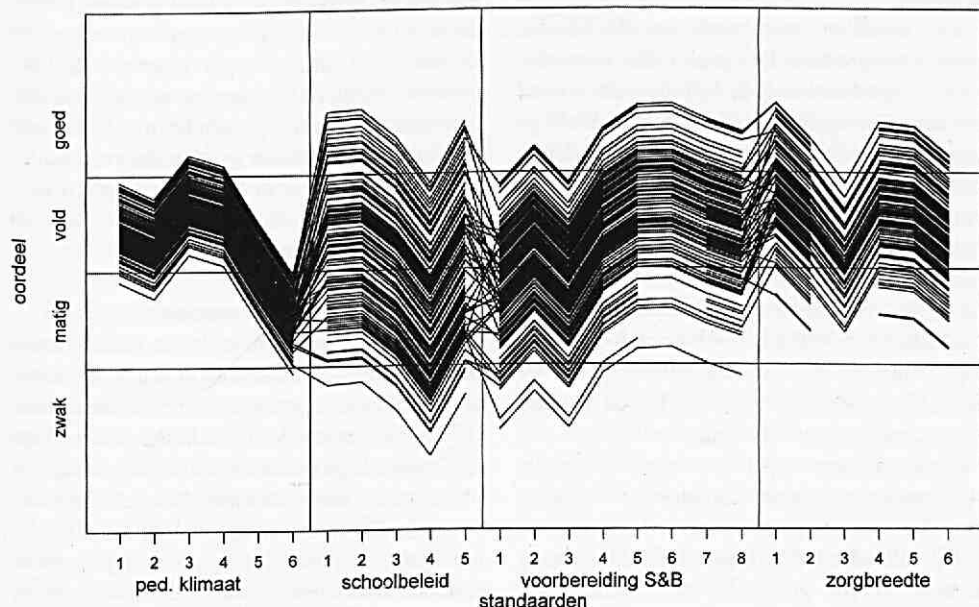
met de basisvorming is ingevoerd. Ook de derde standaard die op OSB betrekking heeft (condities voor OSB gerealiseerd) krijgt een relatief zwakke beoordeling. De standaarden die betrekking hebben op selectie en determinatie worden daarentegen overwegend positief gewaardeerd.

Bij Pedagogisch klimaat worden twee standaarden relatief matig gerealiseerd. Beide hebben betrekking op actief leren, de een op het bevorderen van actief leren door de leraar, de ander op het actief leren door de leerlingen. De aspecten die betrekking hebben op het Pedagogisch klimaat in engere zin worden duidelijk beter gerealiseerd.

Bij Schoolbeleid is de standaard: 'de evaluatie van de organisatie van de basisvorming', het zwakst. Ook de andere standaarden worden echter maar matig gerealiseerd.

2.3 Verschillen tussen scholen

Ook al is er, zoals bij de meeste standaarden, een duidelijke omschrijving van de schaalpunten, toch zegt het gemiddelde lang niet alles. Van even groot belang zijn de verschillen tussen scholen. Zijn slechts enkele scholen als zeer 'matig' beoordeeld, terwijl de meeste als 'voldoende' of misschien zelfs 'ruim voldoende' beoordeeld zijn? Of, zijn er zowel veel scholen die tekort schieten op verschillende standaarden, als scholen die van een meer dan



Figuur 2. De verschillen tussen de scholen per standaard

gewone inspanning blijkt gegeven hebben?

Om deze vragen te beantwoorden zijn de gemiddelden van alle scholen geschat met een meerniveaumodel (Van den Bergh, Peters-Sips & Zwarts, 1998). In dit model vormen de vestigingen en klassen de randomeffecten. De gemiddelden zijn weergegeven in Figuur 2.

In Figuur 2 stelt iedere lijn een school voor. Omdat elke school slechts op twee aspecten is geëvalueerd, zijn de lijnen soms onderbroken. De figuur is in vier horizontale banden verdeeld op de punten 1.5, 2.5, en 3.5, overeenkomend met de kwalificaties op de oordelen. (In de figuur zijn de kwalificaties: 'vooral zwak', 'zwakte overheerst', 'sterkte overheerst' en 'vooral sterk' vervangen door kortere omschrijvingen: 'zwak', 'matig', 'vold(oende)' en 'goed'.) Uit Tabel 5 is af te leiden hoeveel procent van de scholen binnen deze banden valt.

In Figuur 2 valt als eerste op dat de lijnen elkaar betrekkelijk weinig kruisen. De volgorde van de scholen op de verschillende standaarden is binnen een bepaald aspect over het algemeen hetzelfde en wisselt ook niet sterk tussen aspecten. Verder blijken er voor de meeste van de 25 standaarden flinke verschillen tussen scholen. Bij enkele standaarden van pedagogisch klimaat is de variatie veel kleiner. Dit is een gevolg van het feit dat bij deze standaarden het oordeel per school het gemiddelde is van de beoordeling van de bezochte klassen en lessen.

De oordelen voor pedagogisch klimaat lopen weinig uiteen. Dat geldt in het bijzonder voor de standaarden 1, 4, 5 en 6, waar vrijwel alle oordelen dicht bij elkaar liggen. Voor de laatste drie is dit een gevolg van het feit dat de schooloordelen klassengemiddelden zijn (zie Van den Bergh & Zwarts, 1998). Voor standaard 1 'condities voor goed schoolklimaat' geldt dat echter niet: voor veel scholen geldt dat sterkte overheerst. Ook voor het klassenklimaat liggen vrijwel alle oordelen aan de sterke kant. Voor de andere standaarden geldt dat niet. Bij 'actief leren' van leerlingen worden alle scholen negatief beoordeeld. Ook bij 'bevorderen' van actief leren door leraren is bijna een kwart van de scholen aan de zwakke kant.

Bij schoolbeleid zijn de verschillen tussen de scholen erg groot. Bij alle standaarden komen alle of bijna alle kwalificaties voor.

Daarbij komen vrij veel negatieve scores voor: een kwart (standaard 1) tot drie kwart (standaard 4) van de scholen zijn volgens het oordeel van de inspecteurs aan de zwakke kant.

Bij voorbereiding op studie beroep is over het geheel de variatie kleiner dan bij schoolbeleid. Verder is er een duidelijk onderscheid tussen OSB (standaard 1 - 4) en selectie en determinatie (standaard 5 - 8). Grosso modo wordt OSB matig tot zwak beoordeeld terwijl selectie en determinatie positief beoordeeld worden. Binnen OSB worden de voorwaarden wat beter gerealiseerd dan de uitvoering. Echter, niet vergeten mag worden dat voor alle standaarden de verschillen tussen scholen erg groot zijn.

De oordelen voor zorgbreedte zijn grotendeels positief. Alleen de derde standaard (differentiatie in inhoud, omvang, methodiek en tijd) is minder positief beoordeeld dan de andere standaarden. De verschillen tussen scholen op de zorgbreedte standaarden zijn duidelijk kleiner dan de verschillen tussen scholen op de standaarden voor schoolbeleid en voorbereiding op studie en beroep. Dat komt vooral omdat de kwalificatie 'vooral zwak' nauwelijks is gebruikt.

Wanneer we alle vier de onderzochte aspecten overzien, dan blijkt dat een school die voor één van de aspecten positief beoordeeld is, dat ook is voor andere aspecten. Dat wil niet zeggen dat er geen enkele school positief beoordeeld is op de ene standaard en negatief op een andere standaard, maar er is een duidelijke positieve samenhang tussen de standaarden (per aspect en over aspecten heen). Hieruit kan de conclusie getrokken worden dat verscheidene scholen het over de hele linie 'goed doen', terwijl andere scholen over de hele linie als 'zwak' betiteld kunnen worden.

2.4 Samenhang tussen aspecten

Omdat het aannemelijk is dat de verschillende deelsteekproeven uit één en dezelfde populatie afkomstig zijn, is getracht tot een verdere data-reductie te komen (Van den Bergh, Peters-Sips & Zwarts, 1998). Een eerste poging daartoe is ondernomen door per aspect één totaaloordeel te geven. Dit resulteert dus in vier totaaloordeelen. Deze totaaloordeelen zijn gerelateerd aan de residuele scores per standaard. En vervolgens is de zo ontstane covariantie-matrix in een

Tabel 6

De geschatte correlaties tussen de onderscheiden factoren

	P	S	V	Z	T-P	T-S	T-V	T-Z
Pedagogisch klimaat	1.00							
Schoolbeleid	.87	1.00						
Voorbereiding S&B	.41	.53	1.00					
Zorgbreedte	.53	.91	.69	1.00				
Totaal-Pedagogisch klimaat	.89	.37	.36	.33	1.00			
Totaal-Schoolbeleid	.72	.90	.41	.78	.31	1.00		
Totaal-Voorbereiding S&B	.46	.43	.98	.52	.38	.35	1.00	
Totaal-Zorgbreedte	.46	.63	.35	.89	.60	.80	.47	1.00

(confirmatieve) factor-analyse nader geanalyseerd. In Tabel 6 wordt een overzicht gegeven van de (geschatte) correlaties tussen factorscores gebaseerd op de standaarden en deze totaal-oordelen per aspect.

Beschouwing van de correlaties in Tabel 6 leert dat de factorscores op de standaarden van een aspect en het totaal-oordeel voor alle vier de projecten bijzonder hoog correleren. Bij voorbeeld de factorscore op de standaarden voor pedagogisch klimaat en het totaal-oordeel voor dit aspect correleert .89, en de correlatie tussen de factorscore voor oriëntatie studie en beroep met het totaaloordeel is maar liefst .98. Op het eerste gezicht lijken de totaal-oordelen derhalve een adequate samenvatting van de standaarden.

Bij modellering van de covariantiematrix (zie Tabel 6 voor de correlaties) bleek voor elk van de projecten één factor noodzakelijk. Daarbij waren de regressiecoëfficiënten van de totaal-oordelen op deze factoren lager dan die van de 'losse' standaarden (vergelijk: pedagogisch klimaat: .66 vs. 1.34; schoolbeleid: .90 vs. 1.00; voorbereiding studie en beroep: .92 vs. 1.07; zorgbreedte: .84 vs. 1.06). De totaal-oordelen representeren de factoren, of constructen, dus minder goed dan de oordelen op de individuele standaarden.

De correlaties tussen de vier factoren bleken allen positief, en te variëren van .32 ($r_{\text{ped., voorb.}}$) tot .85 ($r_{\text{schoolb., zorgb.}}$). Gezien de grootte van de standaardfout (.17) blijken schoolbeleid en zorgbreedte empirisch ononderscheidbaar (de geschatte correlatiecoëfficiënt tussen deze factoren wijkt niet significant af van 1.0).

Dit leidt tot de volgende conclusie. De oordelen op de standaarden per project correleren positief. Een school die op een standaard positief beoordeeld is, is dat (overwegend) ook op

een andere standaard, en een school die op een standaard minder goed beoordeeld is, is dat (overwegend) ook op de andere standaarden (zie ook Figuur 2). Echter, de samenvatting tot één totaaloordeel is niet geheel geslaagd; deze totaal-oordelen bevatten meer ruis dan gewenst. Hierdoor is een onderscheid tussen scholen op basis van het totaaloordeel minder valide dan de onderdelen waarop deze gebaseerd is.

3 De samenhang met enkele verklarende variabelen

Getracht is de verschillen tussen scholen te verklaren met behulp van tien variabelen. De informatie voor deze variabelen is door middel van vragenlijsten verzameld bij de evaluatie van de vier aspecten. Het betreft de samenstelling van de klassen, aantallen zorgleerlingen en het gebruik van ondersteuning. Tabel 7 geeft een overzicht van deze variabelen en van de schaal waarop zij gerepresenteerd zijn.

Het effect van deze variabelen is getoetst met het eerder genoemde meerniveaumodel. Het bleek dat geen van de in Tabel 7 genoemde variabelen een effect heeft op alle standaarden in één van de vier aspecten. Resteert de mogelijkheid van interacties tussen deze tien verklarende variabelen en de oordelen op de 25 standaarden. Alle ($25 \times 10 =$) 250 interacties zijn op hun merites getoetst. Slechts een klein aantal hiervan (10) bleek een significante bijdrage aan de verklaring van de verschillen tussen scholen te leveren. Dit komt overeen met wat op grond van kans verwacht mag worden (12). In Tabel 8 worden deze significante effecten gepresenteerd. Daaruit blijkt dat geen ervan erg groot is. Wat wel opvalt is een zekere clusterings van effecten. De variabele 'Vier grote ste-

Tabel 7

Een overzicht van tien verklarende variabelen

Variabele	Min	Max
Vier grote steden versus de rest	0	1
Percentage allochtone leerlingen ¹	0	100
Percentage categoriaal ingedeelde leerlingen in leerjaar 2 ¹	0	100
Aantal (dis)locaties	1	7
Aantal contacten ondersteuning	1	12
Aantal zwakke leerlingen ¹	0	550
Aantal gehandicapte leerlingen ¹	0	263
Aantal leerlingen met sociaal emotionele problemen ¹	0	198
Aantal hoog begaafde leerlingen ¹	0	42
Samenstelling van de school ²	0	1

¹ deze variabelen zijn in de analyse gecentreerd rond het gemiddelde² de samenstelling van de school is meegenomen door vier 'dummy' variabelen

den' komt drie keer voor, 'schoolsamenvestelling' vijf keer, en 'het percentage categoriaal ingedeelde leerlingen' twee keer. Daarnaast komt de standaard S3 (coördinatie van het onderwijskundig beleid) drie keer voor. Echter, gezien het geringe aantal aangetoonde effecten kan aan deze resultaten niet veel betekenis worden gehecht.

4 Conclusie

Het landelijk beeld van de realisatie van enkele aspecten van de basisvorming verschilt anno 1996 van aspect tot aspect.

- Bij pedagogisch klimaat worden de standaarden die betrekking hebben op het school- en klassenklimaat gemiddeld positief beoordeeld, waarbij er echter grote verschillen zijn tussen scholen. De standaarden die betrekking hebben op actief leren worden niet positief beoordeeld. De verschillen

tussen schoolgemiddelden zijn daarbij klein.

- Bij schoolbeleid is de variatie in oordelen erg groot, er zijn zowel scholen die als 'vooral sterk' beoordeeld worden als scholen die 'vooral zwak' zijn. Daarbij is bijna de helft van de scholen negatief beoordeeld. Dit geldt in het bijzonder voor de evaluatie van het beleid (standaard 4).
- Voorbereiding op studie en beroep valt in twee clusters uiteen. De standaarden die betrekking hebben op oriëntatie op studie en beroep (OSB) worden duidelijk negatiever beoordeeld dan de standaarden die selectie en determinatie betreffen. Bij de eerste groep standaarden zijn de verschillen tussen scholen ook groter dan bij het tweede cluster.
- De standaarden voor zorgbreedte worden over het geheel positief gewaardeerd. Relatief weinig scholen zijn onder de maat. De

Tabel 8

Parameterschattingen van enkele significante verklarende variabelen

verklarende variabele	standaard	effect	t-waarde
Vier grote steden versus de rest	P3	0.337	2.07
	S5	-0.329	-2.12
	V6	-0.417	-2.08
Schoolsamenvestiging met HAVO	S3	0.404	3.07
	V6	-0.353	-2.59
	Z4	0.418	2.62
	V7	0.236	1.99
	S3	-0.472	-3.62
Percentage categorale leerlingen in leerjaar 2	S3	-0.004	-2.47
	Z3	-0.006	-3.98

verschillen tussen scholen zijn ook kleiner dan bij de andere aspecten.

De relatie tussen de oordelen op de standaarden en een aantal verklarende variabelen is zwak. Geen enkele variabele heeft effect op alle standaarden in één van de vier aspecten. Per standaard zijn er wel enkele factoren die enig effect hebben. Dit betreft de vier grote steden, en de samenstelling en het determinatiebeleid van de school.

In het algemeen kan worden geconcludeerd dat de scholen in ruime mate voldoen aan de eisen die ook al vóór de basisvorming als kenmerken van goed onderwijs in de onderbouw van het voortgezet onderwijs werden beschouwd. In mindere mate worden de standaarden gerealiseerd die door de basisvorming een extra accent kregen of relatief nieuw zijn in de onderbouw, zoals bijvoorbeeld P6 ('de leerlingen nemen actief aan het leerproces deel'), P5 ('de leraren bevorderen in hun lessen een actief leerproces bij de leerlingen'), V3 ('de leraren besteden in de vakken voldoende aandacht aan de relatie tussen vakinhouden en de praktijk van de verschillende beroepen'), en Z3 ('het onderwijs voor leerlingen met extra onderwijsbehoeften kenmerkt zich door differentiatie in inhoud, omvang, methodiek en tijd').

De vier deelevaluaties laten interessante resultaten zien, die zeker een plaats verdienen in de evaluatieve rapportage die in 1999 wordt opgesteld. Ten aanzien van onderdelen van de deelevaluaties bestond behoefte aan verdieping, verificatie (door meer observaties in de klassen naast gesprekken met de schoolleiding, het middenkader en leraren) en/of ruimere analyses. Mede omdat verschillende standaarden uit de gehanteerde toetsingskaders – bijvoorbeeld de hiervoor genoemde P5, P6, V3 en Z3 – raken aan de kwaliteit van het onderwijsleerproces, heeft de inspectie deze standaarden – al dan niet met verbeterde indicatoren – ook opgenomen in het toetsingskader voor de evaluatie van het onderwijsleerproces. Dat onderzoek heeft in het schooljaar 1997-1998 plaatsgevonden. De oordelen over de betreffende standaarden kunnen nu worden gespecificeerd voor alle vakken in de basisvorming en voor de verschillende schoolsoorten. Aan de analyse van de resultaten wordt thans nog gewerkt.

Natuurlijk is het nog te vroeg om verstrek-

kende uitspraken over de kwaliteit van de basisvorming te doen. De scholen waren ten tijde van de evaluaties nog maar drie jaar bezig met de invoering van de basisvorming en zaten dus nog middenin het innovatieproces. Ook is het voorbarig aan het beeld dat uit deze evaluaties opdoemt, verwachtingen ten aanzien van nog niet geëvalueerde aspecten te koppelen. Een echte discussie over de stand van zaken van de basisvorming stelt de inspectie daarom uit tot het moment in 1999 waarop de geprogrammeerde evaluatiestudies zijn afgerond. Wat daarbij echter zeker in het oog zal worden gehouden is de duidelijke conclusie, die de nadere analyse van de vier deelevaluaties heeft opgeleverd, namelijk dat sommige scholen het over de gehele linie goed doen, terwijl andere scholen over de gehele linie zwak staan.

Literatuur

- Bergh, H. van den & Zwarts, M. (1998). *Pedagogisch klimaat: secundaire analyse*. Zwolle: Inspectie van het Onderwijs.
- Bergh, H. van den, M.W.H. Peters-Sips & Zwarts, M. (1998). *Overall analyse evaluatie schoolorganisatorische aspecten basisvorming*. Zwolle: Inspectie van het Onderwijs.
- Donk van den, J & Visser, J. (1997). *Pedagogisch klimaat: technisch rapport*. Intern rapport. Zwolle: Inspectie van het Onderwijs.
- Grift, W. van der & Plomp, H. (1997). *Schoolbeleid: technisch rapport*. Intern rapport. Zwolle: Inspectie van het Onderwijs.
- Heuvelmans, A.P.J.M. & Sanders, P.F. (1993). Beoordelaarsovereenstemming. In: Th.J.H.M. Eggen & P.F. Sanders, *Psychometrie in de praktijk*. Arnhem: Cito instituut voor toetsontwikkeling.
- HM Inspectors of Schools (1995) *Assessing Quality. A Handbook on Inspection for Secondary Education*. Edinburgh: Scottish Office, Education Department.
- Inspectie van het Onderwijs (1997a). *Schoolbeleid voor de basisvorming*. Zwolle: Inspectie van het Onderwijs
- Inspectie van het Onderwijs (1997b). *Een evaluatie van het pedagogisch klimaat in de leerjaren 1 en 2 van de scholen voor voortgezet onderwijs*. Zwolle: Inspectie van het Onderwijs.

- Inspectie van het Onderwijs (1997c). *Vorbereiding op studie en beroep in de basisvorming*. Zwolle: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (1997d). *Zorgbreedte in de basisvorming*. Zwolle: Inspectie van het Onderwijs.
- Inspectie van het onderwijs (1997e). *De basis gelegd*. Zwolle: Inspectie van het onderwijs.
- Janssens, F.J.G. (1997). Zo zijn onze manieren: een 20-jarige zoektocht naar de inspectiemethodologie voor onderwijsbeoordeling. In F.J.G. Janssens (ed), *Van onderzoek naar evaluatie. De methodologie van de onderwijsinspectie*. Den Haag: Delwel.
- Karstanje, P.N. (1996). *Beleidsstheorie basisvorming*. Een proeve van reconstructie. Amsterdam: SCO-Kohnstamm-instituut.
- Linden, W.J. van der, Meijer, R. R. & Vos, H.J. (1997). *Normeringsmethoden voor Inspectie-evaluaties*. Enschede: Vakgroep Onderwijskundige Meetmethoden en Data-analyse, Universiteit Twente.
- Wal, G. van der & Zwarts, M. (1997a). *Vorbereiding op studie en beroep: technisch rapport*. Intern rapport. Zwolle: Inspectie van het Onderwijs.
- Wal, G. van der & Zwarts, M. (1997b). *Zorgbreedte: technisch rapport*. Intern rapport. Zwolle: Inspectie van het Onderwijs.

Manuscript aanvaard 30-4-1999

Auteurs

Huub van den Bergh is verbonden aan het Utrechts Instituut voor Linguïstiek (OTS).

Adres: Universiteit Utrecht, Utrechts Instituut voor Linguïstiek (OTS), Trans 10, 3512JK Utrecht. E-mail: Huub.vandenBerg@Let.uu.nl

Marielle Peters-Sips is werkzaam als coördinerend inspecteur voor het voortgezet onderwijs.

Adres: Rijksinspectiekantoor Utrecht, Postbus 2730, 3500 GS Utrecht. E-mail: M.Peters-Sips@owinsp.nl

Michel Zwarts is als stafinspecteur verbonden aan de inspectie van het onderwijs.

Adres: Districtskantoor Inspectie van het onderwijs, Postbus 10048, 8000 GA Zwolle.
E-mail: M.Zwarts@owinsp.nl

Abstract

Evaluation of the 'basisvorming'

H. van den Bergh, M. Peters-Sips en M. Zwarts.
Pedagogische Studiën, 1999, 76, 224-236.

The evaluation of the implementation of the 'basisvorming' (national curriculum for the junior phase) will take place in 1999. The report to the Second Chamber will be based on research carried out by the Dutch inspectorate and by 'external' evaluators. In 1996 the inspectorate carried out a first series of aspect evaluations in 330 secondary schools focused on: pedagogical climate; educational policy; preparation for further education and work; pupil guidance and coaching.

Four major conclusions were drawn:

- The vast majority of the schools appear to have sound pedagogical climates both in the schools and the classrooms.
- Most schools provide satisfactory pupil guidance and coaching.
- The streaming and setting of pupils are generally well executed; however, only half of the schools prepare them well enough for further education and work.
- Only half of the schools adequately include the introduction and implementation of 'basisvorming' in their schoolpolicies.

Didactiek van authentiek leren in de Basisvorming

Stand van zaken bij docenten Nederlands en wiskunde

E.C. Roelofs en A.A.M. Houtveen

Samenvatting

Basisvorming betekent behalve een vernieuwing van het onderwijsaanbod ook dat van docenten veranderingen in didactisch handelen worden gevraagd. Van docenten wordt verwacht dat ze toepassingsgericht en vaardighedsgericht onderwijs verzorgen, waarbij bovendien sprake is van samenhang in het onderwijsaanbod en tussen verschillende onderwijs- en begeleidingsactiviteiten. In dit artikel wordt verslag gedaan van de mate waarin docenten erin slagen hun didactisch handelen vorm te geven volgens de kenmerken van authentiek leren, die als dekkend voor het verlangde handelen worden beschouwd. Bovendien wordt nagegaan in hoeverre realisatie van deze kenmerken samenhangt met voorwaarden op docent- en schoolniveau. Gebaseerd op een literatuuronderzoek en de uitkomsten van eerdere studies op dit terrein zijn een vragenlijst en een interviewleidraad ontwikkeld, aan de hand waarvan de realisatie van kenmerken van authentiek leren, en relevante beïnvloedende factoren kunnen worden vastgesteld. 88 Docenten Nederlands en 85 wiskundedocenten uit een representatieve steekproef van scholen voor voortgezet onderwijs namen deel aan het onderzoek in de periode van oktober tot en met december 1996. Hoewel de resultaten verschillen per vak, wijzen deze uit dat twee kenmerken van authentiek leren regelmatig worden gerealiseerd: het verwijzen naar de leefwereld van leerlingen en het hanteren van procesgerichte instructievormen. Vormen van samenwerking en communicatie vinden soms plaats. Echter, docenten creëren zelden leersituaties, waarin de grenzen van vakken, binnen- en buitenschoolse situaties, en de gehanteerde methode worden doorbroken, zoals dat in de karakteristiek 'samenhang' wenselijk zou zijn. De belangrijkste verklarende factoren voor de realisatie van een constructieve leeromgeving betreffen de mate waarin randvoorwaarden

zijn vervuld, zoals de kwaliteit en uitrusting van klaslokalen, en de mate waarin docenten zich deskundig achten voor lesgeven in de basisvorming. De mogelijkheden van varianten binnen de didactiek authentiek leren worden besproken.

1 Inleiding

Onder invloed van constructivisme en situationalisme vindt wereldwijd een verschuiving plaats in het denken over leren. Met name wordt de actieve rol van de leerling bij het verwerven en het gebruiken van kennis benadrukt, evenals het belang van interactie tussen docent en leerlingen en leerlingen onderling voor het verwerven van kennis. Leersituaties zouden vooral levensecht moeten zijn. Als gevolg van deze verschuiving worden andere eisen gesteld aan onderwijsinhoud en docentgedrag.

In hedendaagse onderwijsvernieuwingen vinden we deze opvattingen terug. Van docenten wordt verwacht dat zij leersituaties inrichten, waarin leerlingen een actieve inbreng kunnen hebben, onderling veel interacteren, en realistische taken uitvoeren, met als doel dat leerlingen hun kennis en vaardigheden toepassen in situaties buiten de school. Korthedshalve duiden we de bedoelde didactiek aan als didactiek van authentiek leren.

In dit artikel beantwoorden we de vraag in welke mate docenten in de Basisvorming de didactiek van authentiek leren realiseren. Tevens gaan we na welke condities op docent- en schoolniveau relevant zijn voor realisatie van een didactiek van authentiek leren.

Hieraan voorafgaand wordt een nadere beschrijving gegeven van de Basisvorming als vernieuwing, en wordt een beknopt overzicht gegeven van recente inzichten over leren. Vervolgens worden de didactische consequenties van deze recente inzichten in verband gebracht met de veranderingen in docentgedrag die de

basisvorming vraagt. Daarbij wordt het begrip 'authentiek leren' als overkoepelend didactisch concept gehanteerd.

2 De Basisvorming als vernieuwing

Met de Basisvorming wordt onder meer een algehele verhoging van het peil van het jeugd-onderwijs beoogd. Dit betreft vooral een verandering van de inhoud en de aanpak van het onderwijs, opdat leerlingen beter toegerust worden voor deelname aan een complexe en snel veranderende samenleving (Tweede Kamer, 1991).

Ter realisering van de Basisvorming zijn voor 15 vakken kerndoelen geformuleerd. In aanvulling op deze kerndoelen heeft het ProcesManagement Basisvorming (PMB) de zogenoemde TVS-karakteristieken geformuleerd: toepassingsgerichtheid, vaardigheden en samenhang (PMB, 1993a; 1993b). Deze karakteristieken geven de gewenste richting aan met betrekking tot het didactisch handelen.

Toepassingsgerichtheid wordt omschreven aan de hand van drie kenmerken: herkenbaarheid, toepasbaarheid en leren door doen. De manier waarop kennis aan de leerlingen wordt aangeboden dient levensecht te zijn, dat wil zeggen gekoppeld te worden aan voor de leerling herkenbare situaties uit de leefwereld. Toepasbaarheid drukt de eis uit, dat de leerling iets met de kennis moet kunnen doen, zowel als persoon als in het maatschappelijk functioneren. Leren door doen geeft aan dat leerlingen actief met de leerstof om moeten gaan, door erover te communiceren, leerervaringen te presenteren en met de leerstof handelingen uit te voeren.

Vaardigheden. Niet alleen kennis en inzicht, maar ook vaardigheden zijn van belang om aan te leren. Benadrukt wordt dat dit in samenhang gebeurt met 'toepassing'. Dit kunnen vakspecifieke vaardigheden zijn, maar ook vakoverstijgende vaardigheden. Meer concreet worden vijf algemene vaardigheden genoemd: een onderzoek verrichten, een standpunt verwoorden, samenwerken aan opdrachten, relaties leg-

gen tussen studie en beroep en beoordelingscriteria hanteren.

Samenhang. De derde karakteristiek betreft samenhang. Het PMB noemt drie aspecten van samenhang: samenhang door toepassing en vaardigheid, samenhang in leerinhoud, en samenhang in onderwijs en begeleiding. De in de kerndoelen beschreven kennis en vaardigheden zijn van belang voor alle vakken. Het wordt van belang geacht dat deze kennis en vaardigheden voor alle vakken en in samenhang met elkaar worden aangeleerd. Samenhang in leerinhoud komt tot uiting in samenhang tussen vakken en leerstofgebieden. Dit kan uitmonden in overleg tussen vaksecties op basis van overeenkomstige kerndoelen, in het gebruik van thema's (projectonderwijs), in vakkenintegratie, en in versterking van relaties tussen onderdelen van de leerinhoud. Samenhang in onderwijs en begeleiding heeft betrekking op de volgende elementen: versterking van de relatie tussen vaklessen en mentorlessen, het realiseren van een gezamenlijke keuzebegeleiding in de school en het verbinden van binnenschoolse en buitenschoolse activiteiten.

Het bovenstaande gedachtegoed vond zijn inspiratiebron in veranderde opvattingen over leren en onderwijzen. In het onderstaande beschrijven we deze opvattingen in hoofdlijnen.

3 Theorievorming op het gebied van leren

In de leerpsychologie is de laatste jaren een veranderde kijk op leren ontstaan die met behulp van drie principes samengevat kunnen worden. Hoewel verschillende stromingen blijven bestaan is over deze principes steeds meer consensus¹.

1 Leren als een constructief en cumulatief proces

Onder invloed van het constructivistische stromingen in de leerpsychologie² is de opvatting ontstaan dat de leerling zelf actief kennis construeert door nieuwe kenniselementen te relateren aan reeds bestaande cognitieve structuren (Bruer, 1993). Kennis is daarmee niet het gevolg van overdracht, maar van betekenisver-

lening door de leerling zelf. Kennis krijgt vorm binnen een voortdurend wisselende sociale context, door zelf doen, door onderzoek en actief ervaren (Vanderbilt Cognition and Technology Group, 1990). Samengevat gelden de volgende principes binnen het constructivisme (zie ook Boekaerts & Simons, 1993):

- leerlingen zijn actief organiserende wezens en geen passief opnemende wezens;
- begrijpen is meer dan informatie opnemen. Leerlingen leggen zelf structuur op aan de informatie die ze binnen krijgen;
- de gevormde kennisnetwerken veranderen in de loop van de ontwikkeling van een kind;
- begrijpen is nooit definitief. Netwerken worden steeds opnieuw georganiseerd door ervaring en het nadenken daarover;
- de hoeveelheid kennis en ervaring die een leerling al heeft, bepaalt wat geleerd kan worden;
- reflectie en reconstructie stimuleren het leren. Door na te denken over het eigen leren bouwen leerlingen een theorie op over hoe ze zelfstandig kunnen leren. Ze weten van daaruit wat ze wel en niet aankunnen bij een bepaalde leertaak. Vervolgens gaan ze na of dit klopt en stellen ze hun theorie over het eigen leren bij.

Deze principes kunnen ten onrechte suggereren dat het leerproces geheel dient te worden overgelaten aan de lerende zelf. Dit is niet het geval. Erkend wordt dat het zelfstandig construeren en opbouwen van een flexibel kennisbestand bij een groot aantal leerlingen langs geleidelijke weg aangeleerd moet worden (Boekaerts & Simons, 1993; De Jong, 1992; Veenman, 1992).

Samengevat komt de vernieuwing in het denken over leren hierop neer, dat leren als een constructief en cumulatief proces wordt beschouwd, waarbij de leerling in toenemende mate leerfuncties zelfstandig gaat beheren.

2 Leren ingebed in een specifieke socio-culturele context

In de laatste tien jaar is een steeds sterkere leertheoretische basis ontstaan voor het leren in betekenisvolle contexten.

Een aantal moderne leerpsychologen aangeduid als situationisten (Greeno, 1988, 1997; Lave, 1991; Brown, Collins & Duguid, 1989) benadrukt dat kennis het product is van activiteiten in een cultureel bepaalde context. Kennis is onlosmakelijk verbonden met de sociale en fysieke omgeving waarin de betreffende kennis wordt ontwikkeld en gebruikt en is niet een abstract gegeven op zich. De context bepaalt voor een belangrijk deel de structuur, inhoud en samenhang van gehanteerde concepten. Kennis is gekoppeld aan groeiende en veranderende inzichten in de gebruikscultuur. Leren vindt in deze opvatting altijd gesitueerd plaats ('situated learning').

Een consequentie voor leersituaties is dat kennis idealiter wordt aangeboden in betekenisvolle situaties, van waaruit leerlingen tot leren komen. Daarbij kan een onderscheid worden gemaakt tussen persoonlijk betekenisvol en maatschappelijk of cultureel betekenisvol. Bij het eerste gaat het om het aanspreken van leerlingen op hun eigen ervaringen en leefwereld. Bij maatschappelijk of cultureel betekenisvol gaat het om het inlijven van leerlingen binnen een cultuur van kennisgebruikers (Lave, 1991; Newmann, Marks & Gamoran, 1996).

3 Leren als proces van samenwerking en communicatie

In een veel geciteerd artikel van Resnick (1987) wordt meer aandacht bepleit voor taken waarin leerlingen met elkaar samenwerken, zoals bij veel buitenschools leren het geval is. In constructivistische benaderingen van leren wordt beklemtoond dat kennis ontstaat door onderhandeling ('negotiation') en ook voortdurend onderhevig is aan slijtage. Bovendien zouden leerlingen door confrontaties van inzichten via een cognitief conflict tot een nieuw en beter inzicht komen. Dit maakt dat kennisverwerving wordt gezien als een confrontatie van meningen, maar ook van samenwerking en co-constructie van kennis (Brown & Campione, 1994).

De bovenstaande ontwikkelingen in de kijk op leren hebben we in eerdere publicaties samengevat met behulp van het begrip 'authentiek leren' (Roelofs, Franssen & Grootsholten, 1996; Roelofs & Franssen, 1994). Authentiek

leren betreft in essentie het proces waarbij een mens in toenemende mate een actieve deelnemer aan onze cultuur wordt (zie ook Lave & Wenger, 1991). Op basis van de boven beschreven ontwikkelingen komen we tot de volgende meer gedetailleerde omschrijving van authentiek leren: een proces van leren waarbij de lerende voor hem- of haarzelf betekenisvolle inzichten verwerft, primair startend vanuit de intrinsieke motivatie en voortbouwend op bestaande inzichten. Authentiek leren vindt plaats in voor de lerende relevante, praktijkgerichte, en levensechte contexten, waarbij hij/zij een actieve constructieve en reflectieve rol vervult, mede via communicatie en interactie met anderen. De gevolgen van deze kijk op leren worden hieronder besproken.

4 Kenmerken van didactisch handelen dat authentiek leren bevordert

De gevolgen van de verandering in de kijk op leren voor onderwijssituaties worden in verschillende Nederlandse bronnen uitgewerkt, samengevat met begrippen als 'procesgericht onderwijs', 'krachtige leeromgeving', en 'activerende instructie' (Bolhuis & Kluvers, 1998; De Corte, 1990a, 1990b; Verschaffel, 1995; Lamberigts, Den Brok, Derksen & Bergen, 1999). De volgende aspecten van de onderwijs-situatie worden belangrijk geacht teneinde authentiek leren te bevorderen:

- een samenhangende onderwijsaanpak, waarin aandacht is voor leerinhoud en -strategie;
- vormen van procesgerichte instructie, waarbij sprake is van een toenemende zelfregulatie door leerlingen;
- hanteren van reële contexten, ofschoon ook aandacht wordt besteed aan decontextualiseren, teneinde transfer te bewerkstelligen;
- ruimte voor communicatie en interactie tussen (mede)leerlingen;
- een positief leerklimaat, waarin fouten maken mogelijk is, en waarin leerlingen leren dynamisch-affectieve processen te reguleren.

Hoewel de principes van een leeromgeving

voor authentiek leren wel worden beschreven, wordt niet duidelijk wat de gewenste veranderingen in het didactisch handelen zijn. Om hierin meer duidelijkheid te krijgen leren is een raamwerk ontwikkeld, waarin de belangrijkste inzichten over authentiek leren zijn uitgewerkt naar componenten van didactisch handelen. Geïnspireerd op een indeling van Newmann, Marks en Gamoran (1996) onderscheiden Roelofs et al. (1996) vier kenmerken van didactisch handelen, waardoor authentiek leren wordt bevorderd: 1. constructie van kennis in complete taaksituaties, 2. gerichtheid op de leefwereld, 3. relevantie van leren voor buitenschoolse situaties en 4. communicatie en samenwerking. Deze kenmerken zijn uitgewerkt voor drie componenten van didactisch handelen: 1) instructie en begeleiding, 2) opdrachten en 3) beoordelingsvormen. Deze didactiek van authentiek leren kan beschouwd worden als de theoretische inbedding van de TVS-karakteristieken. In het onderhavige onderzoek maken we gebruik van deze operationalisering³. De mate van realisatie van de TVS-karakteristieken meten we dan ook af aan de mate van realisatie van de didactiek van authentiek leren. Om deze reden werken we in het onderstaande deze kenmerken nader uit, waarbij de consequenties van elk van de kenmerken voor instructie/begeleiding, opdrachten en beoordelingsvormen worden aangegeven. De belangrijkste informatie is samengevat in de schema's 1 tot en met 4.

Constructie van kennis in complete taaksituaties

Het eerste kenmerk van een didactiek die authentiek leren bevordert, betreft constructie van kennis in complete taaksituaties (zie schema 1). Leren wordt zinvol wanneer leerlingen te maken krijgen met *complete of complexe taken* waarin meerdere kennis- en vaardigheidsaspecten zijn te onderscheiden, zoals bij (divergente) opdrachten met globale richtlijnen en eindcriteria, bij het zelfstandig verzamelen van informatie en verrichten van onderzoek en bij projectmatig werken. Uitgangspunt is dat een taak een betekenisvolle eenheid is die niet bij voorbaat wordt versnipperd in deeltaken. Om het constructieproces van kennisverwerving zichtbaar te maken wordt van de leerling verwacht dat hij verantwoording aflegt van de gevolgde procedure. De mate van complexiteit

kan variëren: in de beginfase van een leerproces kunnen leerlingen geconfronteerd worden met taken die niet complex zijn, maar wel een compleet geheel vormen en een beeld geven van de aan te leren kennis of vaardigheid.

Om het constructieproces van kennis te realiseren krijgt de leraar (geleidelijk aan) een *begeleidende en ondersteunende rol* in plaats van een puur overdragende rol. Instructie en begeleiding zijn procesgericht van aard (Vermunt, 1992). De docent geeft met name procesgerichte ondersteuning, hetgeen tot uiting komt in het verwoorden van denkprocessen, laten reflecteren op leerprocessen, en het geven van hints bij vragen in plaats van complete antwoorden. Om recht te doen aan het constructieproces van kennis en de complexiteit van leer-situaties biedt de docent ruimte voor inbreng van leerlingen en besteedt hij aandacht aan verbanden met andere vakken.

In leersituaties waarin leerlingen een kennisbestand opbouwen aan de hand van complete leertaken past een wijze van beoordeling waarin ten eerste complexe en divergente opgaven, en de gevolgde aanpak worden meegenomen en ten tweede het proces van persoonlijke voortgang zichtbaar is. Dat houdt in dat evaluatie een min of meer continu karakter krijgt. Een vergaande variant is het werken met portfolio's of leerlingdossiers, waarin de leerlingen een persoonlijk dossier aanleggen met verschillende producten, die tezamen de groei van de leerling illustreren.

Gerichtheid op de leefwereld

Aansluiten bij de leefwereld van leerlingen (zie schema 2) en hun voorkennis houdt in dat docenten bij de keuze van hun lesonderwerpen aansluiten bij interesses en leefwereld van de leerlingen. Leerlingen krijgen mogelijkheden om ervaringen uit te wisselen met elkaar en met de docent. Bij de uitleg van nieuwe leerinhouden binnen het schoolvak sluit de docent aan bij wat leerlingen uit hun leefwereld weten over het onderwerp. Daarbij is aandacht voor preconcepties van leerlingen, die misconcepties kunnen blijken te zijn. Er zijn gelegenheden voor leerlingen om persoonlijk stelling te nemen bij een bepaald vraagstuk. *Opdrachten en taken* zowel in klasverband als bij huiswerk kenmerken zich door een relatie met de leefwereld van leerlingen. Ook bij *beoordeling van leerresultaten* worden opdrachten gehanteerd die verband houden met de leefwereld van leerlingen en is idealiter sprake van inbreng van leerlingen bij het samenstellen van een persoonlijk dossier.

Belang van leren voor buitenschoolse situaties

Om leerervaringen en opbrengsten te realiseren die betekenis hebben voor de wereld buiten school en daarmee verder gaan dan het werken voor een cijfer (zie schema 3) kunnen *taken en opdrachten* worden gehanteerd waarbij de leerling in de rol van kennisgebruiker wordt gezet en een reëel professioneel probleem wordt opgelost. De resultaten van dergelijke opdrach-

Schema 1

Consequenties van constructie van kennis in complete taaksituaties voor didactisch handelen

<i>instructiel/begeleiding (procesgericht)</i>	<i>taken/opdrachten (integratief en compleet)</i>	<i>beoordeling (proces en product)</i>
- articuleren denkprocessen, reflecteren op leerprocessen - vragen beantwoorden met procesantwoord (hints voor oplossingsprocedure) - inzet leerlingenexpertise (docent mede-lerende) - uitleg verbanden met andere vakken	- divergente opdrachten - opdrachten met richtlijnen en eindcriteria - leerlingen zelfstandig informatie laten verzamelen - leerlingen onderzoek laten verrichten - nadruk leggen op ver(ant)woorden aanpak - projectmatig werken	- proces, gevolgde aanpak meenemen - mee betrekken divergente opdrachten - hanteren complexe opgaven (aan leefwereld en toepassings situatie gerelateerd) - continu karakter: hanteren persoonlijke leerlingdossier met verschillende producten

Schema 2*Consequenties van gerichtheid op leefwereld en persoonlijk voorkennis voor didactisch handelen*

<i>instructielbegeleiding</i>	<i>taken/opdrachten</i>	<i>beoordeling</i>
• keuze lesonderwerpen leefwereld • uitwisselen ervaringen leerlingen uit leefwereld • relatie leefwereld voor vak verduidelijkt • keuze onderwerpen aansluiten bij interesses leerlingen • ruimte voor persoonlijke stellingname • aandacht voor preconcepties bij leerlingen	• opdrachten in relatie tot leefwereld • huiswerk in relatie tot leefwereld	• opdrachten in relatie met de leefwereld • leerling heeft inbreng in samenstellen persoonlijk dossier

ten kunnen worden voorgelegd aan een 'echt' publiek buiten de school. Een minder vergaande variant is dat een publiek binnen de school, maar buiten de directe klassituatie wordt gehanteerd. Bij beoordeling van resultaten worden beroepsbeoefenaren en andere personen buiten de school betrokken.

Bij *instructie en begeleiding* legt de docent contacten met de actualiteit en verschillende beroepssituaties. De docent hanteert hiertoe verschillende klassieke media zoals kranten, tijdschriften, televisie en radio, maar ook moderne media, waaronder computers, met toepassingen als CD-ROM, internet en E-mail. Leerlingen kunnen bij het leren gebruik maken

van deze media en andere professionele hulpmiddelen.

Communicatie en samenwerking

Didactisch handelen vanuit het perspectief van communicatie en samenwerking (zie schema 4) houdt in dat bij opdrachten sprake is van onderlinge afhankelijkheid: bij de realisering van een (groeps)opdracht zijn leerlingen afhankelijk van de inzet en expertise van medeleerlingen. Bij de uitvoering van leertaken is sprake van onderlinge communicatie (overleg, discussie). Er doen zich situaties voor waarbij tijdens de uitvoering van taken veel interactie is tussen leerlingen of situaties waarbij leerlingen de resultaten van een taak elk vanuit een

Schema 3*Consequenties van aandacht aan relevantie voor buitenschoolse/professionele situaties voor didactisch handelen*

<i>instructielbegeleiding</i>	<i>taken/opdrachten</i>	<i>beoordeling</i>
• ingaan op actuele ontwikkelingen • gebruik (moderne) media en professionele hulpmiddelen • relaties leggen met beroepssituaties	• opdrachten die verder gaan dan het werken voor een cijfer • opdrachten waarin een reëel professioneel probleem wordt opgelost • opdrachten voor een publiek buiten school • opdrachten voor een publiek buiten de directe klascontext	• beoordeling ook door anderen dan de docent zelf buiten de directe klassituatie • beoordeling door beroepsbeoefenaren buiten de school

<i>instructie/begeleiding</i>	<i>taken/opdrachten</i>	<i>beoordeling</i>
• docent meer begeleider en facilitator • controle voortgang naar leerlingen delegeren • leerlingen laten samenwerken in groepen	• onderlinge afhankelijkheid • resultaten presenteren • leerlingen krijgen rol toegewezen	• overleg tussen leerlingen en docent als basis • zowel individuele beoordeling als beoordeling van het groepsresultaat

eigen rol presenteren aan medeleerlingen of een ander publiek.

Wat geldt voor authentieke leersituaties in het algemeen geldt voor samenwerking en communicatie in het bijzonder: de docent is begeleider van het leerproces. Hij biedt de faciliteiten om leerlingen in groepen samenwerkend te laten leren, bewaakt het groepsproces, en spreekt leerlingen aan op de door hen gevolgde aanpak. In een ideale situatie wordt de controle over de voortgang van het leerproces gedelegeerd naar de leerlingen. Beoordeling van leerresultaten is in het meest vergaande geval een kwestie van overleg tussen docent en leerlingen. Daarnaast wordt bij beoordeling rekening gehouden met zowel de individuele bijdrage als de groepsbijdrage.

In de paragraaf 5 wordt beschreven hoe het begrip authentiek leren geoperationaliseerd is in een onderzoeksinstrument.

5 Opzet en uitvoering van het onderzoek

In het onderzoek worden de volgende vragen beantwoord:

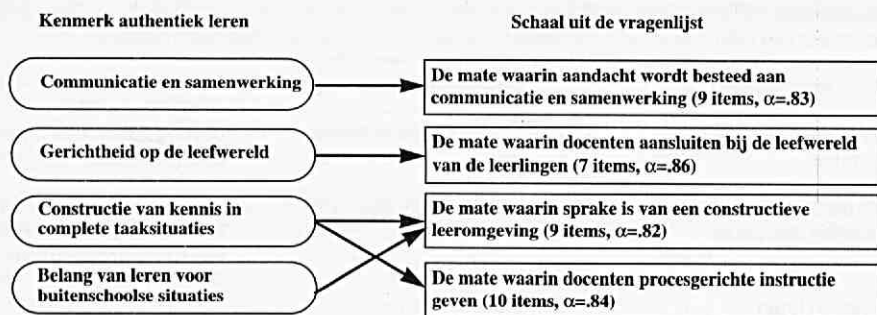
1. In welke mate realiseren docenten een didactiek van authentiek leren binnen de basisvorming?
2. In welke mate worden condities voor de invoering van basisvorming op docent- en schoolniveau gerealiseerd?
3. Wat is de samenhang tussen de realisatie van een op authentiek leren gerichte didactiek enerzijds en de invulling van condities op school- en docentniveau anderzijds?

Deze onderzoeksvragen maken deel uit van een grootschalig onderzoek gericht op de opvattingen van docenten over de Basisvorming en de wijze waarop zij de Basisvorming in de praktijk realiseren (Roelofs, Vermeulen & Houtveen, 1998). Om deze vragen te beantwoorden zijn docenten Nederlands en Wiskunde die lesgeven in de Basisvorming schriftelijk en mondeling bevraagd. Om na te gaan welke factoren op school- en docentniveau de gevonden resultaten zouden kunnen verklaren zijn met de docenten interviews gehouden. De interviews zijn afgenomen door derdejaars onderwijskundestudenten, die hiertoe een training ontvingen.

Instrumenten

Om zicht te krijgen op de mate waarin didactisch handelen wordt gerealiseerd dat ondersteunend is voor de Basisvorming is gebruik gemaakt van een bewerkte versie van de vragenlijst 'Toepassingsgericht onderwijs' (Roelofs, Franssen & Grootsholten, 1996). Enkele items zijn geherformuleerd, omdat ze achteraf bezien voor meer dan één uitleg vatbaar waren. Daarnaast bleek dat niet voor alle kenmerken van authentiek leren items beschikbaar waren met betrekking tot instructie en begeleiding, de aard van leertaken en opdrachten, en de aard van de beoordeling.

Met behulp van exploratieve factoranalyses (principale componentenanalyses) en betrouwbaarheidsanalyses (Cronbach's alpha) is vervolgens nagegaan in hoeverre langs empirische weg theoretisch verdedigbare schalen konden worden gevormd. De resulterende set subschalen week in twee opzichten af van de oorspronkelijke indeling. De oorspronkelijke subschaal 'constructie van kennis in complete taaksitu-



Figuur 1. Operationalisering van kenmerken van authentiek leren

aties' werd gesplitst in een subschaal die we hebben benoemd als 'procesgerichte instructie' en in een subschaal 'constructieve methode-overstijgende leeromgeving'. Deze laatste schaal bevatte tevens de items uit de oorspronkelijke schaal 'relevantie van leren voor buitenschoolse situaties' en enkele items uit de schaal 'aansluiting bij de leefwereld'. De schalen 'samenwerking en communicatie' en 'aansluiting bij de leefwereld' bleven na deze ingrepen overeind.

Aldus resulteerde een instrument dat bestaat uit de volgende vier intern consistente schalen: 1) de mate waarin docenten aansluiten bij de leefwereld van de leerlingen; 2) de mate waarin docenten procesgerichte instructie geven; 3) de mate waarin de docent samenwerking en communicatie bevordert; 4) de mate waarin de docent een constructieve, methode-overstijgende leeromgeving creëert. Iedere subschaal bevat uitspraken over onderwijsactiviteiten die aan docenten worden voorgelegd en welke beantwoord worden naar frequentie van voorkomen (van 1=nooit tot 6=altijd). Er is dus sprake van percepties van het eigen onderwijsgedrag. De items behorend bij de subschalen zijn te vinden in de tabellen 1 tot en met 4. De subschaalscore geeft de gemiddelde mate van voorkomen van verschillende onderwijsactiviteiten weer, ressorterend onder die subschaal. In figuur 1 wordt de operationalisering van de kenmerken van authentiek leren met behulp van subschalen gevisualiseerd, inclusief gegevens over aantallen items en de Cronbach's alpha-coëfficiënten per schaal.

Inventariseren van gedrag via zelfpercepties brengt onder meer als probleem met zich mee dat vragen – al dan niet bewust – door betrok-

kenen sociaal wenselijk ingevuld kunnen worden. Om een beeld te krijgen van de mate waarin de in ons onderzoek betrokken docenten geneigd zijn vragen sociaal wenselijk te beantwoorden, is een daartoe eerder ontwikkeld instrument afgenomen (Houtveen, Vermeulen & Van de Grift, 1993). De schaal bestaat uit acht uitspraken, die docenten kunnen invullen naar frequentie van voorkomen. Voorbeelden van items zijn: 'Ik ben voor iedere les gemotiveerd' en 'Ik verlies nooit mijn geduld bij de leerlingen'.

In de interviews kwam aan de orde of de docenten condities op schoolniveau als bevorderend dan wel als belemmerend ervaren voor de realisatie van de Basisvorming. Condities op schoolniveau hebben betrekking op de invulling van randvoorwaarden: de kwaliteit van lokalen, het gehanteerde lesrooster, de groepsgrootte, het aantal uren voor het vak, de op school gehanteerde methode, de kwaliteit van traditionele leermiddelen en computers, de aanwezigheid van een schooldocumentatiecentrum. Daarnaast is gevraagd naar de invulling van coördinatie en ondersteuning op schoolniveau: de verkregen schoolinterne ondersteuning, implementatie-bevorderende ondersteuning door de schoolleiding, afspraken op sectie- en schoolniveau over lesgeven in de huidige situatie, overleg en of samenwerking met sectiegenoten in huidige situatie. Tot slot is gevraagd naar condities op schoolniveau die betrekking hebben op de verkregen gelegenheid tot nascholing en de genoten vormen daarvan.

De antwoorden van de docenten zijn zoveel mogelijk gescoord op vooraf onderscheiden

antwoordcategorieën. Dit maakte het mogelijk groepjes items inzake condities op schoolniveau in te dikken tot de volgende samengestelde variabelen: 'het aantal randvoorwaarden dat als bevorderend is ervaren' (bereik 0-7) 'het aantal aspecten van interne ondersteuning dat als belemmerend is ervaren' (bereik 0-3) 'het aantal gevolgde externe ondersteuningsvormen' (bereik 0-4).

Conditie op docentniveau hebben betrekking op de mate waarin docenten zich toegerust voelen om de Basisvorming te realiseren en welke gevolgen de invoering van de basisvorming heeft gehad voor hun beroepsopvatting. Om zicht te krijgen op de mate waarin docenten zich toegerust voelen om les te geven in overeenstemming met de didactiek van authentiek leren zijn tijdens het interview docenten gescoord op 9 uitspraken, waarin een aspect van de didactiek van authentiek leren aan de orde komt. Voorbeelden hiervan zijn: 'De docent voelt zich toegerust om leerlingen in de klas een actieve rol te geven', 'De docent voelt zich toegerust om gebruik te maken van andere lesvormen dan frontaal lesgeven'. De uitspraken vormen een homogene schaal ($\alpha=.76$).

Onderzoeksgroep

Het onderzoek is uitgevoerd in het schooljaar 1996/1997 op 90 scholen voor voortgezet onderwijs. De representativiteit van de onderzoeksgroep is nagegaan aan de hand van de volgende achtergrondkenmerken van de scholen: schooltypesamenstelling, schoolgrootte, denominatie en geografische ligging. Alleen ten aanzien van de schoolgrootte bleek de onderzoeksgroep enigszins af te wijken van de populatie. Bij het analyseren van de onderzoeksresultaten is om deze reden nagegaan of er sprake is van invloed van de schoolgrootte. Dit bleek niet het geval te zijn (zie Roelofs, Vermeulen & Houtveen, 1998).

We kunnen derhalve de conclusie trekken dat de onderzoeksbevindingen gegeneraliseerd kunnen worden naar de populatie van scholen voor Voortgezet Onderwijs.

De keuze van docenten verliep via de sectie-leider en leidde tot een bijna volledige respons. Per school is steeds een docent Nederlands ($n=88$) en een docent wiskunde ($n=85$) geïnterviewd ($n=173$).⁴ Daarnaast retourneerde 85% van deze docenten de vragenlijst met

gegevens over authentiek leren en sociale wenselijkheid. 31.8% van de docenten vulde de vragenlijst in voor het VBO, 40.5% voor de MAVO, 16.8% voor de HAVO en 11.0% voor het VWO.

Data-analyse

Ter beantwoording van de eerste en de tweede onderzoeksvraag (resp. realisatie kenmerken van authentiek leren en realisatie condities) zijn beschrijvende analyses uitgevoerd. Verschillen in de mate van realisatie van authentiek leren tussen vakdocenten Nederlands en wiskunde zijn getoetst met behulp van t-toetsen voor onafhankelijke steekproeven. Deze relatief eenvoudige analysemethode bleek toereikend, nadat uit variantie-analyses was gebleken dat docenten uit verschillende onderwijsstromen (VBO, MAVO, HAVO/VWO) niet van elkaar verschilden.

Om de in onderzoeksvraag 3 bedoelde samenhang te onderzoeken zijn stapsgewijze multiële regressieanalyses uitgevoerd, waarbij condities op school- en docentniveau als voorspellers optraden in regressievergelijkingen met als afhankelijke variabelen de afzonderlijke schalen voor authentiek leren. Daarbij is 'sociale wenselijkheid' steeds als eerste voorspeller ingevoerd in de vergelijking. Aldus kon het relatieve gewicht van condities bij de voorspelling van verschillen in realisatie worden nagegaan.

6 Resultaten

6.1 Realisatie van de didactiek van authentiek leren

In de tabellen 1 tot en met 4 zijn de resultaten op de schalen voor authentiek leren weergegeven. Over het geheel genomen geven docenten aan dat ze vrij vaak onderwijsactiviteiten ondernemen om *aan te sluiten bij de leefwereld van de leerlingen* (zie Tabel 1, $\bar{x}=3.88$; range=1-6).

Docenten Nederlands scoren gemiddeld significant hoger dan wiskundedocenten (resp. $\bar{x}=4.12$ en 3.64). Het verschil is niet alleen significant voor het schaalgemiddelde, maar eveneens voor zes van de zeven items uit de schaal afzonderlijk.

Docenten Nederlands gaan bij een nieuw

Tabel 1

Mate waarin tijdens de lessen wordt aangesloten bij de beleavingswereld van leerlingen; toetsing verschil met behulp van t-toets voor onafhankelijke steekproeven

	wiskunde (n=73)		Nederlands (n=74)		Totaal (n=147)		t
	$\bar{x}$	Sd	$\bar{x}$	Sd	$\bar{x}$	Sd	
Aansluiten bij de leefwereld van de leerlingen (7 items, $\alpha = .86$)	3.64	.72	4.12	.70	3.88	.74	4.1**
6. Ik geef leerlingen opdracht problemen op te lossen die voor hen betekenisvol zijn	3.32	.92	3.44	.83	3.38	.88	.8
2. Ik stem de inhoud van mijn lessen zoveel mogelijk af op de beleavingswereld van de leerlingen	3.62	.93	3.99	.96	3.81	.96	2.4*
7. Ik probeer bij elk onderwerp dat ik behandel aan te geven waarom het voor leerlingen belangrijk is dat ze dit leren	3.69	.91	4.03	.91	3.86	.92	2.3*
1. Ik geef leerlingen gelegenheid om zelf ervaringen of informatie naar voren te brengen	3.50	.94	4.22	.90	3.87	.99	4.8**
3. Bij keuze van onderwerpen probeer ik zoveel mogelijk aan te sluiten bij de interesses van leerlingen	3.57	1.31	4.26	1.00	3.92	1.21	3.7**
5. Als ik begin met een nieuw onderwerp ga ik eerst na wat de leerlingen er al van weten	3.69	1.18	4.36	1.04	4.03	1.16	3.7**
4. Ik kies voorbeelden die leerlingen aanspreken	4.04	.93	4.58	.84	4.32	.92	3.7**

N.B. 1 nooit; 2= bijna nooit, 3=af en toe, 4= vaak, 5= bijna altijd, 6= altijd. * = $p < .05$, ** = $p < .01$.

onderwerp vaak na wat leerlingen daar al van weten. Bij hun keuze van onderwerpen sluiten deze docenten vaak aan bij de interesse van leerlingen, ze stemmen de inhoud van hun lessen af op de beleavingswereld van leerlingen, ze geven leerlingen de kans zelf ervaringen of informatie naar voren te brengen. Ook geven ze vooraf het belang van een onderwerp aan. Tijdens de uitleg kiezen ze zeer vaak voorbeelden die in hun ogen de leerlingen aanspreken. Bij al deze activiteiten geldt dat wiskundedocenten ze significant minder frequent maar toch nog regelmatig uitvoeren.

Zowel wiskundedocenten als docenten Nederlands zijn iets terughoudender als het gaat om het laten oplossen van problemen die voor leerlingen zelf betekenisvol zijn. Kennelijk gebeurt aansluiting bij de leefwereld vooral binnen de kaders van het schoolvak.

Als we kijken naar het tweede aspect van de didactiek van authentiek leren, *procesgerichte instructie* (zie Tabel 2) dan zien we dat docenten hier gemiddeld een nog positiever beeld schetsen van hun eigen lesgedrag. Het verschil tussen wiskundedocenten en docenten Nederlands is opnieuw significant, met dien verstande dat het nu de wiskundedocenten zijn die gemiddeld hoger scoren (resp. $\bar{x}=4.57$ en $\bar{x}=4.05$).

Wiskundedocenten schenken bij hun uitleg aandacht aan de wijze waarop leerlingen een taak moeten aanpakken en maken de leerlingen duidelijk dat een goede aanpak net zo belangrijk is als het juiste antwoord. Tijdens interactie met leerlingen laten deze docenten hun leerlingen zeer vaak verwoorden hoe ze tot een oplossing zijn gekomen. Bij vragen geven wiskundedocenten meestal een aanwijzing in plaats van het antwoord, en bij foute antwoorden krijgen leerlingen de kans om alsnog een goed antwoord te geven. Ze stimuleren daarnaast leerlingen vaak zelfstandig achter de oplossing van een probleem of opdracht te komen. Hun proefwerken bestaan zeer vaak uit een klein aantal complexe opdrachten die het zetten van meerdere denkstappen vereisen. Bij al deze activiteiten geldt dat ook docenten Nederlands ze vaak uitvoeren, maar significant minder vaak dan wiskundedocenten (zie Tabel 2). Gezien de aard van de vakdidactiek wiskunde lijkt deze relatief grote aandacht voor oplossingsprocessen in de lijn der verwachtingen te liggen.

Het meest duidelijk komt het verschil in procesgerichtheid tussen de onderscheiden categorieën docenten tot uitdrukking in de waardering van prestaties. Waar wiskundedocenten aangeven bijna altijd een juiste manier van aanpak en denken zwaarder mee te tellen voor het cijfer dan de goede uitkomst op zich

($\bar{x}=4.92$), doen docenten Nederlands dit slechts af en toe ($\bar{x}=2.96$). Ook vragen docenten Nederlands hun leerlingen minder vaak dan hun collega's naar de denkstappen die zij hebben genomen om tot een goede oplossing te komen (resp 3.92 en 4.82). Mogelijk leent het vak wiskunde zich hier beter voor.

De mate waarin docenten *aandacht besteden aan samenwerking en communicatie* is het derde aspect van de didactiek van authentiek leren. De gemiddelde score op deze schaal is beduidend lager dan de gemiddelde scores op de beide voorafgaande aspecten. Bovendien zijn de verschillen tussen de docenten in de mate waarin ze hier, naar eigen zeggen, aandacht aan besteden groter dan op beide andere schalen. Dit blijkt uit de grootte van de standaarddeviatie.

Opnieuw is het verschil tussen docenten wiskunde en Nederlands significant. Gemiddeld geldt dat wiskundedocenten af en toe ($\bar{x}=2.97$) en docenten Nederlands vrij regelmatig ($\bar{x}=3.44$) hun leerlingen samenwerkend laten leren in hun klassen. Het verschil is significant op 5 van de 9 uitspraken uit de schaal. Zo laten docenten Nederlands hun leerlingen vrij

regelmatig werken in groepen, terwijl dit bij wiskunde heel weinig voorkomt. Bij de voorbereiding van samenwerkend leren geven docenten Nederlands aan vrij vaak zorg te besteden aan goede groepsopdrachten. Vrij regelmatig verwachten ze van de leerlingen dat ze bij groepsopdrachten tot een gezamenlijk resultaat komen. Docenten wiskunde besteden veel minder aandacht aan groepsopdrachten en stellen minder vaak de eis dat (eerst) de medeleerling wordt geraadpleegd. Docenten Nederlands eisen af en toe van de leerlingen dat zij alleen als groep vragen mogen stellen, waar docenten wiskunde dit zelden doen.

Op de overige vier uitspraken bestaan nauwelijks verschillen tussen docenten Nederlands en wiskunde. Het verloop van de samenwerking tussen leerlingen is zowel bij Nederlands als wiskunde gemiddeld genomen zelden onderwerp van gesprek. De beoordeling van resultaten van groepswork geschiedt bij beide vakken af en toe door middel van overleg tussen docent en leerlingen. Wanneer gewerkt wordt in groepen, lopen docenten meestal rond in de klas om overzicht te houden. In die zin hanteren zowel docenten wiskunde als docenten Nederlands een begeleidende of supervise-

Tabel 2
Mate waarin aandacht wordt besteed aan procesgerichte instructie

	wiskunde (n=71)		Nederlands (n=71)		Totaal (n=142)		
	$\bar{x}$	Sd	$\bar{x}$	Sd	$\bar{x}$	Sd	t
procesgerichte instructie (10 items, $\alpha=.84$)	4.57	.58	4.05	.59	4.31	.64	-5.2**
10. Als ik een vraag stel tijdens mijn les, geef ik even bedenktijd voordat ik een leerling de beurt geef	3.66	1.06	3.34	1.02	3.50	1.05	1.9
8. Ik geef opdrachten waarbij leerlingen zelf kunnen bepalen hoe ze tot de oplossing komen	3.77	.91	3.40	.73	3.58	.84	-2.7**
3. Bij het geven van cijfers tel ik de juiste aanpak en juiste manier van denken zwaarder mee dan goede antwoord	4.92	1.06	2.96	1.16	3.93	1.48	-11.1**
6. Ik probeer zo les te geven dat leerlingen gaan nadenken over de wijze waarop ze leertaken aanpakken	4.34	.93	4.15	.71	4.24	.83	-1.4
2. Ik vraag leerlingen hoe ze tot oplossing zijn gekomen, welke denkstappen ze hebben genomen	4.82	.93	3.92	.93	4.37	1.03	-5.9**
9. Mijn proefwerken bestaan uit een klein aantal complexe opdrachten die meerdere denkstappen vereisen	4.80	1.23	4.63	1.04	4.51	1.14	1.3
5. Als een leerling een vraag stelt zeg ik het antwoord niet voor, maar geef ik een aanwijzing	4.70	.79	4.45	.82	4.58	.81	-1.9
4. Ik stimuleer leerlingen om zelf achter de oplossing van een probleem/opdracht te komen	4.85	.81	4.56	.73	4.70	.78	-2.3*
7. Als een leerling tijdens de les een fout antwoord geeft, geef ik hem/haar gelegenheid om alsnog goed antwoord geven	4.70	1.03	4.70	.88	4.70	.95	0.0
1. Ik maak de leerlingen duidelijk dat aanpak opdracht minstens zo belangrijk is als juiste antwoord	5.26	.79	4.22	.93	4.73	1.01	-7.3**

N.B. 1 nooit; 2= bijna nooit, 3=af en toe, 4= vaak, 5= bijna altijd, 6= altijd. *= $p<.05$, **= $p<.01$.

Tabel 3

Mate waarin aandacht wordt besteed aan samenwerking en communicatie

	wiskunde (n=72)		Nederlands (n=73)		Totaal (n=145)		t
	$\bar{x}$	Sd	$\bar{x}$	Sd	$\bar{x}$	Sd	
Samenwerking en communicatie (9 items, $\alpha = .83$)	2.97	1.01	3.44	.81	3.21	.94	3.1**
6. Als leerlingen in groepen werken, bespreek ik aan het eind van de les hoe de samenwerking in de groepen is verlopen	2.59	1.74	2.41	1.47	2.50	1.61	-.70
9. Als leerlingen in groepen werken, mogen zij mij alleen als groep vragen stellen, niet als individu	1.95	1.27	3.08	1.27	2.52	1.39	5.4**
4. Als een leerling tijdens het zelfstandig werken iets niet begrijpt, vraagt hij/zij in principe eerst een medeleerling om uitleg	2.11	1.07	3.12	1.15	2.62	1.22	5.6**
7. Ik laat de leerlingen samenwerken in groepen	2.55	1.32	3.44	1.26	3.01	1.36	4.2**
3. Ik besteed veel zorg aan goede groepsopdrachten	2.77	1.38	3.74	1.38	3.26	1.46	4.3**
12. Bij groepsopdrachten geef ik de leerlingen opdracht om tot een gezamenlijk resultaat te komen	3.01	1.65	3.58	1.35	3.30	1.53	2.3*
8. De waardering van resultaten van een groepsproces is een zaak van gezamenlijk overleg tussen leraar en leerlingen	3.30	1.38	3.39	.99	3.34	1.19	.50
5. Ik laat de leerlingen de resultaten van opdrachten aan elkaar presenteren	3.85	1.36	3.74	1.16	3.79	1.26	-.50
1. Als leerlingen in groepen werken loop ik door de klas om overzicht te houden	4.47	1.72	4.60	1.21	4.53	1.48	.50

N.B. 1 nooit; 2= bijna nooit, 3=af en toe, 4= vaak, 5= bijna altijd, 6= altijd. * = $p < .05$, ** = $p < .01$.

rende rol. Zowel docenten wiskunde als docenten Nederlands laten leerlingen vrij regelmatig de resultaten van opdrachten aan elkaar presenteren.

Het laatste aspect van de didactiek van authentiek leren dat we hebben onderscheiden betreft de mate waarin er sprake is van een *constructieve leeromgeving* (zie Tabel 4). Hiermee wordt bedoeld een leersituatie waarin leerlingen zelfstandig werken aan constructieve complete opdrachten, waarbij de grenzen van het vak, de methode en ook de klassituatie worden doorbroken, en waarbij een relatie wordt gelegd met maatschappelijk functioneren. Van een dergelijke leeromgeving is, zoals uit Tabel 4 blijkt, nauwelijks sprake. Dit aspect scoort van de onderscheiden aspecten het laagst. De verschillen tussen de docenten onderling zijn bovendien op dit aspect het kleinst ($SD = .55$). Opvallend is dat ook op dit aspect het verschil tussen de gemiddelde scores van de wiskunde-docenten en de docenten Nederlands significant is. Bij Nederlands is nog heel af en toe sprake van een constructieve leeromgeving ($\bar{x} = 2.67$), bij wiskunde bijna nooit ($\bar{x} = 2.26$).

De volgende activiteiten die verder gaan

dan de oefeningen in de methode zien we bij wiskunde bijna nooit gebeuren en bij Nederlands heel af en toe: contacten van leerlingen met kennisgebruikers buiten de school organiseren, opdrachten buiten de school geven die maatschappelijk relevant zijn, onderzoek laten verrichten of zelfstandig informatie laten verzamelen.

Activiteiten die wat vaker plaatsvinden (af en toe) en waarin geen significant verschil bestaat tussen Nederlands en wiskunde zijn het laten meenemen van materialen ter verduidelijking van leerstof en het geven van opdrachten waarbij leerlingen ook kennis vanuit andere vakken moeten gebruiken. Deze activiteiten kunnen in tegenstelling tot bovenstaande activiteiten binnen de reguliere schoolse context plaatsvinden.

Ten slotte, projectactiviteiten waarbij leerstof uit verschillende vakken wordt geïntegreerd vinden bij Nederlands ($\bar{x} = 3.21$) wat meer en bij wiskunde ($\bar{x} = 2.73$) wat minder plaats dan af en toe. Het verschil is significant.

We moeten constateren dat de in het onderzoek betrokken docenten in hoge mate geneigd zijn tot sociaal wenselijke beantwoording, getuige

Tabel 4

Mate waarin sprake is van een constructieve leeromgeving

	wiskunde (n=73)		Nederlands (n=72)		Totaal (n=145)		t
	$\bar{x}$	Sd	$\bar{x}$	Sd	$\bar{x}$	Sd	
Constructieve leeromgeving (9 items, $\alpha = .82$)	2.26	.55	2.67	.48	2.46	.55	4.9**
3. Leerlingen krijgen opdracht om zelf concrete voorbeelden bij bepaalde leerstofonderdelen te bedenken	1.45	.58	2.00	.88	1.72	.79	4.5**
7. In mijn lessen gebruik ik andere leermiddelen buiten het leerboek	1.76	.79	2.08	.87	1.92	.84	2.4*
5. Leerlingen hebben in lessen contact met kennisgebruikers buiten de formele klas-situatie	1.76	.92	2.61	.78	2.19	.95	6.1**
1. Ik geef opdrachten buiten de school die maatschappelijk relevant zijn	2.11	.77	2.62	.63	2.37	.74	4.5**
4. Ik laat leerlingen zelfstandig informatie verzamelen over onderwerpen die in mijn lessen aan de orde komen	2.33	.83	2.79	.91	2.57	.90	3.2**
2. Ik geef opdrachten waarbij leerlingen zelf een klein onderzoekje moeten uitvoeren	2.42	.97	2.91	.61	2.67	.84	3.7**
8. Ik geef opdrachten waarbij de leerlingen ook kennis vanuit andere vakken moeten gebruiken	2.91	.76	2.82	.79	2.86	.77	-.70
9. Ik neem voorwerpen mee ter verduidelijking van bepaalde leerstof of laat deze meenemen door de leerlingen	2.91	.86	3.00	1.00	2.95	.93	.60
6. Leerlingen werken aan projecten waarin leerstof uit verschillende vakken geïntegreerd wordt	2.73	.85	3.21	.79	2.97	.85	3.6**

N.B. 1 nooit; 2= bijna nooit, 3=af en toe, 4= vaak, 5= bijna altijd, 6= altijd. * = $p < .05$, ** = $p < .01$.

een gemiddelde score van 5.04 (range 1-6) op deze schaal en een geringe spreiding van .57. Het ligt dan ook in de rede te verwachten dat de schalen van het instrument voor het meten van de didactiek van authentiek leren eveneens in zekere mate sociaal wenselijk zijn beantwoord. Om hier enig zicht op te krijgen zijn correlaties berekend. Voor drie van de vier schalen geldt een positief en significant verband met sociale wenselijkheid: aansluiten van het lesprogramma bij de beleavingswereld ($r=.34$), samenwerking en communicatie ($r=.23$) en met name procesgerichte instructie ($r=.47$). De inschattingen van de docenten zullen voor wat betreft deze drie aspecten van de didactiek van authentiek leren overschattingen vormen van de werkelijkheid. Dit geldt niet voor de realisatie van een constructieve leeromgeving. De score op de schaal 'constructieve leeromgeving' vertoont geen samenhang met sociale wenselijkheid ($r=.02$).

Wanneer we rekening houden met dit laatste gegeven, dan kunnen we op zijn minst concluderen dat het meest vernieuwende element binnen de gewenste didactiek van authentiek leren, te weten het werken met activerende, constructieve opdrachten die betekenis hebben

in het alledaags maatschappelijk leven, het minst gerealiseerd wordt. Een beperkte graad van realisering geldt ook voor het element van samenwerken en communicatie binnen de basisvorming. Rekening houdend met een overschatting van de overige aspecten kunnen we concluderen dat realisatie van de gewenste didactiek in de basisvorming over het geheel genomen beperkt is.

6.2 Invulling van condities

Vervolgens is de vraag aan de orde, welke factoren van invloed zijn op de implementatie van de gewenste didactiek in de klas. Om hier meer zicht op te krijgen is bij de docenten nagegaan in hoeverre aan condities op school- en docentniveau voldaan is, waarvan op basis van innovatieliteratuur verwacht mag worden dat die implementatiebevorderend kunnen werken (vraagstelling 2). Voorts is nagegaan in welke mate deze condities in dit onderzoek daadwerkelijk predictoren vormen voor de implementatie van docentgedrag dat recht doet aan de achterliggende ideeën van Basisvorming: de didactiek van authentiek leren (vraagstelling 3). In het onderstaande gaan we verder in op de resultaten van dit deel van het onderzoek.

Conditie op schoolniveau

Als aspecten van de schoolorganisatie die als voorwaarden kunnen worden beschouwd voor de realisatie van de didactiek van authentiek leren, hebben we de volgende zaken onderscheiden: randvoorwaarden, coördinatie van het vernieuwingsproces en professionalisering van docenten.

Randvoorwaarden

Onder randvoorwaarden rekenen we de waardering voor de methode, de gemaakte rooster-technische keuzes, de kwaliteit van lokalen, en leermiddelen in het algemeen. Tijdens de interviews is nagegaan of (de realisering van) deze randvoorwaarden volgens de docenten bevorderend of belemmerend werken voor de realisatie van de Basisvorming.

Uit de resultaten komt naar voren, dat er grote verschillen bestaan in de mate waarin deze als bevorderend dan wel belemmerend voor de invoering van Basisvorming worden gezien. De kwaliteit van de lokalen (spreiding, inrichting, grootte, mogelijkheid om snel te wisselen) wordt vaker als belemmerend dan als bevorderend gezien (zie Tabel 5).

De (te omvangrijke) groeps-grootte wordt het vaakst genoemd als belemmerende factor (66% van de docenten). Van de leermiddelen

en media worden de audiovisuele middelen en andere (traditionele) leermiddelen vaker als bevorderend dan als belemmerend voor de realisatie van Basisvorming gezien (resp. 46.5% en 49.7%). De methode wordt meestal (bijna 80%) als bevorderend gezien. Daarentegen is de aanwezigheid of beter de afwezigheid en het weinig 'up-to-date' zijn van computers belemmerend (ruim 50%)

Coördinatie en ondersteuning

Voor de implementatie van Basisvorming is daarnaast de vraag van belang op welke wijze het vernieuwingsproces wordt gecoördineerd. Meer in het bijzonder is nagegaan in welke mate overleg op school- en sectieniveau als belemmerend of bevorderend worden ervaren voor de implementatie van de Basisvorming.

Afspraken over lesgeven op schoolniveau worden door slechts 38% van de docenten als bevorderend ervaren, en door 20% als belemmerend. Een groot percentage docenten (71.2%) schat de effecten van overleg en samenwerking met sectiegenoten in als bevorderend voor implementatie van de Basisvorming. Ook afspraken op sectieniveau over lesgeven worden door tweederde van de docenten (64.3 %) als ondersteunend voor de Basisvorming gezien.

Tabel 5

Mate waarin randvoorwaarden en coördinatie en ondersteuning ervaren zijn als belemmerend of bevorderend voor de implementatie van de basisvorming

	belemmerend	bevorderend
Randvoorwaarden	%	%
De grootte van lokalen:	43.4	23.7
De mogelijkheid om van opstelling wisselen	43.9	28.3
De inrichting van de lokalen	42.4	24.7
De spreiding van de lokalen	35.1	14.6
Het gehanteerde lesrooster	13.3	24.3
De groeps-grootte	66.1	16.4
Het aantal uren voor het vak	31.8	35.8
De op school gehanteerde methode	13.4	79.7
Aanwezigheid van traditionele leermiddelen	26.0	49.7
Aanwezigheid van audiovisuele middelen	26.7	46.5
Aantal, bereikbaarheid en het up-to-date zijn van computers	50.3	39.8
Aanwezigheid van schooldocumentatiecentrum	31.9	31.9
Coördinatie en ondersteuning		
Schoolinterne ondersteuning/begeleiding bij invoering Basisvorming	24.2	46.0
Implementatie-bevorderende ondersteuning door de schoolleiding	56.1	20.0
Afspraken op schoolniveau over lesgeven in huidige situatie	20.5	38.0
Afspraken op sectieniveau over lesgeven in huidige situatie	16.7	64.3
Overleg en of samenwerking met sectiegenoten in huidige situatie	18.8	71.2

N.B. N=173; De antwoordcategorie 'geen effect' is weggelaten uit de tabel

Tevens is nagegaan in welke mate docenten gelegenheid hebben gehad tot professionalisering. We onderscheiden hierbij externe ondersteuning, zoals nascholingscursussen en interne ondersteuning. Bijna 33% van de docenten was beperkt of niet in de gelegenheid om gebruik te maken van externe ondersteuning, de rest in ruime mate. Van de feitelijk gebruikte vormen maakten nascholingscursussen en voorlichtingsbijeenkomsten het grootste deel uit (resp. 57.8% en 62.4% van de respondenten noemden deze vormen).

Ondersteuning door de schoolleiding houdt met name in dat de schoolleiding de Basisvorming op de vergaderagenda plaatst, de voortgang controleert en stimuleert, ruimte biedt voor professionalisering. Ten tijde van de invoering van de Basisvorming was er volgens meer dan de helft van de docenten (56.1%) sprake van een implementatie-bevorderende vorm van ondersteuning door de schoolleiding. 20% vindt dat ondersteuning vanuit de schoolleiding belemmerend heeft gewerkt.

Iets minder positief is het beeld als het gaat om andere vormen van schoolinterne ondersteuning en begeleiding. Dit houdt in dat docenten via daartoe in het leven geroepen ondersteunende commissies en groepen geholpen worden bij concrete vraagstukken, zoals realisatie

van de TVS-karakteristieken, de wijze van determineren van leerlingen, realisatie van de leerlingbegeleiding. Van de docenten acht 46% de gegeven begeleiding en ondersteuning op de eigen school bevorderend, 24% acht deze belemmerend voor de realisatie van de Basisvorming.

Conditie op docentniveau: deskundigheidsbeleving en beroepsopvatting in de Basisvorming

Uit de interviews kwam naar voren dat docenten zich gemiddeld redelijk deskundig achten om les te geven in de Basisvorming. Op de samenvattende schaal bestaande uit het gemiddelde op negen uitspraken over lesgeven in de basisvorming (lopend van 1 tot 4) scoren docenten gemiddeld 2.79 ($SD = .56$).

Naast deskundigheid in lesgeven is van belang welke rol docenten aan zichzelf toekennen tijdens lesgeven in de Basisvorming. De Basisvorming veronderstelt een meer begeleidende rol voor de docent, om de leerlingen in een meer actieve (kennis construerende) rol te plaatsen. Desgevraagd geeft 42.2% van de docenten aan de eigen rol primair als die van begeleider van leerprocessen te zien, terwijl 5.8% de docent zich primair als een overdrager van kennis ziet. 43.4% van de docenten ziet zichzelf als begeleider van leerprocessen én als overdrager van kennis.

Tabel 6
Gestandaardiseerde regressiegewichten in de voorspelling van implementatie van authentiek leren

	Aansl	Proc	Sam	Clo
Predictoren	β	β	β	β
Sociale wenselijkheid	.30**	.50**	.19*	
Leermiddelen en randvoorwaarden				
Documentatiecentrum is bevorderend (0=nee, 1=ja)	.14			
Aantal randvoorwaarden als bevorderend ervaren (bereik: 0-7)				.19*
Coördinatie en ondersteuning				
Aantal aspecten interne ondersteuning als belemmerend ervaren (bereik: 0-3)		-.12		
Aantal gevolgd externe ondersteuningsvormen (bereik: 0-4)	.24**			
Deskundigheid en rolopvatting				
Docent ziet zichzelf primair als begeleider (0=nee, 1=ja)	.23**			
De mate waarin docenten zichzelf deskundig achten om les te geven in de Basisvorming (9 items, $\alpha=.76$)	.26**		.44**	.37**
R^2	.21	.36	.24	.19

N.B. * = $p < .05$, ** $p < .01$; Aansl= aansluiting bij leefwereld, Proc= procesgerichte instructie, Sam= samenwerking en communicatie, Clo= constructieve leeromgeving. Bij de dichotome variabelen kan de correlatie opgevat worden als een weergave van een verschil tussen twee groepen

6.3 Verbanden tussen invulling van condities en realisatie van authentiek leren

In Tabel 6 worden de resultaten van multiple regressieanalyses samengevat, waarbij als predictoren de onderscheiden condities op docent- en schoolniveau optreden en de vier subschalen voor authentiek leren als afhankelijke variabelen. In de tabel zijn de gestandaardiseerde regressiegewichten opgenomen. Deze geven het relatieve gewicht van elke voorspeller aan temidden van andere voorspellers. Een positieve waarde wijst op een positieve bijdrage aan de voorspelling van realisatie van een kenmerk van authentiek leren.

Zoals op te maken valt uit Tabel 6, blijkt bij voorspelling van verschillen in procesgerichte instructie en aansluiting bij de belevingswereld de geneigdheid tot het geven van sociaal wenselijke antwoorden de relatief krachtigste predictor te zijn. Voor samenwerking en communicatie geldt eveneens een voorspelkracht zij het in mindere mate. De schaal constructieve leeromgeving blijkt ongevoelig voor sociale wenselijkheid. Nadat rekening gehouden is met de factor sociale wenselijkheid resteert het volgende beeld.

Allereerst is opvallend dat slechts een deel van de in de interviews bevraagde condities voorspellende waarde heeft voor verschillen in realisatie van authentiek leren. Zo heeft de waardering voor de gemaakte afspraken over het lesgeven op school- en sectieniveau voor geen van de aspecten van authentiek leren een significant voorspellende waarde.

Een algemene factor op docentniveau die bij realisatie van drie aspecten van authentiek leren een rol speelt betreft de mate waarin docenten zich deskundig voelen om onderwijs te geven in de Basisvorming. Naarmate de docenten zich meer deskundig voelen sluiten zij meer aan bij de belevingswereld van de leerlingen, realiseren ze vaker een constructieve leeromgeving, en laten ze leerlingen in sterkere mate samenwerken en communiceren.

Verschillen in de realisatie van procesgerichte instructie kunnen daarentegen niet worden teruggevoerd op verschillen in de ervaren eigen deskundigheid. Het zijn vooral factoren op het gebied van coördinatie en ondersteuning op schoolniveau en de rolopvatting van de docent die voorspellend zijn voor verschillen in procesgerichte instructie. Naarmate men meer

externe ondersteuningsvormen heeft gevolgd, is sprake van meer procesgerichte instructie. Docenten die zichzelf beschouwen als begeleider realiseren deze vorm van instructie frequenter dan zij die dit niet doen. Om aan te sluiten bij de belevingswereld blijkt naast deskundigheid een adequaat documentatiecentrum van belang, hoewel het bijbehorende bètagewicht net niet significant is.

De enige significante voorspeller van realisatie van een constructieve leeromgeving is behalve de ervaren deskundigheid, de mate waarin voldaan is aan randvoorwaarden voor lesgeven in de Basisvorming. Deze hebben betrekking op de inrichting, de grootte en spreiding van lokalen, de mogelijkheid om snel van opstelling te wisselen, het lesrooster en groeps-grootte. Bij samenwerking en communicatie blijken opvallend genoeg behalve ervaren deskundigheid geen andere condities op docent- en schoolniveau gepaard te gaan met verschillen in realisatie.

7 Conclusies en discussie

We keren terug naar de vraag in hoeverre docenten een didactiek van authentiek leren realiseren in de basisvorming (onderzoeksvraag 1). De resultaten op de vier onderscheiden aspecten van de didactiek van authentiek leren kunnen we als volgt samenvatten. In de perceptie van de docenten zelf, is de mate van realisatie van de didactiek van authentiek leren sterk verschillend per aspect. Er is gemiddeld nog nauwelijks sprake van een constructieve leeromgeving. Hiermee wordt een leersituatie bedoeld, waarin leerlingen (zelfstandig) werken aan constructieve opdrachten, waarbij de grenzen van het vak, de methode en ook de klassituatie worden doorbroken, en waarbij een relatie wordt gelegd met maatschappelijk functioneren. Het besteden van aandacht aan communicatie en samenwerking komt gemiddeld slechts af en toe voor. De beide overige aspecten van deze didactiek worden echter volgens de docenten in aanzienlijk hogere mate gerealiseerd: docenten sluiten gemiddeld vrij vaak aan bij de leefwereld van de leerlingen en geven gemiddeld vaak procesgerichte instructie.

Gezien de redelijk hoge samenhang tussen de score op de sociale wenselijkheid-schaal en

drie van de vier schalen voor het meten van de didactiek van authentiek leren, moeten we aan-nemen dat de feitelijke realisatie van elementen van deze didactiek lager ligt. Hardere conclusies hierover kunnen getrokken worden als verschillende databronnen worden vergeleken. Zo wordt de validiteit van onze bevindingen verhoogd als we kijken naar andere onderzoeken naar de implementatie van Basisvorming. De conclusie dat van een constructieve leeromgeving nauwelijks sprake is wordt ondersteund door de bevindingen van het onderzoek van de Inspectie van het Onderwijs (1997a, 1997b) waarover ook in dit nummer gerapporteerd wordt. Van den Bergh, Peters-Sips en Zwarts komen namens de inspectie tot de conclusie, dat het bevorderen van een actief leerproces bij de leerlingen slechts in beperkte mate plaatsvindt. Een zelfde conclusie resulteert uit de dieptestudie van Roelofs, Franssen, Houtveen & Lagerweij (elders in dit nummer), in welke studie ook gebruik gemaakt is van lesbezoeken. Meer algemeen zijn lesbezoeken en les-observaties, en analyses van methodegebruik gewenst om een dieper en meer betrouwbaar inzicht te krijgen in de aard van het didactisch handelen van docenten.

Een opvallende bevinding van dit onderzoek is, dat op elk van de besproken aspecten de verschillen tussen de gemiddelde scores van de docenten Nederlands en wiskunde significant zijn. De richting van de verschillen is echter niet steeds gelijk voor alle aspecten van authentiek leren, hetgeen kan erop kan wijzen dat de uitwerking van authentiek leren vakgebonden is. Docenten Nederlands scoren gemiddeld significant hoger dan wiskundeleraars op drie van de vier onderzochte aspecten: aansluiten bij de leefwereld van de leerlingen; samenwerking en communicatie; en de mate van realisatie van een constructieve omgeving. Wiskundeleraars geven in hogere mate proces-gerichte instructie dan docenten Nederlands. Een voor de hand liggende verklaring voor de sterkere procesgerichtheid bij wiskunde is dat in dit vakgebied al langer aandacht gevraagd wordt voor oplossingsstrategieën als een kern van het wiskundeonderwijs. Niettemin zijn ook bij Nederlands tal van strategieën denkbaar waarop instructie en begeleiding zich zouden kunnen richten, zoals het afleiden van

hoofddedachten uit een tekst, het opstellen van persoonlijke en zakelijke brieven.

Opmerkelijk is dat hoewel het Nederlandse wiskundeonderwijs al enige tijd een ontwikkeling kent richting contextrijk onderwijs, als contextrijk bekend staat, onder meer door het gedachtegoed van Freudenthal (1973, 1991), het feitelijke karakter van het wiskundeonderwijs niet sterk leefwereldgericht is. De gerichtheid op de leefwereld is beduidend minder sterk dan bij Nederlands. Een verklaring hiervoor kan zijn dat bij de uitwerking van de vakdidactiek wiskunde dagelijkse contexten weliswaar als ondersteuning worden gehanteerd bij leerprocessen, maar dat uiteindelijk van de leerlingen wordt verlangd kennis in formele abstracte klassituaties toe te passen. Mogelijk zijn vakinhouden van Nederlands meer ingebed in het dagelijks leven van leerlingen. De kerndoelen liggen op het terrein van dagelijkse mondelinge en schriftelijke communicatie en zijn wellicht daardoor per definitie meer leefwereldgericht. In verband hiermee is ook de hogere mate van samenwerking en communicatie bij Nederlands niet verbazingwekkend, aangezien de kerndoelen van het vak zelf gericht zijn op communicatieve kennis en vaardigheden, waarvoor nu eenmaal een wederzijdse betrokkenheid tussen leerlingen nodig is.

Hoewel verschillen tussen schooltypen niet centraal stonden in deze studie, willen we toch wijzen op het feit dat nergens significante verschillen werden aangetroffen in de mate van realisatie van de didactiek van authentiek leren tussen de schooltypen (VBO; MAVO; HAVO; VWO) waarvoor de vragenlijsten werden ingevuld. Of onderwijs nu aan leerlingen van lagere of hogere schooltypen wordt gegeven, het kenmerkt zich door een gerichtheid op basisvaardigheden, waarbij voor zelfstandige kennisconstructie in complexe taaksituaties weinig plaats is. Verwacht zou mogen worden dat aan VWO-leerlingen ander onderwijs wordt gegeven dan aan VBO-leerlingen, ervan uitgaande dat voor deze groepen een verschillend regime (Boekaerts & Simons, 1993; Simons, 1995) passend zou zijn. Mogelijk heeft dit resultaat te maken met de keuze van de vakken, die met name cognitieve vaardigheden betreffen. Zo blijkt uit een dieptestudie door Franssen, Eijkelhof, Houtveen en Duijmelinck (1996) dat

het didactisch handelen bij technieklessen wel verschilt tussen docenten van verschillende schooltypen. Franssen et al. vonden dat het didactisch handelen in VBO-scholen voornamelijk was op verwerving van praktische vaardigheden, terwijl op gymnasia een sterk accent gelegd werd op de cognitieve aspecten van het vak.

We mogen concluderen dat, ondanks duidelijke verschillen tussen vakken, de destijds wenselijk geachte didactische karakteristieken 'toepassen', 'vaardigheden', 'samenhang' in de Basisvorming slechts mondjesmaat worden gerealiseerd. Met name de combinatie van deze karakteristieken, in ons onderzoek aangeduid als 'constructieve leeromgeving', komt nauwelijks van de grond.

Ter verklaring van deze beperkte realisatie zijn verschillende factoren aan te wijzen. Daartoe zijn condities op docent- en schoolniveau onderzocht en hun samenhang met realisatie (vraagstelling 2 en 3). Het blijkt dat de relevantie van condities op school- en docentniveau varieert per te realiseren aspect van authentiek leren. De meest consistente conditie is er een op docentniveau: een hoog competentiegevoel. Deze factor is ook veelvuldig in ander onderzoek aangetoond (Fullan, 1992; Fullan & Pomfret, 1977; Gibson & Dembo, 1984; Imants & van Zoelen, 1995). Docenten die het gevoel hebben voldoende deskundig te zijn om les te geven in de basisvorming realiseren in hogere mate een didactiek van authentiek leren, dan zij die dit gevoel minder hebben. Dit geldt met name voor het meest complexe en omvattende aspect van authentiek leren, de realisatie van een constructieve leeromgeving. Voor de realisatie van dit aspect blijken voorts materiële randvoorwaarden van cruciaal belang: goede kwaliteit van lokalen, een passend lesrooster en een niet te grote groep.

Voor de realisatie van procesgericht onderwijs blijken roloppvatting van docenten en het gevolgd hebben van externe ondersteuning van belang te zijn.

Een vraag willen we nog kort aanstippen aan het slot van dit artikel: hoe uniform zijn de onderscheiden kenmerken van authentiek leren toe te passen op verschillende leersituaties?

Leersituaties kunnen zoals gezien verschillen naargelang de vakinhoud, maar ook als gevolg van leerlingkenmerken, kenmerken van docent en van de schoolorganisatie. Anders geformuleerd: bestaan er verschillende wenselijke configuraties van kenmerken van authentiek leren gegeven al deze kenmerken? Een rigide prescriptie van onderwijsleersituaties uitgaande van de vier kenmerken van authentiek leren doet geen recht aan de vraag naar adaptiviteit van onderwijs. Veeleer zouden de kenmerken van authentiek leren gehanteerd moeten worden als dimensies waarop instructie, begeleiding en beoordeling kunnen variëren, afhankelijk van behoeften van leerlingen en wellicht ook te verwerven leerdoelen. Docenten kunnen daarbij hun aanpak variëren tussen constructie en imitatie, tussen leefwereld en formele vakstructuren, tussen individueel en coöperatief leren, en tussen complexe taken en enkelvoudige leertaken. Een criterium voor goed onderwijs zou daarbij kunnen zijn of leerlingen in staat zijn de opgedane kennis en vaardigheden in vele verschillende betekenisvolle situaties toe te passen (zie ook Roelofs & Terwel, 1999). Een verdere verfijning van het begrippenapparaat rondom authentiek leren is dan ook gewenst. Van daaruit kunnen didactische arrangementen worden ontwikkeld waarbij naast een aantal algemene richtlijnen recht wordt gedaan aan vele bronnen van verschillen die het onderwijs kent van vakspecifieke, leerlingsspecifieke, docentspecifiek en schoolspecifieke aard.

Noten

- 1 Ook in het beleidsmatige denken over leren en onderwijzen worden de geschetste ontwikkelingen meegenomen. De Advies Raad Onderwijs (1994) formuleerde in haar advies 'Ruimte voor leren' het ontwikkelen van kennis als gereedschap als een hoofdtaak van het funderend onderwijs. De raad formuleert drie principes aan de hand waarvan deze taak gerealiseerd kan worden: 1) leren als uitbouw en reconstructie van kennis, 2) leren als sociaal proces en 3) leren problemen aan te pakken op eigen kracht.
- 2 Overigens is er geen sprake van hét constructivisme, maar bestaan daarbinnen verschillende stromingen (Phillips, 1995).

- 3 We willen benadrukken dat de operationalisering niet opgevat moet worden als een rigide prescriptie van onderwijsleersituaties uitgaande van vier kenmerken van authentiek leren. De kenmerken van authentiek leren zijn bedoeld als een dimensie of meetlat waarop instructie, begeleiding en beoordeling kunnen variëren, afhankelijk diverse factoren zoals behoeften van leerlingen en te verwerven leerdoelen.
- 4 De keuze voor de betrokken vakgebieden is bepaald door de opdrachtgever om een vergelijking mogelijk te maken met resultaten uit de VOCL-studies.

Literatuur

- Advies Raad Onderwijs [Advisory Council for Learning] (1994). *Ruimte voor leren*. Utrecht: ARO.
- Boekaerts, M. & Simons, P.R.J. (1993). *Leren en instructie: psychologie van de leerling en het leerproces*. Assen: Dekker & Van de Vegt.
- Bolhuis, S. & Kluvers, C. (1998). Procesgericht onderwijs. In J. Vermunt & L. Verschaffel, *Onderwijzen van kennis en vaardigheden. Onderwijskundig Lexicon*, editie III, pp. 87-107. Alphen aan den Rijn: Samson H.D. Tjeenk Willink
- Brown, A.L. & Campione, J.C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (229-270). Cambridge, MA: MIT Press/Bradford Books.
- Brown, J.S., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18 (1), 32-41.
- Bruer, J.J. (1993). *Schools for thought. A science of learning in the classroom*. Cambridge, Mass.: Bradford.
- Corte, E. de (1990a). Ontwerpen van krachtige leeromgevingen. In M.J. Ippel & J.J. Elshout (Red.). *Training van hogere-orde denkprocessen* (pp. 133-147). Bijdragen aan de onderwijsresearch, nr. 32. Amsterdam: Swets & Zeitlinger.
- Corte, E. de (1990b). Learning with new information technologies in schools: perspectives from the psychology of learning and instruction. *Journal of Computer Assisted Learning*, 6, 69-87.
- Essen, A. van, (1990). Three decades of foreign-language teaching in the Netherlands. *Toegepaste Taalwetenschap in Artikelen (applied Linguistics in Articles)*, 36, 72-85. ERIC reproduction services, no.338021.
- Franssen, H.A.M., Eijkelhof, H.M.C., Houtveen, A.A.M., & Duijmelinck, H.J.T. (1996). *Techniek in de basisvorming*. Utrecht: ISOR.
- Freudenthal, H. (1973) *Mathematics as an educational task*. Dordrecht: Reidel.
- Freudenthal, H. (1991) *Revisiting Mathematics Education. China Lectures*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Fullan, M.G. (1992). *Successful School Improvement: The implementation perspective and beyond*. Buckingham, Philadelphia: Open University Press.
- Fullan, M. & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research*, 47 (1), 335-397.
- Gibson, S. & Dembo, M. (1984). Teacher efficacy. A construct validation. *Journal of Educational Psychology*, 76, 569-582.
- Greeno, J.G. (1988). *A perspective on thinking*. California: Institute for Research on Learning.
- Greeno, J.G. (1997). On claims that answer the wrong questions. *Educational Researcher*, 26(1), 5-17.
- Houtveen, Th., Vermeulen, C. & Grift, W., Van de (1993). *Bouwstenen voor onderzoek naar de kwaliteit van scholen. Constructie van een instrument voor het meten van het onderwijsleer-klimaat*. Utrecht: ISOR.
- Imants, J., & Zoelen, A. van (1995). Teachers' sickness absence in primary schools, school climate and teachers' sense of efficacy. *School Organisation*, 15, 77-86.
- Inspectie van het Onderwijs (1997a). *Inventarisatie van onderzoek naar de Basisvorming*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (1997b). *De basis gelegd. Evaluatie van school-organisatorische aspecten van de Basisvorming*. Utrecht: Inspectie van het Onderwijs.
- Jong, F.P.C.M. de (1992). *Zelfstandig leren: Regulatie van het leerproces en leren reguleren, een procesbenadering*. Proefschrift. Tilburg: Katholieke Universiteit Brabant.
- Lamberigts, R., Brok, P.J. den, Derksen, K., & Bergen, T. (1999). Het concept activerende instructie gemeten via de perceptie van de leerlingen. *Pedagogische Studiën*, 76(1), 36-50.
- Lave, J. (1991). Situating learning in communities of practice. In L.B. Resnick, J.M. Levine, & S.D. Teasley (Eds.), *Perspectives on socially shared cognition* (63-82). Washington DC: American Psychological Association.

- Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Newmann, F.M., Marks, H.M. & Gamoran, A.G. (1996). Authentic pedagogy and student performance. *American Journal of Education*, (104), 280-312.
- Phillips, D.C. (1995). The good, the bad, and the ugly: the many faces of constructivism. *Educational Researcher*, 24 (7), 5-12.
- Procesmanagement Basisvorming (1993a). *Van kerndoel naar praktijk, deel A*. Almere: Procesmanagement Basisvorming.
- Procesmanagement Basisvorming (1993b). *TVS-karakteristieken, deel B*. Almere: Procesmanagement Basisvorming.
- Resnick, L. (1987). Learning in school and out. *Educational Researcher*, 16 (9), 13-19.
- Roelofs, E.C. & Franssen, H. (1994). *Bevordering van toepassingsgerichtheid door authentiek leren*. In J. Ahlers et al. (Red.), *Handboekbasisvorming voor de praktijk van het voortgezet onderwijs*. Houten: Bohn Stafleu Van Loghum.
- Roelofs, E.C., Franssen, H., Grootsoorten, J. (1996). *Basisvorming op de voet gevolgd. Deelrapport in Utrechts onderzoek naar de invoering van de basisvorming. Toepassingsgerichtheid en authentiek leren na één jaar basisvorming*. Utrecht: ISOR/Vakgroep Onderwijskunde.
- Roelofs, E.C. & Terwel, J. (1999). Constructivism and authentic pedagogy: State of the art and recent developments in the Dutch national curriculum in secondary education. *Journal of Curriculum Studies*, 31(2), 201-227.
- Roelofs, E.C., Vermeulen, C.J. & Houtveen, A.A.M. (1998). *Basisvorming op weg Utrecht*: ISOR.
- Simons, R.J. (1995). Leerlingenmerken. In N. Verloop, & J. Lowyck, (red.), *Onderwijskunde. Een kennisbasis voor professionals* (pp. 15-38). Groningen: Wolters-Noordhoff.
- Tweede Kamer (1991). *Wijziging van onder meer de Wet op het voortgezet onderwijs en de Wet op het basisonderwijs in verband met de invoering van basisvorming in het voortgezet onderwijs, de invoering van kerndoelen in het basisonderwijs en voortgezet onderwijs en de invoering van het voorbereidend beroepsonderwijs* (kamerstuk 20381, nr. 141). 's Gravenhage: Staatsuitgeverij.
- Vanderbilt Cognition and Technology Group (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher*, 19 (6), 2-9.
- Veenman, S.A.M. (1992). Effectieve instructie volgens het directe instructiemodel. *Pedagogische Studiën*, 69(4), 1-28.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs: Naar procesgerichte instructie in zelfstandig denken*. Amsterdam: Swets & Zeitlinger.
- Verschaffel (1995). Beïnvloeden van leerprocessen [Influencing learning processes] In J. Lowyck, N. Verloop & W.Th.J.G. Hoeben (Red.), *Onderwijskunde: Een kennisbasis voor professionals* (152-187). Groningen: Wolters-Noordhoff.

Manuscript aanvaard 6-4-1999

Auteurs

E.C. Roelofs is universitair docent-onderzoeker aan de capaciteitsgroep Onderwijskunde van de Universiteit Utrecht

A.A.M. Houtveen is senior onderzoeker bij het ISOR, Instituut voor Onderwijsresearch, Universiteit Utrecht.

Adres: Capaciteitsgroep Onderwijskunde, Universiteit Utrecht, Postbus 80.140, 3508 TC Utrecht

Abstract

Authentic pedagogy in Dutch secondary education: a state of the art

E.C. Roelofs en A.A.M. Houtveen. Pedagogische Studiën, 1999, 76, 237-257.

The national curriculum for the first stage of Dutch secondary education, Basic Education, implies apart from a change in the educational content, also a change in the instructional process. The teacher is expected to create a learning environment that asks for active learners engaging in meaningful and strategic learning. The main question to be addressed is to what extent do teachers realize authentic instruction, to accomplish this kind of learning environment. In addition, the relation between the implementation of authentic instruction on the one hand and conditions at schools level and teacher level on the other hand was studied.

Based on a literature survey a questionnaire and

an interview guide have been developed, covering characteristics of authentic instruction, and school and teacher level factors considered important for implementation. 88 Dutch language teachers and 85 math teachers from a representative sample of Dutch secondary schools were included in the study. Data collection took place in the fall of 1996. Although the results differ across schools subjects, it is found that three aspects of authentic instruction are implemented to a reasonable extent: referring to students' personal worlds, using process-oriented instruction, and sometimes using forms of cooperative learning. However, teachers seldom create learning environments in which the boundaries between school subjects, between in and out of school learning are broken, and the content matter in textbooks is transcended.

Realization of a constructive learning environment, regarded as the most important innovative pedagogical characteristic of Basic Education is related to the fulfillment of preconditions regarding the quality of classrooms and (modern) media present, and with the extent to which teachers feel capable of teaching in the innovative aspects of Basic Education. An instructional approach is discussed, that takes into account problems of student, teacher and school characteristics, and still is aimed at meaningful and strategic learning, without falling into the dichotomy of transmission versus construction of knowledge.

Een dieptestudie naar authentiek leren in de Basisvorming

Docentgedrag, methodengebruik, en leerlingpercepties

E.C. Roelofs, H.A.M. Franssen, A.A.M. Houtveen & N.A.J. Lagerweij*

Samenvatting

Als onderdeel van een longitudinale dieptestudie op drie brede scholengemeenschappen is onderzocht in welke mate docenten hun didactisch handelen in de Basisvorming vormgeven volgens kenmerken van authentiek leren, een didactisch concept waarbinnen de gewenste veranderingen zijn te plaatsen. Vier kenmerken van de didactiek van authentiek leren stonden centraal: 1) constructie van kennis in complete taaksituaties, 2) gerichtheid op de leefwereld van leerlingen, 3) waarde van leeractiviteiten voor situaties buiten school, 4) communicatie en samenwerking. Gebruik is gemaakt van verschillende databronnen: in 1994 en 1996 zijn vragenlijsten bij docenten voor alle bavo-vakken afgenomen. Daarnaast zijn leerlingen bevraagd over de lessen Engels en wiskunde. Begin 1995 zijn docenten wiskunde en Engels geïnterviewd over hun methodegebruik tegen de achtergrond van kenmerken van de didactiek van authentiek leren. Resultaten uit de verschillende databronnen wijzen uit dat geen van de scholen hoog scoort op kenmerken van de didactiek van authentiek leren. De methode biedt volgens docenten weinig gelegenheid tot authentiek leren. Ze passen de methode hierop echter niet aan. Verder blijkt dat in beperkte mate wordt voldaan aan condities voor authentiek leren op klas- en schoolniveau.

1 Inleiding

De Basisvorming betekende in 1993 allereerst een verandering in het onderwijsaanbod voor de eerste fase van het voortgezet onderwijs. Destijds zijn voor 15 vakken kerndoelen vastgesteld, die door alle leerlingen in twee tot vier jaar bereikt zouden moeten worden. Daarnaast werd ook een verandering verwacht in de aard van het onderwijsleerproces. Hoewel in de wet geen didactische werkwijzen worden voorgeschreven, heeft het Procesmanagement Basisvorming (PMB) in 1993 drie kernkarakteristieken geformuleerd, aan de hand waarvan wenselijke didactische activiteiten konden worden afgeleid: toepassen, vaardigheid en samenhang (TVS).

Onderwijs volgens deze kernkarakteristieken houdt in dat de te verwerven kennis levensrecht moet zijn, dat de leerling er iets mee moet kunnen in zijn eigen persoonlijk leven en in de maatschappij waarin hij leeft. Daarnaast wordt een gerichtheid op vaardigheden bepleit, waarbij gedoeld wordt op te besteden aandacht aan bijvoorbeeld het proces van kennisverwerving, samenwerken en onderzoek verrichten. Tot slot wordt samenhang verlangd tussen het onderwijsaanbod van verschillende vakken (bijvoorbeeld via vakoverstijgende thema's), tussen binnen- en buitenlesactiviteiten en tussen onderwijs en begeleiding.

In dit artikel wordt verslag gedaan van een onderzoek naar de realisering van onderwijs volgens de TVS-karakteristieken op drie grote scholengemeenschappen in de periode 1994-1996. Het onderzoek vormt een onderdeel van een longitudinale dieptestudie naar de invoering van de basisvorming in de periode van 1993 tot 1997.

* De volgende leden van het onderzoeksproject is dank verschuldigd: dr. R. de Jong, prof. dr. G. Kantselaar, dr. J. Van der Linden, dr. J. Voogt, prof. dr. L. Vriens, drs. O. Treep, dr. L. Van Wessum en dr. R. Wierstra. Dank gaat uit naar de toenmalige student-assistenten van die het project hebben ondersteund: drs. Judith Grootsholten, drs. Ilse van Eekelen, drs. Dana Janssen, drs. Camille Rietmeijer, drs. Karin Spanjersberg, drs. Viva Wessels.

2 Theoretische achtergrond

Elders in dit nummer (Roelofs & Houtveen) wordt een uitgebreide analyse gegeven van de relatie tussen de voor basisvorming wenselijk geachte TVS-karakteristiek en hedendaagse ontwikkelingen in de wetenschappelijke benadering van onderwijzen en leren. De gevolgen van veranderd denken over leren en onderwijzen voor het didactisch handelen zijn daar samengevat met behulp van vier kenmerken, die te samen een didactiek van authentiek leren kenmerken: 1) constructie van kennis in complete taaksituaties, 2) gerichtheid op de leefwereld, 3) relevantie van leren voor buitenschoolse situaties en 4) communicatie en samenwerking (zie ook Roelofs, Franssen & Grootsholten, 1996; Roelofs & Terwel, 1999). Elk van de kenmerken is gerelateerd aan drie aspecten van didactisch handelen: instructie en begeleiding, de keuze van leertaken en de beoordeling van leerresultaten.

Samenvattend, betreft het eerste kenmerk van de didactiek van authentiek leren het streven naar constructie van kennis in complete taaksituaties. Dit betekent dat leerlingen te maken krijgen met een betekenisvolle taak die niet bij voorbaat wordt versnipperd in deeltaken. Instructie en begeleiding hebben een procesgericht karakter waarbij sprake is van verwoording van denkprocessen en reflectie op regulatieprocessen. De rol van de docent is hierbij begeleidend van aard. Bij deze wijze van kennisverwerving passen beoordelingsvormen die een continu en persoonlijk karakter hebben, zoals het geval is bij gebruik van portfolio's of leerlingdossiers.

Het tweede kenmerk is de gerichtheid op de leefwereld van leerlingen en hun persoonlijke voorkennis. Deze houdt dit in dat docenten bij de keuze van hun lesonderwerpen, opdrachten en leertaken aansluiten bij interesses van leerlingen. Er is aandacht voor preconcepties van leerlingen, die misconcepties kunnen blijken te zijn. Ook bij beoordeling van leerresultaten worden opdrachten gehanteerd die verband houden met de leefwereld van leerlingen.

Het derde kenmerk van de didactiek van authentiek leren is de aandacht voor de betekenis van leerervaringen en opbrengsten voor de wereld buiten school. Leeropbrengsten reiken in de optiek van de didactiek van authentiek

leren verder dan het verdienen van een schoolcijfer, gegeven door de docent. Bij taken en opdrachten worden leerlingen in de rol van kennisgebruiker geplaatst. Bij instructie en begeleiding legt de docent contacten met de actualiteit en hanteert hij verschillende media waarlangs contacten met de wereld buiten school mogelijk zijn (computer, radio, televisie). Leerlingen kunnen bij het leren gebruik maken van deze media en van hulpmiddelen die professionals ze ook hanteren (bijvoorbeeld een spreadsheet programma).

Het vierde kenmerk van de didactiek van authentiek leren, communicatie en samenwerking, houdt in dat kennisverwerving gebeurt door middel van leertaken waarin sprake is sprake van onderlinge interactie en samenwerkend leren. De docent biedt de faciliteiten voor communicatie en samenwerking, bewaakt het groepsproces, en spreekt leerlingen aan op de door hen gevolgde aanpak. Bij de beoordeling van groepsresultaten zijn zowel leraar als leerlingen betrokken, en wordt rekening gehouden met zowel de individuele bijdrage als de groepsbijdrage.

De veronderstelling is dat door de didactiek van authentiek leren (hierna afgekort als 'authentiek leren') de verworven kennis en inzichten lang beklijven en flexibel worden toegepast in uiteenlopende nieuwe situaties. Deze veronderstelling krijgt geleidelijk aan een empirische basis. In de Verenigde Staten zijn een aantal inmiddels beroemde onderwijsprojecten opgezet, waarin leerlingen leren volgens een didactiek van authentiek leren en de bedoelde effecten ook optreden. In deze projecten wordt de klas gekenmerkt als een lerende gemeenschap, waarin leerlingen leren aan de hand van voor hen betekenisvolle problemen. Er zijn multi-mediale communicatieve leeromgevingen en leerlingen hebben snel toegang tot verschillende informatiebronnen. Verwerven van kennis is er een persoonlijke, actieve en strategische activiteit (Brown & Campione, 1994; Lamon, 1995; Scardamalia, Bereiter, McLean & Woodruff, 1989).

Ontwikkelingen in vakdidactieken Engels en wiskunde

De wenselijk geachte didactiek kan per vak verschillend worden uitgewerkt. In de toelichting op de kerndoelen per vak kunnen elementen van authentiek leren worden herkend.

Een van de kerndoeldomeinen Moderne vreemde talen heeft betrekking op 'gebruik van compenserende strategieën om met een beperkte taalvaardigheid om te gaan (waaronder opzoekvaardigheden)'. Dit domein verwijst naar constructie van kennis. De domeinen 'aandacht voor land en volk van Engels spreken' en 'oriëntatie op de zin van het leren van vreemde talen' wijzen op aandacht voor het belang van leren voor buitenschoolse situaties (zie Lenaerts, 1992). In het hedendaagse moderne taalonderwijs ligt het accent minder op het kennen van de grammatica en meer op taalverwerving in communicatieve situaties. Leren van een taal vindt plaats door gebruik binnen een interactieve context. Leerlingen moeten met elkaar in een vreemde taal leren spreken en waar mogelijk communiceren met een 'native speaker'. Nieuwe werkvormen, zoals het luisteren naar uitzendingen in een vreemde taal op de radio, het bekijken en bediscussiëren van buitenlandse tv-programma's en het schrijven van brieven, passen in deze opvatting. Het gaat vooral om de vraag of voorzien wordt in een communicatieve behoefte (Van Essen, 1990; Van Els & Buis, 1987).

In de wiskundendidactiek zien we een verschuiving van de structuur van het vak in de richting van realistische wiskunde, wiskunde als menselijke activiteit. Deze verschuiving is te herkennen in de toelichting op de kerndoelen (Ten Hove & Van der Zwaart, 1992). In de meeste kerndoeldomeinen wordt veelvuldig verwezen naar gebruik van kennis in praktische situaties. Dit wijst op een versterkte aandacht voor de leefwereld en gebruik van kennis buiten de school. Het domein 'geïntegreerde wiskundige activiteiten' (waaronder o.a. projectwerk) wijst op kennisconstructie in complete taaksituaties. Deze kenmerken van wiskundeonderwijs kennen in Nederland al een langere traditie, geïnspireerd door het werk van Freudenthal (1973, 1991). Bij wiskundeonderwijs gelden principes als het opnieuw ontdekken, uitvinden van principes, het ontwikkelen van persoonlijke oplossingsstrategieën. Door structureren, decontextualiseren en generaliseren begeleidt de docent de leerling naar een meer formeel begrip van wiskundige concepten (zie Gravemeijer, 1995; Van der Kooij, 1995).

3 Onderzoeksvragen

In dit artikel doen we verslag van een onderzoek naar de mate waarin docenten erin slagen kenmerken van authentiek leren te realiseren tijdens de invoering van basisvorming in de klas. De veronderstelling is dat naarmate onderwijs meer voldoet aan kenmerken van authentiek leren, er meer sprake is van toepassingsgericht onderwijs. Het onderzoek is verricht in het kader van het BAVO-project, een longitudinale dieptestudie (1993-1997) op drie brede scholengemeenschappen (Lagerweij, 1993; Voogt & Franssen, 1994; Lagerweij et al., 1996). In onderhavige studie is gebruik gemaakt van verschillende databronnen: zelf gerapporteerd onderwijsgedrag, percepties van leerlingen en gegevens over het gebruik van de methoden voor Engels en wiskunde. De volgende vragen staan centraal:

1. In welke mate komen kenmerken van authentiek leren voor in de eerste fase van het voortgezet onderwijs resp. één (1994) en drie jaar (1996) na de invoering van de basisvorming, gelet op percepties van docenten en leerlingen inzake het didactisch handelen in het algemeen en het methodegebruik bij Engels en wiskunde in het bijzonder?
2. In welke mate worden condities voor authentiek leren vervuld volgens docenten?

4 Methode

4.1 Onderzoeksgroep

In de schooljaren 1993-1994 en 1995-1996 hebben docenten (resp. $n=254$ en $n=317$) en leerlingen (resp. $n=906$, $n=1992$) van drie brede scholengemeenschappen deelgenomen aan een enquêtestudie. Daarnaast was een klein deel van de docenten wiskunde en Engels ($n=18$) in 1994-1995 betrokken bij een interviewstudie inzake methodegebruik. De resultaten van deze studies worden in dit artikel gecombineerd besproken.

De docenten en leerlingen zijn afkomstig uit drie brede scholengemeenschappen. School A is een openbare scholengemeenschap, gesticht in 1983, vanuit een onderwijsvisie die stelde op de ideeën van de Middenschool. Er bestaan

vijf niveaustromen: ivbo, vbo, mavo, havo, vwo. In de school werken 100 docenten, 5 directieleden en 20 personen die het onderwijs ondersteunen (e.g. technische en administratieve ondersteuning). De school wordt bevolkt door ongeveer 1500 leerlingen, zo'n 10% daarvan is allochtoon.

School B is een rooms-katholieke scholengemeenschap met vijf niveaustromen: ihno, lhno, mavo, havo, vwo. De school is in 1985 ontstaan uit een fusie van twee scholen en heeft ten tijde van het onderzoek ongeveer 75 docenten, 6 directieleden en 12 onderwijsondersteunende personen in dienst. De leerlingpopulatie van ongeveer 1000 leerlingen is volledig autochtoon.

School C is een scholengemeenschap op protestant-christelijke grondslag met vier niveaustromen: vbo, mavo, havo, vwo. De identiteit is weloverwogen gekozen en speelt in alle facetten van het schoolleven een beslissende rol. De school is in 1984 gesticht en sindsdien snel gegroeid tot 1400 leerlingen. In het begin van schooljaar 1994-1995 is een tweede locatie gestart met vijf eerste klassen. In de school werkten in het schooljaar 1992-1993 ongeveer 80 docenten, 5 directieleden en 14 personen die het onderwijs ondersteunen.

4.2 Instrumenten en dataverzameling

Authentiek leren is geoperationaliseerd aan de hand van de eerder beschreven vier kenmerken. Twee typen vragenlijsten werden ontwikkeld, waarin indicatoren voor de vier kenmerken zijn uitgewerkt: een docentenlijst voor docenten van alle vakken en een leerlingenlijst voor de vakken Engels en wiskunde.

In de docentenlijst wordt gevraagd naar meningen, intenties en zelf gerapporteerd gedrag inzake de kenmerken van authentiek leren. De docentenlijst bestaat uit twee onderdelen: 1) de realisatie van kenmerken van authentiek leren, 2) condities en knelpunten op klas- en schoolniveau met betrekking tot authentiek leren.

Het onderdeel 'realisatie van kenmerken van authentiek leren' start met een korte beschrijving van een leersituatie, waarin kenmerken van authentiek leren in hoge mate aanwezig zijn. Dit om docenten een referentiepunt te geven voor het beantwoorden van de vragen. Vervolgens worden vragen gesteld die op een

zes-puntsschaal (1=nooit, 6= altijd) kunnen worden beantwoord over het proces van instructie en begeleiding, de opgegeven (huiswerk)taken, en de manier waarop leerresultaten worden beoordeeld. Alle vragen over één kenmerk van authentiek leren vormen een schaal. Schaalgegevens zijn vermeld in Tabel 1.

Een voorbeelditem uit de schaal *constructie van kennis* is: 'In mijn lessen komen de aanpak en de te zetten stappen bij het uitvoeren van opdrachten aan de orde, doordat ik de leerlingen uitnodig hun manier van denken te vertellen.' Een voorbeeld uit de schaal *gerichtheid op de leefwereld* is: 'Ik kies bij mijn lessen onderwerpen uit de leefwereld van mijn leerlingen.' Een voorbeelditem uit de schaal *waarde van leeractiviteiten buiten school*: 'Het is mogelijk taken te geven aan leerlingen die gericht zijn op het leveren van een bijdrage aan de oplossing van een maatschappelijk probleem. Ik geef dergelijke opdrachten, waarbij instanties van buiten de school commentaar leveren'. Een voorbeelditem uit de schaal *communicatie en samenwerking* is: 'Ik geef de leerlingen de gelegenheid om gegevens en/of resultaten die ze op grond van opdrachten hebben verkregen, aan elkaar te presenteren.'

In het tweede gedeelte worden docenten bevraagd over de mate waarin condities voor authentiek leren op klas- en schoolniveau zijn vervuld, respectievelijk welke problemen en knelpunten zij in verband hiermee zien.

De docentenlijst is in 1994 ($n=254$) en 1996 ($n=317$) afgenomen bij het voltallige onderwijzend personeel van de drie deelnemende scholen. De responspercentages op de docentenvragenlijst bedroegen op school A in 1994 20% en in 1996 36%, hetgeen laag genoemd kan worden. Op school B bedroegen deze percentages respectievelijk 71% en 79%, op school C respectievelijk 84% en 90%. Redenen voor non-response van met name school A waren: de moeilijkheidsgraad van de vragen, de afstand tussen de vragenlijst en de realiteit, en geen tijd.

In de leerlingenlijsten wordt aan leerlingen gevraagd hoe zij het didactisch handelen van hun docent tijdens lessen Engels en wiskunde ervaren. Ter oriëntatie op de vragen is in de lijsten een klassituatie beschreven. In de klassituatie zijn kenmerken van authentiek leren ver-

werkt. Bij elke vraag wordt steeds een aspect uit deze klassituatie herhaald, waarna de leerling de vraag krijgt of de beschreven situatie ook in zijn klas voorkomt. Eerste versies van de lijst zijn afgenomen bij en besproken met leerlingen van 12 tot 14 jaar en hun docenten. Bijstellingen vonden plaats op grond van leesbaarheid en begrijpelijkheid. Een voorbeeld-item (wiskundelijst): 'Bij Petra in de klas gebruiken de leraar en de leerlingen dingen van buiten de school die met de opgaven te maken hebben. De leerlingen hebben wel eens allemaal iets van huis meegenomen met een ronde vorm, zoals een bierviltje, een blikje of een ring. Ze moesten toen de omtrek en de diameter van de voorwerpen berekenen. Nemen jullie ook dingen van buiten school mee die met wiskunde te maken hebben?' De antwoordmogelijkheden lopen van 1 ('nooit') tot 5 ('heel vaak').

Voor de onderscheiden vier kenmerken van authentiek leren konden geen betrouwbare subschalen worden gevormd. Wel zijn redelijke betrouwbare totale schalen gevormd, aangeduid als 'realisatie didactiek van authentiek leren'. Voor wiskunde bedraagt de betrouwbaarheidscoëfficiënt α .77 (20 items), voor Engels .78 (22 items).

In 1994 en 1996 kreeg de helft van alle eerste klassen de leerlinglijst voor wiskunde voorgelegd, de andere helft die voor Engels, evenredig verdeeld over schooltypen (mavo, havo etc.). In 1996 werd de lijst bovendien afgenomen in de leerjaren 2 en 3. Op verzoek van de docenten zijn de leerlinglijsten in 1995-1996 zoveel mogelijk bij andere docenten afgenomen om de belasting van docenten te spreiden. Dit had als nadeel dat vergelijking van resultaten uit leerjaar 1 betrekking heeft op verschillende docenten. De verdeling van leerlinglijsten over klassen- met resulterende steekproefgrootten is weergegeven in bijlage 1.

Voorts is een leidraad ontwikkeld ten behoeve van een semi-gestructureerd interview met 9 docenten Engels en 9 docenten wiskunde uit leerjaar 1, halverwege het schooljaar 1994-1995. Daarbij stonden de vier kenmerken van authentiek leren in relatie tot methodegebruik centraal. Gevraagd werd in welke mate docenten in hun methodegebruik bij Engels en wiskunde maatregelen nemen ter realisatie van

deze kenmerken. Daarnaast werd gevraagd welke mogelijkheden de methode hiertoe biedt en hoe de docent de methode in dit opzicht hanteert. Ook is gevraagd naar de eigen opvattingen en motieven inzake methodegebruik. In het interview zijn naast algemene vragen over de kenmerken per vak een achttal vakspecifieke vragen gesteld. Deze vragen zijn afgeleid uit de toelichting op de kerndoelen en recente vakdidactische literatuur inzake communicatief taalonderwijs en contextrijk wiskundeonderwijs. De vraag over het bestaan van een communicatieve behoefte in een leersituatie is een voorbeeld van een vraag voor Engels. Aandacht voor persoonlijke oplosstrategieën is bijvoorbeeld een typisch wiskundig gespreksonderwerp.

Als voorbereiding op het interview werd de handleiding van de gebruikte methode grondig bestudeerd. Voorafgaande aan het interview werd de les bezocht, na een grondige bestudering van de geplande leerstof en opdrachten. Bij de introductie van het interview bracht de interviewer kort verslag uit van de activiteiten die hij had waargenomen tijdens de les, met de vraag of deze een goede afspiegeling waren van wat in een doorsnee les gebeurt. Na deze introducerende vraag volgde de rest van het interview, waarin eventueel terugverwezen kon worden naar les passages.

4.3 Data-analyse

Bij de analyses op de docentgegevens zijn alleen gegevens van docenten betrokken die lesgeven in de basisvorming en die zowel in '94 als in '96 een vragenlijst ingevuld retour hebben gezonden ($n=86$). Om een beeld te verkrijgen van de realisatie van authentiek leren, wordt de individuele docent als analyse-eenheid gehanteerd. Veranderingen in de realisatie van authentiek leren in de periode 1993-1994 tot 1995-1996 zijn geanalyseerd met behulp van variantie-analyses voor een herhaalde metingen design. Het schooljaar (1993-1994 en 1995-1996) gold daarbij als binnen-subjecten factor. De school (A, B of C) is genomen als tussen-subjecten factor, waarmee verschillen in realisatie tussen scholen worden getoetst. Verschillen in veranderingen tussen scholen zijn getoetst door de interactieterm schooljaar* school in het variantie-analytisch model op te nemen.

De realisatie van authentiek leren bij Engels en wiskunde in de perceptie van leerlingen wordt in beeld gebracht met behulp van datasets uit 1993-1994 (een leerjaar) en 1995-1996 (drie leerjaren). Als analyse-eenheid is de leerling gehanteerd.

Verschillen in de wijze van lesgeven in leerjaar 1 tussen 1993-1994 en 1995-1996 kunnen alleen worden getoetst met een tussen subjecten design omdat in beide leerjaren sprake is van grotendeels verschillende docenten. Om die reden worden variantie-analyses gehanteerd met een tussen subjecten design, waarbij verschillen tussen twee schooljaren worden getoetst met behulp van de factor schooljaar. Overall verschillen tussen scholen in realisatie van authentiek leren worden getoetst door opname van de factor school (A, B, C). Variaties tussen scholen wat betreft de verschillen tussen 1993-1994 en 1995-1996 worden getoetst door opname van de interactieterm school*schooljaar.

Gebruikmakend van de dataset van 1995-1996 zijn ook de gegevens van leerjaar 2 en 3 meegenomen, teneinde een completer beeld te verkrijgen van verschillen tussen scholen. Dit is gedaan met behulp van variantieanalyses, waarbij naast de school ook het leerjaar (1, 2, of 3) als factor kon worden meegenomen.

Uitspraken van geïnterviewden zijn per school, vak en docent ingedikt tot de vier kenmerken van authentiek leren, onderverdeeld naar de aard van de instructie en begeleiding,

en de aard en keuze van leertaken. Daarbij is onderscheid gemaakt tussen de wijze waarop men de methode beoordeelt op aspecten van authentiek leren en de wijze waarop men de methode hanteert in de klas.

5 Resultaten

De resultaten hebben betrekking op drie onderdelen. Allereerst de realisatie van authentiek leren volgens de docenten, daarna de realisatie van authentiek leren volgens de leerlingen. De beschrijving van resultaten wordt afgesloten met een bespreking van condities en knelpunten inzake authentiek leren.

5.1 Realisatie van authentiek leren in de ogen van docenten

In dit deel wordt verslag gedaan over de mate waarin docenten naar eigen waarneming een didactiek van authentiek leren hanteren in de klas, aangevuld met gegevens over methodegebruik door een deel van de docenten. De bespreking vindt plaats aan de hand van de vier kenmerken van de didactiek van authentiek leren. Bij de bespreking van gemiddelde scores wordt de betekenis van de schaalpunten op de 6-puntsschaal aangehouden. Een score van 3.5, het schaal midden, wordt geïnterpreteerd als 'regelmatig'.

Tabel 1
Implementatie van kenmerken van authentiek leren (gemiddelden en tussen haakjes de standaarddeviaties)

Schaal	Totaal (n=86)		School A (n=12)		School B (n=32)		School C (n=42)	
	'94	'96	'94	'96	'94	'96	'94	'96
Constructie van kennis in complete taaksituaties (15 items, $\alpha = .78$)	3.3 (0.6)	3.4 (0.5)	3.6 (0.5)	3.4 (0.5)	3.5 (0.4)	3.4 (0.4)	3.2 (0.6)	3.3 (0.7)
Gerichtheid op de leefwereld (10 items, $\alpha = .81$)	3.6 (0.7)	3.6 (0.6)	3.8 (0.5)	3.7 (0.7)	3.7 (0.8)	3.7 (0.7)	3.5 (0.6)	3.6 (0.5)
Waarde van leeractiviteiten voor situaties buiten school (4 items, $\alpha = .68$)	2.9 (0.7)	3.0 (0.6)	2.9 (0.5)	2.8 (0.5)	3.2 (0.8)	3.3 (0.8)	2.7 (0.5)	2.8 (0.5)
Samenwerking en communicatie (7 items, $\alpha = .71$)	3.2 (0.8)	3.3 (0.7)	4.0 (0.6)	3.9 (0.8)	3.3 (0.8)	3.3 (0.7)	2.9 (0.8)	3.1 (0.7)

N.B. 1= nooit, 2= bijna nooit, 3= soms 4= vaak 5=zeer vaak, 6= altijd

1. Constructie van kennis

Instructiegedrag

In Tabel 1 worden de gegevens op de vier subschalen voor authentiek leren gepresenteerd. Uit de antwoorden van docenten blijkt dat in de leersituatie af en toe sprake is van constructie van kennis (gemiddeld 3.3 en 3.4 in '94 en '96). Docenten ondernemen naar eigen zeggen een aantal activiteiten ten behoeve van kennisconstructie. Deze hebben voornamelijk betrekking op het activeren van denkactiviteiten bij leerlingen: het geven van hints in plaats van complete antwoorden op vragen van leerlingen, gebruik maken van expertise van leerlingen, aandacht besteden aan de wijze waarop leerlingen tot een antwoord of oplossing van een opdracht zijn gekomen. Voorts geven docenten aan meer belang te hechten aan de manier waarop leerlingen tot een oplossing van een vraagstuk komen dan aan het correcte antwoord op zich. Echter, hardop laten verwoorden van de aanpak waarmee een probleem is opgelost, gebeurt slechts af en toe.

Leersituaties bevatten geen opdrachten waarbij inhouden van meerdere vakken aan de orde komen. Docenten laten zelden vakoverstijgende onderzoekjes uitvoeren. Er is niet duidelijk sprake van zelfstandige verzameling van informatie door leerlingen. Bij instructie, het uitvoeren van opdrachten en bij beoordelingsprocedures wordt geen gebruik gemaakt van integratieve opdrachten. Kortom de complexiteit van de taaksituatie is beperkt. Tussen 1994 en 1996 zijn geen significante verschuivingen opgetreden in onderwijsgedrag wat betreft constructie van kennis ($F=1.6$, $df=1$, 83 , $p=.69$). In school A zien we in 1994 frequenter onderwijsactiviteiten die duiden op constructie van kennis dan in school C en B, maar dit verschil neemt af in 1996 (interactie schooljaar* school: $F=2.9$, $df=2$, 83 , $p=.06$).

Methodegebruik

Uit de interviews met docenten van het eerste leerjaar blijkt dat de lessen Engels en wiskunde het karakter hebben van het aanleren van basis-kennis en -vaardigheden. Complete taken en leersituaties waarin leerlingen zeggenschap hebben over het lesverloop komen zelden voor in de leerboeken voor Engels en wiskunde. Docenten geven aan dat leerlingen, en dan

vooral de zwakkeren, het meeste profijt hebben van het aanleren van basiskennis, alvorens gewerkt kan worden met complexe taken. Een andere reden voor de beperkte aandacht voor kennisconstructie is tijdgebrek, die toe te schrijven is aan de overladenheid van het programma.

Leersituaties waarin compenserende strategieën moeten worden ingezet om een beperkte communicatieve vaardigheid te overbruggen komen dan ook beperkt voor. Wiskundeboeken bieden zeer beperkt de mogelijkheid aan leerlingen om eigen oplosstrategieën voor een probleem te ontwikkelen, aldus de docenten.

Docenten Engels geven er – in afwijking van de methode – de voorkeur aan eerst een aantal basale grammaticaregels aan te leren voordat gecommuniceerd wordt. Wiskundeleraars oriënteren zich vooral op de vakinhoud zelf. Wiskundige strategieën en begrippen worden slechts af en toe verbonden met andere schoolvakken.

2. Gerichtheid op de leefwereld

Instructiegedrag

De docenten onderschrijven gemiddeld genomen het belang om aan te sluiten bij de leefwereld van de leerlingen. De maatregelen hiertoe worden redelijk frequent getroffen (gemiddelde 3.6 zowel in 1994 als in 1996). Een nadere inspectie van de itemgemiddelden leert dat docenten aansluiten bij de leefwereld voorzover dat past in het strakke programma dat door het leerboek wordt voorgeschreven. Bij de uitleg van vakinhouden worden voorbeelden uit de ervaringswereld van de leerlingen betrokken. Een grote inbreng van leerlingen bij de lessen, stoelend op de leefwereld, wordt door docenten in mindere mate gerealiseerd, zoals het kiezen van (huiswerk)opdrachten die rechtstreeks betrekking hebben op de leersituatie en inspraak in het lesverloop.

Gemiddeld genomen verandert er weinig tijdens de implementatie van basisvorming als het gaat om het betrekken van de leefwereld bij het onderwijs. De gemiddelde scores van de docenten verschillen niet significant tussen '94 en '96 ($F=.31$, $df=1$, 83 , $p=.58$). De verschillen tussen de drie scholen zijn gering en blijven bovendien gelijk in de tijd (interactie schooljaar*school: $F=.87$, $df=2$, 83 , $p=.42$).

Methodegebruik

Methodes en de wijze waarop deze worden gebruikt weerspiegelen niet in alle opzichten een gerichtheid op de leefwereld. De leersituaties in wiskundeboeken worden door docenten als concreet voorstelbaar gezien, omdat veel illustraties uit het dagelijkse leven zijn opgenomen. Docenten maken gebruik van deze illustraties. Daarentegen gebruiken ze niet vaak concrete materialen in hun lessen. In enkele gevallen maken docenten Engels gebruik van Engels uit het dagelijkse leven, vooral door middel van het vertellen van anekdotes (zoals over de eigen vakantie in Engeland). Leerlingen hebben geen inbreng in de lesprocedure die verder gaat dan wat is gepland in het leerboek. Docenten vinden dit ook niet wenselijk vanwege het al overladen programma.

Vanuit de optiek van authentiek leren zouden leerlingen bij Engelse lessen het Engels dat ze tegenkomen regelmatig dienen te herkennen uit hun dagelijks leven (bijvoorbeeld van de televisie en de computer). Daarnaast zou lezen, spreken, schrijven op eigen initiatief mogelijk moeten zijn.

Methoden voor Engels geven weinig ruimte voor deze eigen initiatieven, aldus docenten. De geïnterviewde docenten wijken in dit opzicht ook niet af van hun methode. Als motief hiervoor geven docenten aan dat leerlingen eerst een aantal communicatieve basisvaardigheden moeten beheersen, voordat sprake kan zijn van vrijheid in keuze van materialen om te luisteren, lezen, spreken, en schrijven. De volgende bezwaren tegen het betrekken van de leefwereld worden genoemd: de (leef)wereld verandert snel, leefwereld verschilt van leerling tot leerling, meegebrachte materialen zijn van slechte kwaliteit, en leerlingen zouden hun leefwereld juist moeten ontstijgen.

3. Waarde van leeractiviteiten voor situaties buiten school

Instructiegedrag

Docenten besteden gemiddeld genomen soms aandacht aan de relevantie van het leren voor situaties buiten de school getuige de gemiddelde score van 3.0 op de betreffende subschaal. Gelet op de itemgemiddelden geldt ook hier dat hoe verder de maatregelen afwijken van de methode en hoe meer ze aansluiten bij buiten-

schoolse toepassing van kennis, hoe minder vaak ze worden toegepast. Zo is er gemiddeld genomen regelmatig aandacht voor actuele ontwikkelingen in de samenleving en wordt eveneens regelmatig aandacht geschonken aan de relevantie van leerstof voor beroeps- en leefsituaties. Daarentegen wordt slechts soms gebruik gemaakt van media als leermiddelen bij het onderwijsleerproces, waarin beroeps-situaties en actuele ontwikkelingen aan de orde zijn. De meest vergaande wijze om aandacht te besteden aan buitenschoolse toepassing is het geven van taken waarbij een maatschappelijk probleem aan de orde is, waarvan vervolgens het resultaat wordt voorgelegd aan een instantie of persoon buiten de school. Dit gebeurt bijna nooit.

Geconcludeerd kan worden dat de buitenschoolse omgeving soms doordringt in de les. De aandacht voor de betekenis van leren voor situaties buiten de school is dan ook beperkt te noemen. In dit opzicht is er voor alle scholen weinig veranderd tussen '94 en '96 ($F=2.0$, $df=1$, 83 , $p=.16$).

Methodegebruik

Gevraagd is naar het ingeschatte nut van leersituaties in methoden Engels en wiskunde voor buitenschoolse situaties. Bij het leren communiceren in het Engels zou sprake moeten zijn van het vervullen van een communicatieve behoefte, zoals het geven of verkrijgen van informatie. Daarnaast kan aandacht worden besteed aan taalgebruik in Engels sprekende landen, respectievelijk aan gebruikers van wiskunde buiten de school.

Docenten blijken aandacht te besteden aan de relevantie van leren voor buitenschoolse situaties, voorzover de methode dat toelaat. Op deze laatste beperking verschillen docenten Engels en wiskunde van elkaar. Docenten Engels zijn veel meer tevreden over het nut en de bruikbaarheid van leersituaties in de gebruikte methodes dan wiskundedocenten. Dagelijks Engels taalgebruik wordt volgens docenten goed in beeld gebracht, hoewel dit uitsluitend Brits Engels is. Bij opdrachten in het werkboek is echter zelden sprake van een reële communicatieve behoefte, zoals een behoefte aan informatie-uitwisseling tussen spreker en luisteraar, aldus docenten. Docenten veranderen in zo'n situatie de opdracht niet.

Docenten van school A nemen in tegenstelling tot hun collega's van B en C initiatieven om wiskundekennis bruikbaar te maken, waaronder een project waarbij schoolvakken worden geïntegreerd. Enkele wiskundedocenten van B en C twijfelen aan het belang van relaties leggen met de wereld van alledag. Zij zijn van mening dat gebruik van veel contexten verwarrend kan werken bij het aanleren van basisvaardigheden.

4. Communicatie en samenwerking

Instructiegedrag

Wat betreft het vierde kenmerk van authentiek leren, communicatie en samenwerking, zien we een beeld dat verschilt per school, dat constant blijft in de tijd (interactie schooljaar* school: $F=1.5$, $df=2$, 83, $p=.23$). Rapportage van een gemiddeld beeld is daarom weinig zinvol. Op school A wordt zowel in '94 als in '96 de lessituatie vaak gekenmerkt door samenwerking en communicatie (gem. resp. 4.0 en 3.9), terwijl dit op de andere scholen gemiddeld genomen minder vaak het geval is ($F=5.9$, $df=2$, 83, $p=.01$).

Op school A wordt vrij vaak samengewerkt in groepen, waarbij de docent zich als begeleider op de achtergrond houdt. De waardering van resultaten van een groepsproces is daar volgens de docenten vrij sterk een kwestie van overleg tussen docent en leerling. De controle-taak ligt bij groepswerk echter minder sterk bij de leerling dan verwacht zou mogen worden. Ook is opvallend dat bij groepswerk slechts heel soms onderlinge afhankelijkheid (interdependentie) wordt gecreëerd, zoals bij coöperatief leren het geval zou zijn. Interdependentie geldt in nog mindere mate voor de andere scholen. Ook treedt de docent daar wat meer op de voorgrond dan op school A.

Op school A worden leerlingen soms en vaker dan op school C in rollen geplaatst, herkenbaar uit het dagelijks leven, om met elkaar te leren communiceren ($F=4.7$, $df=2$, 83, $p=.01$). Leerlingen krijgen regelmatig de opdracht om resultaten van opdrachten aan elkaar te presenteren.

Methodegebruik

Samenwerkend leren als kenmerk van authentiek leren betekent dat leerlingen in leersitu-

aties onderling afhankelijk zijn van elkaar om leerdoelen te bereiken. Bij Engels kan tot op zekere hoogte gesproken worden van wederzijdse afhankelijkheid. Leerlingen oefenen Engelse uitdrukkingen en korte conversaties in tweetallen, zoals die gepland staan in het leer- en werkboek. Volgens de wiskundedocenten bieden hun methoden weinig gelegenheid aan studenten om samen te werken. Zowel bij Engels als bij wiskunde is het volgens docenten *toegestaan* dat leerlingen overleggen met klasgenoten over individuele (verwerkings)-opdrachten. In dat opzicht is er sprake van communicatie over kennis.

Docenten van school A stimuleren hun leerlingen om elkaar te helpen, vooral wanneer zichzelf niet beschikbaar zijn voor ondersteuning. Docenten doen dit door groepjes te vormen, waarbij regels gelden voor het geven en ontvangen van hulp. In enkele gevallen veranderen deze docenten opdrachten uit de methode op zo'n manier dat leerlingen afhankelijk worden van elkaar. Docenten van school B en C rapporteren geen van deze activiteiten. Integendeel, soms wordt een wederzijdse uitwisseling van informatie vervangen door het om beurten hardop aan elkaar voorlezen van zinnen die tevoren zijn opgeschreven. Docenten geven hierbij als motief aan dat leerlingen moeten beschikken over een aantal standaard uitdrukkingen die ze in de toekomst kunnen gebruiken, hetgeen ze in boven beschreven situaties als belangrijker zien dan de informatie-uitwisseling.

5.2 Leerlingpercepties van authentiek leren

In Tabel 2 zijn de resultaten op de vragenlijsten voor leerlingen weergegeven, die een beeld verschaffen van leerlingpercepties inzake de realisatie van authentiek leren. Het betreft gemiddelden op de schaal 'realisatie van de didactiek van authentiek leren' voor het vak Engels (22 items $\alpha=.78$) en voor wiskunde (20 items, $\alpha=.77$). De schalen zijn niet volledig parallel, aangezien het gaat om vakspecifieke uitwerkingen.

We kunnen in het algemeen constateren dat de leerlingen het onderwijs in Engels niet bestempelen als gericht op authentiek leren, althans volgens onze definitie (gemiddelden 2.2 in 1994 en 2.1 in 1996).

Tabel 2

Mate waarin volgens leerlingen sprake is van authentiek leren bij Engels en wiskunde

	Totaal		School A		School B		School C	
	Gem	Sd	Gem	Sd	Gem	Sd	Gem	Sd
Engels (22 items, $\alpha=.78$)								
'93-'94:								
Leerjaar 1	2.2	.4	2.4	.4	2.0	.4	2.0	.4
'95-'96:								
Leerjaar 1	2.1	.4	2.3	.4	2.1	.4	2.0	.4
Leerjaar 2	2.2	.4	2.3	.4	2.3	.3	2.0	.4
Leerjaar 3	2.1	.5	2.4	.5	2.0	.4	2.0	.4
Totaal '95-'96	2.1	.4	2.3	.4	2.1	.4	2.0	.4
Wiskunde (20 items, $\alpha=.77$)								
'93-'94:								
Leerjaar 1	2.3	.4	2.5	.4	2.3	.3	2.2	.3
'95-'96:								
Leerjaar 1	2.4	.4	2.6	.5	2.3	.4	2.3	.4
Leerjaar 2	2.4	.4	2.6	.4	2.3	.4	2.3	.4
Leerjaar 3	2.1	.4	2.2	.5	2.2	.3	2.1	.4
Totaal '95-'96	2.3	.4	2.5	.5	2.3	.4	2.2	.4

N.B. 1= nooit, 2= heel af en toe, 3= af en toe, 4= vaak, 5= heel vaak

Meer specifiek geldt het volgende beeld. Docenten Engels betrekken volgens de leerlingen bij hun lessen zelden onderwerpen van andere vakken. Van leerlingen wordt zelden verwacht dat ze op eigen initiatief activiteiten starten, teneinde zelfstandig leren te bevorderen. Complexe opdrachten waarin sprake is van een combinatie van lezen en schrijven komen heel af en toe voor. Belevissen en eigen interesses van leerlingen spelen nauwelijks mee tijdens de lessen, aldus de leerlingen. Ook valt op dat leerlingen zelden of nooit voorwerpen van thuis mogen meenemen, die bij de les betrokken worden (bijv. cd's). Daarentegen geven leerlingen aan dat opdrachten soms gaan over onderwerpen die ze herkennen uit de wereld buiten de school. Wat betreft de waarde van leren voor buitenschoolse situaties geven leerlingen aan dat ze bijna nooit opdrachten uitvoeren voor een publiek buiten school. Tot slot werken leerlingen niet regelmatig aan gezamenlijke leertaken.

Wat betreft wiskunde kunnen we eveneens constateren dat volgens leerlingen uit leerjaar 1 tot en met 3 docenten zelden onderwijs geven volgens de kenmerken van authentiek leren (gemiddeld 2.3 zowel in 1994 als in 1996).

Door vergelijking van gegevens uit 1994 en 1996 voor leerjaar 1 kan een uitspraak gedaan worden over veranderingen in lesgedrag. De aard van de lessen Engels in leerjaar 1, zoals

gezien door de leerlingen, is niet veranderd tijdens de invoering van de basisvorming ($F=0$; $df=1$, 792, $p=.99$). Bij wiskunde zien we een kleine positieve, maar significante verschuiving die verschilt tussen scholen ($F=4.1$, $df=2$, 726, $p<.05$).

Een andere vergelijking betreft die tussen de drie scholen en de drie leerjaren. Dit levert het volgende beeld op voor wiskunde. In 1993-1994 geven leerlingen uit school A (gem. 2.5) wat betreft wiskunde in leerjaar 1 significant hogere scores aan hun docenten dan hun leeftijdgenoten in school B en C (gem. 2.3 en 2.2; $F=50.3$, $df=2$, 439, $p=.00$). Uit de gegevens van 1995-1996¹ blijkt dat school A (gem. 2.6) in leerjaar 1 en 2 opnieuw hoger scoort dan de twee andere scholen (gem. 2.3), maar dat dit verschil niet bestaat in leerjaar 3 (resp. 2.2 tegen 2.2 en 2.1). Er is met andere woorden een significante interactie tussen school en leerjaar ($F=4.2$, $df=4$, 921, $p=.00$). Overigens geldt meer algemeen dat in leerjaar 3 (gem. 2.1) wiskundeleraars (volgens hun leerlingen) in mindere mate authentiek leren realiseren dan in leerjaar 1 en 2 (beide gem. 2.4). Het hoofdeffect van leerjaar is significant ($F=35.4$, $df=2$, 921, $p=.00$).

Het beeld bij Engels is in zoverre vergelijkbaar met dat van wiskunde dat ook hier school A hoger scoort dan de andere scholen, maar dan in beide onderzoeksjaren en in alle leerja-

ren (1994: $F=41.9$, $df=2$, 463, $p=.00$; 1996: $F=60.31$ $df=2$, 1011, $p=.00$) . De mate van realisatie van authentiek leren varieert met name op school B tussen leerjaren, waar deze op andere scholen constant blijft (interactie schooljaar*school: $F=4.4$, $df=4$, 1005, $p=.00$).

5.3 Conditie voor authentiek leren

De condities ten behoeve van 'constructie van kennis' zijn in zeer beperkte mate gerealiseerd (zie Tabel 3). Zo geldt dat de gebruikte methode niet veel ruimte biedt om opdrachten aan leerlingen te verstrekken met maximale vrijheid. De lestijd om opdrachten te geven met meerdere goede antwoorden wordt niet toereikend geacht. Schoolfaciliteiten om leerlingen zelfstandig te laten werken zijn volgens docenten in beperkte mate voorhanden.

De gehanteerde methode biedt wel mogelijkheden om onderwerpen uit de leefwereld te kiezen. Mogelijkheden tot het inzetten van media worden als beperkt gezien. Bovendien zijn deze verder verslechterd van 1994 tot 1996, aldus docenten. ($t=2.8$, $df=84$, $p=.00$).

Realisatie van authentiek leren vraagt om ondersteunende maatregelen op schoolniveau, zoals het gebruik van flexibele roostering en samenwerking tussen vaksecties. Docenten geven aan dat gemiddeld genomen weinig sprake is van ondersteuning op schoolniveau. Deze situatie blijft onveranderd tussen 1994 en 1996 ($F=.02$, $df=1$, 81, $p=.89$).

Tenslotte zijn de docenten het vrij sterk eens met de stelling dat leerinhouden in eerste instantie afhankelijk zijn van de gehanteerde methode. In dit opzicht is hun mening niet veranderd tussen '94 en '96 ($F=2.9$, $df=1$, 82, $p=.59$). Voorts blijkt dat condities wat betreft leermiddelen zowel op A als op B in als veel gunstiger worden ingeschat dan op school C, hoewel deze situatie in 1996 op beide scholen verslechterd is (interactie schooljaar*school: $F=5.4$, $df=2$, 82, $p=.01$).

6 Samenvatting en discussie

Uit de gegevens van drie bronnen (docentpercepties, leerlingpercepties en methodegebruik)

Tabel 3

Perceptie van de condities en problemen omtrent de toepassing van de didactiek van authentiek leren (gemiddelden en standaarddeviaties)

Item	Totaal (n=86)		School A (n=12)		School B (n=32)		School C (n=42)	
	'94	'96	'94	'96	'94	'96	'94	'96
- De lestijd voldoende om opdrachten te geven met meerdere goede antwoorden	2.6 (1.4)	2.7 (1.4)	3.2 (1.3)	3.0 (1.5)	2.4 (1.3)	2.9 (1.5)	2.5 (1.5)	2.5 (1.2)
- Schoolorganisatorische mogelijkheden voor geïntegreerd onderwijs zijn in ruime mate aanwezig (aanpassing lesrooster, overleg tussen vaksecties)	2.9 (1.3)	2.8 (1.3)	3.3 (1.4)	3.1 (1.7)	2.7 (1.2)	2.7 (1.1)	2.9 (1.4)	2.9 (1.2)
- De mogelijkheden tot inzetten van media (krant, radio, t.v.) zijn op school ruim voorhanden	3.0 (1.5)	2.6 (1.4)	3.3 (1.3)	2.5 (1.3)	4.3 (1.0)	3.6 (1.3)	1.8 (0.9)	1.9 (1.0)
- Voor vinden van opdrachten met maximale vrijheid leerlingen biedt methode ruimte	3.0 (1.2)	3.1 (1.2)	3.1 (1.1)	2.8 (1.1)	3.3 (1.2)	3.5 (1.1)	2.8 (1.2)	2.8 (1.3)
- Basisvoorzieningen in de school voldoende, zodat leerlingen zelfstandig informatie kunnen verzamelen	3.7 (1.5)	3.4 (1.3)	4.3 (1.4)	3.2 (1.3)	4.1 (1.6)	4.0 (1.2)	3.2 (1.4)	3.1 (1.4)
- De gehanteerde methode nodigt uit tot kiezen van onderwerpen uit de leefwereld	4.1 (1.4)	4.0 (1.1)	4.5 (1.0)	4.0 (1.0)	4.5 (1.2)	4.3 (1.1)	3.6 (1.5)	3.7 (1.1)
- Keuze en volgorde van leerinhouden zijn in eerste instantie afhankelijk van de methode	4.4 (1.4)	4.5 (1.4)	4.9 (1.0)	4.7 (1.0)	4.4 (1.5)	4.4 (1.4)	4.3 (1.5)	4.6 (1.4)

Note 1= geheel mee oneens, 2= grotendeels mee oneens, 3= enigszins mee oneens, 4= enigszins mee eens, 5= grotendeels mee eens, 6= geheel mee eens

mogen we concluderen dat kenmerken van authentiek leren binnen de basisvorming in vrij beperkte mate zijn gerealiseerd op de drie onderzochte scholen en dat in deze situatie weinig is veranderd tussen 1994 en 1996 (vraagstelling 1). Meer in het bijzonder kunnen we de volgende conclusies trekken:

1. Docenten besteden aandacht aan constructie van kennis, voorzover het gaat om de weg waarlangs c.q. de strategie waarmee leerlingen leertaken kunnen vervullen. Complete taken, zoals integratieve opdrachten komen echter weinig voor (bijvoorbeeld eigen onderzoek door leerlingen, verslaggeving van bevindingen en integratie van vakinhouden).

2. Om aan te sluiten bij de leefwereld nemen docenten maatregelen die passen in het strakke programma dat door het leerboek wordt voorgeschreven, zoals het bespreken van voorbeelden uit het dagelijkse leven. Van activiteiten buiten het leerboek om en startend bij individuele interesses en behoeften van leerlingen is in beperkte mate sprake.

3. Voor maatregelen ten behoeve van de relevantie van schoolse leersituaties voor buitenschoolse situaties geldt: hoe verder de maatregelen de gehanteerde methode overstijgen (contacten met beroepsbeoefenaren, authentieke media), des te minder ze worden toegepast.

4. Aspecten van samenwerking en communicatie, waaronder groepswork met het gebruik van taken met onderlinge doelafhankelijkheid, het elkaar laten presenteren van resultaten, komen, school A uitgezonderd, beperkt voor.

5. Leerlingen geven een minder positief oordeel over hun docenten met betrekking tot de realisatie van authentiek leren dan de docenten zelf. Hierbij geldt de beperking dat de leerling-schalen op een andere wijze zijn geconstrueerd dan de docentenlijsten.² De hogere waardering die docenten zichzelf toekennen vergeleken met leerlingen kan een effect van sociale wenselijkheid zijn. Zelfpercepties zijn gevoelig voor dit fenomeen, hetgeen in de bijdrage van Roelofs en Houtveen elders in dit nummer aan de orde komt.

6. Docenten voeren als argument voor hun huidige lespraktijken aan dat eerst basiskennis dient te worden aangebracht alvorens gewerkt kan worden volgens kenmerken van authentiek leren.

Deze bevindingen sporen met gegevens van Newmann, Marks en Gamoran (1996), die aangeven dat op een zeer kleine minderheid van scholen kenmerken van authentiek leren ('authentic pedagogy') zijn waargenomen. In dit opzicht lopen de ideeën en ontwikkelingen van groepen onderzoekers, beleidsmakers en ontwikkelaars nog ver voor op de praktijk. De conclusies liggen op een lijn met die van Kuiper (1993), Terwel, Vermeulen en Volman (1996) en Withagen, Oud-de Glas, Smeets en Buis (1996) en inspectieonderzoek (Inspectie van het Onderwijs, 1997), waaruit bleek dat leefwereldgericht onderwijs, gebruik van activerende en divergente opdrachten tijdens lessen in beperkte mate worden gerealiseerd.

Het gegeven dat authentiek leren in de basisvorming, althans in de onderzochte scholen, tot nu toe nog onvolledig wordt gerealiseerd, is wel enigszins begrijpelijk. Condities voor authentiek leren in de schoolorganisatie (vraagstelling 2) zijn beperkt of zwak ingevuld volgens de docenten. Te denken valt aan onderlinge afstemming tussen verschillende vaksecties, het voeren van een ondersteunend beleid op schoolniveau, roostertechnische aanpassingen voor vak overstijgende activiteiten.

Ook de aard van de leermiddelen en de methode vormt een obstakel. Docenten geven aan dat de methodes – die zij overigens vrij sterk volgen – authentiek leren slechts in beperkte mate mogelijk maken. Nieuwe media en computerondersteunde leeromgevingen zouden perspectieven kunnen bieden als het gaat om interactie, communicatie, authenticiteit van gebruikte leercontexten, zoals aange-toond in het Amerikaanse onderwijsproject 'Fostering Communities of Learners' (Campion, Shapiro & Brown, 1995). Op de onderzochte scholen zijn deze media onvoldoende voorhanden. Dit geldt ook meer algemeen in Nederland (Kleis Jager, 1996; Bronkhorst & Rouwen, 1996).

Ten derde is het de vraag in hoeverre docenten beschikken over de benodigde vaardigheden om een didactiek van authentiek leren vorm te geven.

Tot slot is het mogelijk dat docenten weerstanden hebben tegen een tamelijk fundamentele vernieuwing als het realiseren van TVSkarakteristieken, respectievelijk authentiek

leren in de lessen, die inhouden dat beproefde praktijken moeten worden losgelaten. Uit de interviews werd duidelijk dat docenten de voorkeur geven aan het aanleren van basiskennis en vaardigheden, boven leersituaties met veel vrijheid en eigen initiatief voor leerlingen.

Hiermee zijn we tegelijkertijd aangekomen bij de grenzen van authentiek leren en daarmee aan de haalbaarheid van de TVS-karakteristieken in de basisvorming. Onderwijs in 15 vakken met elk een eigen programma staat vormen van authentiek leren, met aandacht voor vaardigheden en samenhang in de weg. Sterk in vakken gesplitst onderwijs vraagt om beheersbaarheid en efficiency. De huidige wijze waarop scholen zijn georganiseerd, is daarop afgestemd: aparte leslokalen per vak, korte lessen, gevolgd door huiswerk, dat in de volgende lessen wordt gecontroleerd via overhoringen, beperkte omvang van tijdrovende contacten met de buitenwereld.

Ook valt te betwijfelen of tijdens het onderwijsleerproces altijd voor alle leerlingen alle kenmerken van authentiek leren aanwezig moeten zijn. Er zijn, zoals docenten in dit onderzoek ook zelf aangaven, argumenten om niet te kiezen voor een leeromgeving die een sterk beroep doet op zelfontdekkend en zelfgestuurd leren. Er zijn leersituaties waarin leerlingen niet beschikken over voldoende voorkennis en regulatievaardigheden, waarbij het nodig kan zijn het proces van kennisconstructie deels docentgestuurd te laten verlopen. Wanneer deze behoeften van leerlingen zouden worden genegeerd, kunnen problemen ontstaan (Vermunt, 1992; Simons, 1995). Het lijkt dan ook zinvol in voortgezet onderwijs op flexibele wijze invulling te geven aan de kenmerken van authentiek leren, zonder dat constructiviteit, betekenisvolheid, en interactiviteit van het leerproces verloren gaan.

Afsluitend moet gewezen worden op enkele methodologische problemen die de validiteit en betrouwbaarheid van het onderzoek kunnen schaden. Om te beginnen gaat het hier om slechts drie scholen.

Gezien de lage respons op de enquête in school A moeten we vraagtekens zetten bij de representativiteit van het daar verkregen beeld. Gezien het kleine aantal scholen is generali-

seerbaarheid naar de Nederlandse populatie bovendien niet gegarandeerd. Wel worden onze gegevens bevestigd door eerder gerapporteerd onderzoek (Newmann et al, 1996; Kuiper, 1993; Terwel, Vermeulen & Volman, 1996; Withagen et al., 1996) en een andere door ons verrichte studie (zie daarvoor Roelofs en Houtveen elders in dit nummer).

De bevinding dat leerlingpercepties en docentpercepties ten aanzien van realisatie van authentiek leren verschillen, vraagt om vervolgonderzoek waarin via systematische observaties nadere informatie wordt ingewonnen over feitelijk lesgedrag.

Noten

1. We benadrukken nog eens dat dit resultaten van andere docenten betreft.
2. De vakspecifieke uitwerking voor Engels en wiskunde zien we wel terug in een te verwachten samenhang tussen de docentgemiddelden op de leerlingsschalen voor authentiek leren en de subschalen van de docentenlijst. Er bestaat een positieve correlatie tussen de docentgemiddelden voor authentiek leren op de leerlingenlijst wiskunde en de subschaal 'constructie van kennis' van de docentenlijst (.43, $p < .05$; $n = 31$). Voorts bestaat er een significante correlatie (.42; $p < .05$, $n = 27$) tussen de docentgemiddelden op de schaal authentiek leren Engels en de docentscores op de subschaal 'communicatie en samenwerking'.

Literatuur

- Bronkhorst, J., & Rouwen, M. (1996). De noodzaak van telektiek: een didactiek van telematica. *Didaktief*, 26 (2): 10-11.
- Brown, A.L., & Campione, J.C. (1994). Guided discovery in a community of learners. In: K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229-270). Cambridge, MA: MIT Press/ Bradford Books.
- Campione, J.C., Shapiro, A.M., & Brown, A.L. (1995). Forms of transfer in a community of learners: Flexible learning and understanding. In A. McKeough, J. Lupart & A. Marini, *Teaching for transfer. Fostering generalisation in learning*. New Jersey: Lawrence Erlbaum Associates.

- Els van, T.J.M. & Buis, Th.J.M.N. (1987). *De praktijk van het onderwijs moderne vreemde talen in de bovenbouw havo/vwo*. Enschede: SLO.
- Essen, A. van (1990). Three decades of foreign-language teaching in the Netherlands. *Toegepaste Taalwetenschap in Artikelen (applied Linguistics in Articles)*, 36, 72-85. ERIC reproduction services, no.338021.
- Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht: Reidel.
- Freudenthal, H. (1991). *Revisiting Mathematics Education*. China Lectures. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Gravemeijer, K. (1995). *Developing realistic mathematics education*. Culemborg: Technipress.
- Hove, J. ten, & Zwaard, P. van der (1992). *Wiskunde. Handboek Basisvorming, voor de praktijk van het voortgezet onderwijs*. afl 21, hoofdstuk III, 6b. Houten: Bohn Stafleu Van Loghum.
- Inspectie van het Onderwijs (1997). *De basis gelegd. Evaluatie van schoolorganisatorische aspecten van de basisvorming*. Utrecht: Inspectie van het Onderwijs.
- Kleis Jager, P. (1996). CD-i en Cd-rom blijven rare buitenbeentjes in onderwijsland. *Didaktief*, 26 (2), 8-9.
- Kooij, H. van der (1995). *Realistic mathematics education and extended response tasks*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco 1995.
- Kuiper, W. (1993) *Curriculumvernieuwing en lespraktijk* (academisch proefschrift). Enschede: Universiteit Twente.
- Lamon, M. (1995). *Schools for thought: Transforming classrooms into learning communities*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco 1995.
- Lagerweij, N.A.J. (1993). Implementatie van basisvorming: de start van een dieptestudie. *Pedagogische Studiën*, 70(3), 162-180.
- Lagerweij, N.A.J., Kanselaar, G., Linden, J.L. van der, Roelofs, E.C., Treep, O., Voogt, J.C., Vriens, L.J.A., & Wessum, L. van (1996). *Basisvorming op de voet gevolgd, rapport 2. De invoering. Informatie over Utrechts onderzoek naar invoering basisvorming*. Utrecht: vakgroep Onderwijskunde.
- Lenaerts, G. (1992). Engels. *Handboek Basisvorming, voor de praktijk van het voortgezet onderwijs*. afl 21, hoofdstuk III, 4b. Houten: Bohn Stafleu Van Loghum.
- Newmann, F.M., Marks, H.M. & Gamoran, A.G. (1996). Authentic pedagogy and student performance. *American Journal of Education*, 104, 280-312.
- Roelofs, E., Franssen, H. & Grootsholten, J. (1996). *Basisvorming op de voet gevolgd. Deelrapport in Utrechts onderzoek naar de invoering van de basisvorming. Toepassingsgerichtheid en authentiek leren na één jaar basisvorming*. Utrecht: ISOR/Vakgroep Onderwijskunde.
- Roelofs, E.C. & Terwel, J. (1999). Constructivism and authentic pedagogy: State of the art and recent developments in the Dutch national curriculum in secondary education. *Journal of Curriculum Studies*, 31 (2), 201-227.
- Scardamalia, M., Bereiter, C., McLean, R.S. & Woodruff, E. (1989). Computer supported intentional learning environments. *Journal of Educational Computing Research*, 5, 51-68.
- Terwel, J., Vermeulen, A. & Volman, M. (1996) *Success Factors in Curriculum Innovation. The Case of Mathematics, Physics, Chemistry and Biology in Secondary Mathematics*. Paper presented at the Annual Meeting of the AERA, New York, 1996.
- Simons, P.R.J. (1995). Leerlingenmerken. In: N. Verloop, & J. Lowyck, *Onderwijskunde. Een kennisbasis voor professionals*. Groningen: Wolters-Noordhoff
- Vanderbilt cognition and technology group (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher*, 19 (6), 2-9.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs: Naar procesgerichte instructie in zelfstandig denken*. Amsterdam: Swets & Zeitlinger.
- Voogt, J.C., & Franssen, H. (red) (1994). *De basisvorming op de voet gevolgd. Rapport 1: Voorbereidingen*. Utrecht: Vakgroep Onderwijskunde.
- Withagen, V., Oud-de Glas, M., Smeets, E. & Buis, Th.J.M.N. (1996). *Vernieuwingen in het vreemde talenonderwijs*. Nijmegen: ITS.

Auteurs

E. Roelofs is universitair docent-onderzoeker aan de capaciteitsgroep Onderwijskunde van de Universiteit Utrecht

H. Franssen was tot maart 1995 universitair hoofd-docent aan de capaciteitsgroep Onderwijskunde van de Universiteit Utrecht

Th. Houtveen is senior onderzoeker bij het ISOR, Instituut voor Onderwijsresearch, Universiteit Utrecht.

N. Lagerweij is hoogleraar aan de capaciteitsgroep Onderwijskunde van de Universiteit Utrecht

Adres: Capaciteitsgroep Onderwijskunde, Postbus 80.140, 3508 TC Utrecht

Abstract

An in-depth study into authentic pedagogy: teacher behaviour, text book use and students' perceptions

E. Roelofs, H. Franssen, Th. Houtveen en N. Lagerweij. Pedagogische Studiën, 1999, 76, 258-272.

The national curriculum for the first stage of Dutch secondary education introduced in the school year 1993-1994 implies, apart from a change in educational content, a change in teaching practices. The research question addressed concerns the degree to

which teachers realize characteristics of authentic pedagogy, a concept covering the changes aimed for.

An in-depth case study in three large secondary schools was set up, to assess the implementation of four characteristics of authentic pedagogy: 1) construction of knowledge in complete task environments, 2) connection to students' personal worlds, 3) value of learning beyond school, 4) cooperation and communication.

Different data sources were used: in 1994 and 1996 teacher questionnaires have been administered. In addition students have been questioned about teaching practices in their math and English classes. At the start of 1995 English and math teachers have been interviewed on their text book use.

The results show that none of the schools scores high on either one of the characteristics. The text book used appears not to support authentic pedagogy. Teachers do not adapt their text books in this respect. Finally, conditions for authentic pedagogy on class and school level are hardly being met.

Bijlage 1

Aantal klassen en leerlingen (tussen haakjes) per school waaraan de vragenlijst voor Engels en wiskunde is voorgelegd

Vak jaar	leerjaar	Totaal	School A	School B	School C
Engels					
1994:	1	20 (n= 464)	8 (n= 163)	5 (n= 113)	7 (n=188)
1996:	1	17 (n= 340)	6 (n= 101)	4 (n= 75)	7 (n= 164)
1996:	2	15 (n= 306)	5 (n= 98)	4 (n= 88)	6 (n= 120)
1996:	3	19 (n= 390)	8 (n= 127)	5 (n= 98)	6 (n= 165)
Wiskunde					
1994:	1	19 (n= 442)	9 (n= 162)	4 (n= 93)	7 (n= 187)
1996:	1	14 (n= 307)	4 (n= 80)	4 (n= 89)	6 (n= 138)
1996:	2	17 (n= 360)	7 (n= 118)	4 (n=108)	6 (n= 134)
1996:	3	17 (n= 289)	5 (n= 56)	5 (n= 94)	7 (n= 139)

Onderwijsopbrengsten en onderwijskansen voor en na invoering van de basisvorming

G. van der Werf, M. Lubbers en H. Kuyper

Samenvatting

Met behulp van de gegevens uit de bestanden van de cohortonderzoeken VOCL '89 en VOCL '93 is onderzocht in welke mate sedert de invoering van de Basisvorming de opbrengsten van het voortgezet onderwijs zijn verbeterd en de onderwijskansen van achterstandsleerlingen zijn toegenomen. Overall blijkt dat de prestaties voor wiskunde en tekstbegrip Nederlands in het derde jaar en de bereikte onderwijspositie (score op de leerjarenladder) na vijf jaar vrijwel gelijk zijn gebleven. Uitsplitsing naar onderwijstype laat zien dat er sprake is van een lichte verbetering voor (I)VBO en MAVO leerlingen en een kleine achteruitgang voor HAVO en VWO leerlingen. Na controle voor verschillen in instroomniveau (intelligentie, score op de entree-toets en advies van de basisschool) tussen beide cohorten blijkt dat de prestaties zijn gedaald, terwijl de onderwijspositie is gestegen. De daling respectievelijk stijging is sterker voor de leerlingen in de hogere schooltypen dan voor leerlingen in de lagere.

De mate van groepsgebonden ongelijkheid van onderwijskansen is sinds de invoering van de Basisvorming niet afgenomen. De verschillen in prestaties en behaalde onderwijspositie, samenhangend met sekse, SES, en etniciteit, zijn na controle voor instroomniveau in beide cohorten even groot.

Geconcludeerd wordt dat de Basisvorming nog niet tot de gewenste effecten heeft geleid, althans niet in termen van de in het onderzoek gehanteerde criteria, maar dat er ook geen reden is tot pessimisme over de gevolgen van de Basisvorming voor de opbrengsten van het voortgezet onderwijs.

Inleiding

In augustus 1993 is elke school voor voortgezet onderwijs begonnen met de invoering van de

Basisvorming. In de Basisvorming wordt een gemeenschappelijk curriculum aangeboden aan alle leerlingen in de eerste fase van het voortgezet onderwijs, die in het algemeen uit drie leerjaren bestaat. In sommige scholen kan de Basisvorming twee of vier leerjaren omvatten. De invoering van de Basisvorming is gebaseerd op drie sociaal-politieke doelen, namelijk:

1. Verbetering van de algemene kwaliteit van het onderwijs;
2. Versterking en verbreding van de gemeenschappelijke culturele basis, aangepast aan technologische en sociaal-culturele ontwikkelingen;
3. Uitstel van studie- en beroepskeuze.

Het gemeenschappelijk curriculum omvat algemene en vakspecifieke kerndoelen die richtinggevend zijn voor de leerinhouden van de vakken. Elke leerling volgt minstens 15 vakken. De meeste van deze vakken werden ook al in het traditionele curriculum van de eerste fase van het voortgezet onderwijs aangeboden, maar er zijn enkele nieuwe vakken bijgekomen, namelijk techniek, informatica en verzorging. Behalve de 15 vakken, waaraan gemiddeld per jaar minimaal 25 lesuren van 50 minuten moeten worden besteed, beschikken scholen over 7 vrije lesuren die ze mogen besteden aan vakken van hun eigen keuze.

Scholen kunnen uit drie curriculumvarianten kiezen: 1) drie jaar Basisvorming gevolgd door overgang naar het vierde leerjaar MAVO, HAVO of VWO; 2) twee jaar Basisvorming en daarna VBO; of twee jaar Basisvorming en een jaar een combinatie van Basisvorming en VBO.

Door het CITO zijn toetsen ontwikkeld om te bepalen of leerlingen de eindtermen van de Basisvorming hebben behaald. Deze toetsen kunnen echter niet beschouwd worden als afsluitende examens: scholen zijn vrij om te beslissen wanneer ze de toetsen willen afnemen en er zijn geen strikte normen en niveaue-

aanduidingen vastgelegd om te bepalen of een leerling de toets met een voldoende resultaat heeft gemaakt.

Scholen moeten de Basisvorming verbinden met de onderwijstypen die ze daarna aanbieden. Ze moeten keuzen maken over de organisatie van de eerste fase van het voortgezet onderwijs ten aanzien van het groeperen van leerlingen in homogene of heterogene klassen, over de duur van de Basisvorming, over lesroosters en over het inbedden van de Basisvorming in de bestaande onderwijstypen. Deze verschillen kunnen grote gevolgen hebben voor de prestaties en de schoolloopbaan van de leerlingen.

Over de verschillen in implementatie wordt elders in dit themanummer gerapporteerd door Roelofs en Houtveen. Dit artikel gaat over de effecten van de Basisvorming op de prestaties en schoolloopbanen van leerlingen. De volgende vragen zullen worden beantwoord:

1. In welke mate zijn de onderwijsopbrengsten in het voortgezet onderwijs verbeterd sinds de invoering van de Basisvorming in 1993?
2. In welke mate zijn de kansen in het voortgezet onderwijs van leerlingen uit de lagere sociale milieus, allochtone leerlingen en meisjes verbeterd sinds 1993?

De vragen zullen worden beantwoord op basis van de gegevens die zijn verzameld in de cohortonderzoeken VOCL '89 en VOCL '93. Deze cohortonderzoeken zijn onder andere opgezet met als doel informatie te leveren ten behoeve van de evaluatie van de Basisvorming. Bij de beantwoording van de vraagstellingen dient rekening te worden gehouden met het feit dat de effecten worden vastgesteld bij de eerste lichting van leerlingen die Basisvorming hebben gehad. De scholen hadden op het moment dat deze leerlingen in de eerste klas zaten nog niet veel tijd gehad om de Basisvorming te implementeren. Kortom, het gaat hier om een tussenbalans.

Theoretische achtergrond

In de vraagstellingen staan twee begrippen centraal die nader dienen te worden gedefinieerd, namelijk het begrip *onderwijsopbrengsten* en het begrip *onderwijskansen*. Beide begrippen worden hieronder besproken.

Het begrip onderwijsopbrengsten heeft drie dimensies, namelijk de dimensies *effectiviteit*, *individueel rendement* en *efficiëntie*. In de eerste plaats moeten de opbrengsten van het onderwijs blijken uit wat de leerlingen op school leren in termen van meetbare kennis en vaardigheden (De Groot, 1983) die met het curriculum worden beoogd. Het gaat hier dus om de mate van doelrealisering ofwel de *effectiviteit* van het onderwijs. Voor zover verhoging van het peil van het onderwijs het doel is van de Basisvorming dan betekent dit logischerwijs dat de Basisvorming beoogt dat leerlingen meer moeten opsteken van het onderwijs dan ze voorheen deden en dus betere resultaten moeten behalen dan voorheen, uiteraard bepaald met dezelfde criteria. Verhoging van het peil betekent uitdrukkelijk niet verhoging van het rendement in termen van behaalde diploma's, slaag- en zak percentages van scholen en dergelijke. Verhoging van het rendement is wezenlijk iets anders dan verhoging van het peil van het onderwijs. We komen daar straks nog op terug.

Voor het vaststellen van het peil van het voortgezet onderwijs na invoering van de Basisvorming in vergelijking met het peil voorheen dient een vergelijking te worden gemaakt tussen de leerprestaties van leerlingen die Basisvorming hebben gehad en leerlingen die het voormalige curriculum hebben gehad. Hiertegen valt in te brengen – en velen zullen dat ook doen – dat met de Basisvorming heel andere doelen worden beoogd dan met het oude curriculum. Zo wordt bijvoorbeeld in de uitwerking van de kerndoelen van de basisvorming minder accent gelegd op het opdoen van kennis en meer op het leren toepassen van kennis en het leren van vaardigheden (Procesmanagement Basisvorming, 1993). Er is dan echter geen sprake van verhoging van het peil van het onderwijs als doel van de Basisvorming maar van een wijziging in doelen en inhouden. Een dergelijke opvatting maakt een vergelijking tussen de mate van doelrealisatie van het oude curriculum en die van de Basisvorming per definitie onmogelijk. Aan de andere kant is het wel zo dat in maatschappelijk opzicht – bijvoorbeeld door de publieke opinie (vgl. Hofstee, 1992) en vooral ook door de scholen zelf – de effectiviteit van de Basisvorming zeer waarschijnlijk ook – of misschien

wel overwegend – zal worden afgemeten aan de criteria die voorheen ook al werden gehanteerd voor het geven van een oordeel over de mate van doelrealisatie van het onderwijs. Om deze reden (en omdat er anders niets te vergelijken valt ten aanzien van de realisatie van de verhoging van het onderwijspeil) gaan we ervan uit dat met de Basisvorming in ieder geval wordt beoogd dat leerlingen meer opsteken van het onderwijs en dus dat ze betere resultaten behalen in termen van het oude curriculum dan voorheen. Dit sluit overigens niet uit dat eveneens de mate van realisatie ten aanzien van andersoortige doelen die met de Basisvorming worden beoogd dient te worden geëvalueerd. Zo zijn bijvoorbeeld in het VOCL '93 cohort ook de 'algemene vaardigheden' van leerlingen gemeten. Omdat hiervoor echter geen vergelijkingsbasis aanwezig is blijven de resultaten van deze meting in dit artikel buiten beschouwing.

De prestaties van leerlingen in het voortgezet onderwijs zijn uitermate belangrijk voor het niveau van het onderwijs dat ze uiteindelijk behalen. Het gaat hier om het onderwijstype in het voortgezet onderwijs dat ze doorlopen en waarvan ze uiteindelijk een diploma behalen en om het niveau van de vervolgopleiding die ze na het voortgezet onderwijs gaan doen. Dit is de tweede dimensie van het begrip onderwijsopbrengsten. We noemen deze dimensie *individueel rendement* van het onderwijs, waarmee we tot uitdrukking brengen dat het gaat om voor de leerling direct aantoonbaar profijt van het onderwijs.

De derde dimensie van onderwijsopbrengsten is de *efficiëntie*, die betrekking heeft op de tijd die een leerling erover doet om een onderwijstype in het voortgezet onderwijs te voltooien of een zeker opleidingsniveau te behalen. Leerlingen die blijven zitten of tijd verliezen door het horizontaal overstappen van het ene schooltype naar het andere zijn dus minder efficiënt dan leerlingen die een schooltype in de nominale tijd doorlopen.

Zowel de tweede als de derde dimensie van onderwijsopbrengsten zijn gebaseerd op de gedachte dat uitstel van studie- en beroepskeuze – het derde doel van de Basisvorming – moet resulteren in betere keuzen en daardoor in rendabeler en efficiënter loopbanen in het voortgezet onderwijs en in de vervolgopleiding.

Het zal duidelijk zijn dat de drie dimensies van het begrip onderwijsopbrengsten nauw aan elkaar zijn gerelateerd. Het doel van de Basisvorming is om alle drie de typen opbrengsten te verhogen. De vraag die nu gesteld kan worden is voor wie de opbrengsten moeten worden verhoogd. Het antwoord hierop kan worden gebaseerd op het tweede doel van de Basisvorming, het versterken van de gemeenschappelijke culturele basis. Hiermee wordt in eerste instantie bedoeld dat het onderwijsaanbod voor alle leerlingen in de eerste drie jaren verbreed dient te worden en voor alle leerlingen hetzelfde dient te zijn. Men mag dan ook verwachten dat alle leerlingen zoveel mogelijk profiteren van dat aanbod. Met andere woorden, de Basisvorming beoogt *gelijke onderwijskansen* te bieden aan alle leerlingen. Het is echter bekend uit onderzoek naar onderwijskansen dat de prestaties en schoolloopbanen van leerlingen voor een groot deel bepaald worden door de capaciteiten en motivatie van de leerlingen zelf. Niet alle leerlingen zijn in dezelfde mate in staat of bereid om te leren. Daarom moet gelijkheid van kansen liever worden opgevat als 'zoveel mogelijk leerlingen moeten zoveel en zo efficiënt mogelijk leren en profiteren van het onderwijs als hun capaciteiten toelaten'. Dit betekent dat de leerprestaties en schoolloopbanen alleen gerelateerd mogen zijn aan de capaciteiten (en eventueel de motivatie van de leerling) en niet aan SES, sekse of etniciteit. Indien er na controle voor capaciteiten nog een relatie is met SES, sekse of etniciteit is er sprake van milieuspecifieke, seksspecifieke of etnisch specifieke prestaties en schoolloopbanen en dus van groepsgebonden ongelijkheid van kansen. Van de Basisvorming wordt verwacht dat de gelijkheid van kansen toeneemt. Dit zal het geval zijn als de relaties tussen SES, sekse en etniciteit enerzijds en de prestaties en schoolloopbanen anderzijds, na controle voor capaciteiten, sinds de invoering van de Basisvorming zwakker zijn geworden in vergelijking met voorheen. Hiermee is het tweede begrip uit de vraagstellingen gedefinieerd.

Methode

De vraagstellingen zullen worden beantwoord met behulp van de gegevens die zijn verzameld

in de VOCL cohortonderzoeken in het voortgezet onderwijs. Deze cohortstudies zijn specifiek ontworpen voor de evaluatie van de Basisvorming (Peschar, 1988). De cohorten zijn uitgevoerd in opdracht van SVO (sinds 1996 NWO) door het GION (voorheen RION), het CBS, het ITS (alleen VOCL '89) en het OCTO (alleen VOCL '93). De eerste cohortstudie is gestart in 1989, de tweede in 1993, tegelijk met de invoering van de Basisvorming in 1993. De eerste cohortstudie kan worden beschouwd als voormeting waarvan de gegevens kunnen worden vergeleken met die van de tweede cohortstudie, die kan worden beschouwd als name-ting voor het vaststellen van de effecten van de Basisvorming. Hieronder worden de steekproeven, de variabelen en instrumenten, procedures van dataverzameling en analysemethoden beschreven.

Steekproef

De populatie waaruit de steekproef voor de eerste cohortstudie (VOCL '89) getrokken is bestond uit 1800 scholen voor voortgezet onderwijs met ongeveer 185.000 leerlingen in het eerste leerjaar van IVBO (toen nog IBO genoemd), VBO (toen nog LBO), MAVO, HAVO en VWO. De beoogde steekproefomvang bedroeg ongeveer 440 scholen, gestratificeerd naar de onderwijstypen die op de school konden worden gevolgd. De gerealiseerde steekproef bedroeg 381 scholen. Binnen de scholen zijn selecte enkele eerste klassen getrokken waarvan in principe alle leerlingen in de steekproef werden opgenomen. Het totaal aantal eerste klassen bedroeg 868, het totaal aantal leerlingen was 19.524.

Als gevolg van het beleid van schaalvergroting in het voortgezet onderwijs was bij de start van het tweede cohort (VOCL '93) in 1993 de totale populatie van scholen afgenomen tot ongeveer 800. Het was de bedoeling om in de tweede cohortstudie (VOCL '93) zoveel mogelijk scholen op te nemen die ook aan de eerste cohortstudie hadden meegedaan. Dat is niet goed gelukt. Slechts 121 scholen uit de eerste cohortstudie waren bereid ook aan de tweede cohortstudie mee te werken. Uit de 680 resterende scholen moest een nieuwe steekproef worden getrokken om voor de tweede cohortstudie opnieuw een aantal van 400 scholen te realiseren. Slechts 212 scholen van de scholen

uit deze steekproef waren bereid mee te werken, hetgeen de totaal gerealiseerde steekproef voor de tweede cohortstudie bracht op 333. Hoewel het aantal scholen in de tweede cohortstudie lager is dan in de eerste, zijn het aantal klassen en het aantal leerlingen vergelijkbaar. Dit is het gevolg van het feit dat de scholen voor voortgezet onderwijs in 1993 groter waren dan in 1989 en om deze reden meer klassen per school aan het onderzoek deel namen. Het aantal klassen in de tweede cohortstudie bedraagt 870, het aantal leerlingen is 20.331.

Verderop in tabel 1 zullen gegevens worden gepresenteerd waaruit blijkt in hoeverre de twee steekproeven vergelijkbaar zijn naar achtergrondkenmerken van de leerlingen en de onderwijstypen die ze volgden.

Procedures van dataverzameling

De eerste cohortstudie is gestart in het schooljaar 1989/1990 met de dataverzameling bij leerlingen die in dat jaar in de eerste klas zaten. Er zijn entreetoetsen afgenomen en er zijn gegevens verzameld over de achtergrondkenmerken van de leerlingen en de schooltypen waarin ze zaten. De docenten van de scholen hebben de toetsen afgenomen. De informatie over de achtergrondkenmerken van de leerlingen is deels verzameld bij de administratie van de scholen en deels verschaft door de ouders van de leerlingen door middel van het invullen van een vragenlijst. Bij de directeuren van de scholen zijn interviews afgenomen en docenten die les gaven in het eerste leerjaar hebben een vragenlijst ingevuld waarmee informatie is verkregen over school- en klaskenmerken.

In het schooljaar 1991/1992, toen de leerlingen die onvertraagd waren doorgestroomd in de derde klas zaten, zijn opnieuw toetsen afgenomen (wederom door de docenten) als ook interviews bij de directeuren en vragenlijsten bij de docenten van – dit keer – derde klassen. Informatie over de schoolloopbanen van de leerlingen (schooltype en klas, vervolgonderwijs na het voortgezet onderwijs) is elk jaar verzameld. Tot nu toe zijn gegevens beschikbaar tot het schooljaar 1997/1998 (het tiende cohortjaar).

In schooljaar 1993/1994 hebben de leerlingen die in klas 5 van HAVO en VWO zaten een vragenlijst ingevuld over de keuze van het eindexamenpakket, hun toekomstplannen, stu-

dievaardigheden en prestatiemotivatie. Leerlingen die het voortgezet onderwijs ongediplomeerd hadden verlaten en leerlingen die het diploma hadden behaald van VBO of MAVO hebben een vragenlijst ingevuld over hun eind-examenpakket, examenresultaten, overgang naar vervolgonderwijs of arbeidsmarkt en toekomstplannen.

De tweede cohortstudie is begonnen in het schooljaar 1993/1994. De procedure was vrijwel dezelfde als die van de eerste cohortstudie. Er zijn twee verschillen. In de eerste plaats zijn de toetsen in het eerste leerjaar afgenomen door testleiders en de toetsen in het derde leerjaar door hetzij testleiders, hetzij docenten. In dat laatste geval kreeg de school een vergoeding. In de tweede plaats is er in het derde leerjaar een extra toets afgenomen die 'algemene vaardigheden' meet. Van de tweede cohortstudie zijn momenteel de gegevens beschikbaar tot het schooljaar 1997/1998 (het vijfde cohortjaar).

Variabelen en instrumenten

In deze paragraaf worden alleen de variabelen en instrumenten besproken die relevant zijn voor het beantwoorden van de vragenstellingen. Tussen de twee cohorten zullen vergelijkingen worden gemaakt met betrekking tot 1) de prestaties van de leerlingen in het eerste en derde leerjaar, 2) de onderwijsposities van de leerlingen tot het vijfde cohortjaar, en 3) de relaties tussen prestaties en onderwijsposities enerzijds en achtergrondkenmerken van leerlingen anderzijds. Deze drie groepen variabelen zullen hieronder worden besproken.

1. Prestaties van leerlingen

De prestaties in leerjaar 1 zijn gemeten met een entreetoets, bestaande uit de onderdelen rekenen, Nederlands en informatieverwerking. De toets is ontwikkeld door het CITO. Elk toetsonderdeel bestond uit 20 meerkeuze vragen met vier antwoordalternatieven waarvan telkens één antwoord correct was. De inhoud van de entreetoets was in beide cohorten dezelfde. De betrouwbaarheid van de testonderdelen was respectievelijk .76, .84 en .78. De betrouwbaarheid van de hele toets was .91.

De prestaties in het derde jaar van de eerste cohortstudie zijn gemeten met een toets voor tekstbegrip en een toets voor wiskunde. Beide

toetsen zijn speciaal ten behoeve van het cohortonderzoek ontwikkeld door het CITO. De toets voor tekstbegrip bestond uit 40 meerkeuze items met vier antwoordalternatieven waarvan één het correcte antwoord was. De betrouwbaarheid van de toets is .80. De wiskundetoets bestond uit twee versies, versie A voor MAVO-, HAVO- en VWO-leerlingen en versie B voor IBO- en VBO-leerlingen. Versie A bestond uit 12 meerkeuze vragen met vier antwoordalternatieven en uit 20 open vragen; versie B bevatte 2 meerkeuze vragen en 30 open vragen. De puntentoeckenning voor de correcte antwoorden verschilde per vraag, afhankelijk van de moeilijkheidsgraad van het betreffende item. De maximum score bedroeg 140 voor de A-versie en 80 voor de B-versie. De betrouwbaarheid is respectievelijk .80 en .87.

In het derde leerjaar van de tweede cohortstudie zijn de prestaties opnieuw gemeten met een toets voor tekstbegrip en een toets voor wiskunde. De inhoud van de toetsen was in opdracht van SVO inmiddels door het CITO veranderd. Dit om tegemoet te komen aan de veranderingen in de doelen en inhoud van het curriculum die met de Basisvorming werden beoogd. In de toets voor tekstbegrip zijn 13 oude items vervangen door nieuwe items. De betrouwbaarheid is nu .79. In de A-versie van de wiskundetoets zijn 15 items vervangen door 16 nieuwe, in de B-versie zijn 2 oude items vervangen door 3 nieuwe. De maximumscore bedraagt 101 voor de nieuwe A-versie en 77 voor de nieuwe B-versie. De betrouwbaarheid is respectievelijk .81 en .87.

Om vergelijkingen tussen de twee cohorten mogelijk te maken zijn de beide versies (de '89-versie en de '93-versie) van de toets voor tekstbegrip en de vier versies (de A- en B-versie van respectievelijk '89 en '93) van de toets wiskunde onderling vergelijkbaar gemaakt (equivaleren) met behulp van het programma OPLM (One-Parameter Logistic Model; Verhelst, Glas & Verstralen, 1995). Dit was mogelijk doordat in de verschillende versies van de toetsen voor wiskunde en tekstbegrip overlappende items voorkwamen. Met behulp van de scores op deze overlappende items konden ook de scores op de overige items uit de verschillende toetsversies op één onderliggende dimensie worden geschaald. De betrouwbaar-

heid van de ge-equivaleerde toets voor Nederlands is .73 voor de VOCL '89 versie en .75 voor de VOCL '93 versie. De betrouwbaarheden van de A- en B-versie van de beide cohorten variëren tussen .84 en .88. Voor meer informatie over de equivaleringsprocedure wordt verwezen naar het rapport van Van der Werf, Lubbers & Kuyper, 1999.

Voor de vergelijkingen zijn de scores op alle toetsen in leerjaar 1 en 3 na equivalering omgezet naar T-scores met een gemiddelde van 50 en een standaarddeviatie van 10.

2. Onderwijsposities

De onderwijsposities die de leerlingen in het vijfde cohortjaar hebben bereikt (positie 5) zijn uitgedrukt in een score op een aangepaste versie van de leerjarenladder (Bosker, van der Velden & Hofman, 1985). Deze score komt tot stand door een combinatie van het in 1997/1998 gevolgde onderwijstype, het leerjaar waarin de leerling zat en, indien de leerling het onderwijs had verlaten, de examenresultaten van 1996/1997 dan wel de positie in het jaar dat vooraf ging aan het schoolverlaten. De score op de leerjarenladder is een combinatie van het rendement alsook de efficiëntie van de schoolloopbaan van de leerling, zoals gedefinieerd in paragraaf 2. De theoretische range van de leerjarenladder loopt van 1 (advies IBO) tot en met 12 (geslaagd voor het VWO). De maximum score kan worden behaald als een leerling geslaagd is voor het VWO examen zonder te zijn blijven zitten. Het duurt zes jaar om deze score te kunnen bereiken. Omdat in het VOCL '93 cohort de leerlingen in het vijfde cohortjaar zijn, bedraagt de maximum score 10, die bereikt kan worden als leerlingen zonder vertraging in de vijfde klas van het VWO zitten. Leerlingen die zonder vertraging in de vijfde klas van het HAVO zitten hebben de score 9. Leerlingen die geslaagd zijn voor IVBO, VBO of MAVO en daarna het onderwijs hebben verlaten krijgen de waarde van hun laatst bekende positie plus een punt extra. Leerlingen die het onderwijs ongediplomeerd hebben verlaten behouden hun laatst bekende positie. Gediplomeerde leerlingen die een vervolgstudie zijn gaan doen krijgen de waarde van het schooltype en leerjaar waarin ze dan zitten. Voor meer gedetailleerde informatie over deze versie van de leerjarenladder verwij-

zen we naar Van der Werf, Kuyper & Lubbers (1999).

3. Achtergrondkenmerken

Het eerste achtergrondkenmerk is het type voortgezet onderwijs dat aan de leerling is geadviseerd aan het eind van het basisonderwijs. Dit advies is meestal gebaseerd op de prestaties van de leerling in de basisschool en op de score op de CITO eindtoets, die de meeste basisscholen afnemen. Het advies kan tevens worden opgevat als een indicator van het beginniveau van de leerling bij entree in het voortgezet onderwijs. Evenals de onderwijsposities in het vijfde jaar kunnen ook de adviezen worden omgezet in een score op de leerjarenladder. Dit is de startpositie (positie 0) waarvan het scorebereik ligt tussen 1 (advies IVBO) en 5 (advies VWO). Dubbele adviezen krijgen de tussenliggende halve waarde, bijvoorbeeld VBO/MAVO is 2,5 (zie Van der Werf, Kuyper & Lubbers, 1999).

Het tweede achtergrondkenmerk is het onderwijstype (klastype) waarin de leerling in het eerste leerjaar zit. Ook dit klastype in leerjaar 1 kan worden uitgedrukt in een score op de leerjarenladder. Dit is dan positie 1, waarvan het scorebereik ligt tussen 2 (IVBO) en 6 (VWO). Gecombineerde klastypen krijgen de tussenliggende waarde, bijvoorbeeld een MAVO/HAVO klas krijgt de score 4,5.

Het derde achtergrondkenmerk is de sociaal-economische status (SES) van het gezin, geoperationaliseerd als het hoogst voltooide opleidingsniveau van de vader en de moeder. Het scorebereik ligt tussen 2 (alleen lager onderwijs voltooid) en 6 (universiteit voltooid).⁴

Het vierde achtergrondkenmerk is de etniciteit van de leerling, geoperationaliseerd met de combinatie van geboorteland en nationaliteit van de ouders en van het kind zelf. De volgende categorieën worden onderscheiden: 1) Nederland; 2) Suriname, Antillen en Aruba (ASA); 3) Marokko; 4) Turkije; en 5) overige landen.

Het vijfde achtergrondkenmerk is de intelligentie van de leerling, zoals in het eerste leerjaar gemeten met de PSB3 en de PSB8 (Prüf-system für Schul- und Bildungsberatung; Horn, 1969). De eerste subtest meet redeneervermogen, de tweede abstractievermogen. Beide sub-

testen zijn non-verbaal en bestaan uit 40 items.

Alle vijf de achtergrondkenmerken zijn in beide cohorten op dezelfde wijze gemeten, met uitzondering van de nationaliteit van de leerling, die in het '93 cohort bij de school is opgevraagd en in het '89 cohort bij de ouders.

Analyse van de gegevens

Allereerst worden voor de beide cohorten beschrijvende gegevens gepresenteerd: voor de achtergrondkenmerken sekse en etniciteit de frequentieverdelingen en voor de achtergrondkenmerken advies, gevolgd schooltype in leerjaar 1, intelligentie en SES de gemiddelden en standaarddeviaties. Vervolgens worden de verschillen tussen de cohorten in de prestaties in het eerste en derde leerjaar en de onderwijsposities in het vijfde cohortjaar getoetst. Ook worden de gegevens van het derde en vijfde jaar voor de gevolgde onderwijstypen afzonderlijk gepresenteerd, alsmede correlaties tussen de scores op intelligentie en de entreetoets in het eerste jaar en de prestaties en onderwijsposities in het derde respectievelijk vijfde jaar. Daarna worden multiple regressieanalyses gerapporteerd met als predictoren cohort en gevolgd onderwijstype en als criteriumvariabelen de prestaties in het derde jaar en de onderwijspositie in het vijfde jaar en met als controlevariabelen het advies, de score op de intelligentietest en de score op de entreetoets. Met de resultaten van deze regressieanalyses

wordt de eerste vraagstelling beantwoord.

Voor de beantwoording van de tweede vraagstelling worden de beschrijvende gegevens met betrekking tot de prestaties in het eerste en derde leerjaar en de onderwijsposities in het vijfde jaar voor elk cohort afzonderlijk gepresenteerd voor jongens en meisjes en voor de verschillende categorieën van etniciteit, alsmede de correlaties met SES. De verschillen tussen de cohorten en tussen de groepen leerlingen worden getoetst in een drietal multiple regressieanalyses, waarin cohort en de groepsvariabele (respectievelijk sekse, etniciteit en SES) de predictorvariabelen zijn, de scores op de toetsen in leerjaar 3 en de onderwijsposities in het vijfde jaar de criteriumvariabelen en het gevolgde onderwijstype, het advies, de score op de intelligentietest en de score op de entreetoets de controlevariabelen.

Resultaten

Verschillen tussen de cohorten in achtergrondkenmerken

In tabel 1 zijn de gemiddelde scores op de leerlingachtergrondkenmerken gepresenteerd, evenals de standaarddeviaties en de aantallen leerlingen. Voor de variabelen etniciteit en sekse zijn de percentages leerlingen in elke categorie weergegeven.

Tabel 1
Vergelijking tussen de cohorten op achtergrondkenmerken van leerlingen

	VOCL '89			VOCL '93			score bereik
	Gem.	Sd.	N	Gem.	Sd.	N	
Advies	3.0	1.0	18.854	3.0	1.1	19.252	1 - 5
Schooltype Ij1	4.2	1.0	19.484	4.1	1.1	19.914	2 - 6
PSB-3	23.6	5.3	19.084	25.2	5.4	19.205	1-40
PSB-8	20.9	7.0	19.096	23.8	7.1	19.208	1-40
SES	3.6	1.1	16.710	3.9	1.0	16.680	2 - 6
	%		N	%		N	
Sekse			19.524			20.331	
meisjes	48.1		9.384	48.6		9.879	
jongens	51.9		10.140	51.4		10.452	
Etniciteit			16.826			20.331	
Nederlands	89.7		14.785	87.0		17.694	
ASA*	2.0		330	1.9		391	
Marokkaans	2.0		336	2.2		446	
Turks	2.1		347	2.6		528	
Overig	6.1		1.027	6.3		1.272	

* ASA is: Antilliaans, Surinaams of Arubaans

Tabel 2

Vergelijking van de cohorten op prestaties in leerjaar 1 en 3 en onderwijsposities in cohortjaar 5

	VOCL '89			VOCL '93			t-toets	
	Gem.	Sd	N	Gem.	Sd.	N	t	eff.gr
Leerjaar 1								
Nederlands	49.5	10.1	18.526	50.5	9.9	19.131	9.10	.10
Rekenen	49.6	10.0	18.456	50.4	9.9	19.144	8.53	.08
Informatie	49.6	9.9	18.335	50.4	10.1	19.117	7.26	.08
Totale test	49.5	10.0	18.159	50.4	10.0	19.097	8.51	.09
Leerjaar 3								
Tekstbegrip	50.0	9.6	10.496	50.0	9.6	8.903	-2.1	.00
Wiskunde	49.7	10.2	10.831	50.4	9.7	8.100	4.53	.07
Cohortjaar 5								
Onderwijspositie	7.5	1.5	18.926	7.5	1.5	19.692	4.23	.05

Vetgedrukte t-waarden zijn significant bij $p < .001$

De tabel laat zien dat de gemiddelde scores op SES en de beide subtesten van de PSB in 1993 hoger zijn dan in 1989. De verschillen zijn significant met een p -waarde kleiner dan .001. Uitgedrukt in effectgroottes gaat het om relevante verschillen (.30/.40). Het aantal leerlingen van wie bekend is dat ze van niet-Nederlandse afkomst zijn, is in het '93 cohort bijna één procent hoger dan in het cohort '89.

Verschillen tussen de cohorten ten aanzien van prestaties en onderwijsposities

In tabel 2 staan de gemiddelden en standaarddeviaties van beide cohorten van de toetsscores in leerjaar 1 en 3 en van de onderwijsposities in het vijfde jaar. De toetsscores zijn in alle gevallen T-scores met een overall gemiddelde (het gemiddelde over beide cohorten) van 50 en een standaarddeviatie van 10. De verschillen tussen de cohorten zijn getoetst met t -toetsen. De resultaten van deze toetsing staan eveneens in de tabel.

Met uitzondering van tekstbegrip in leerjaar 3 zijn alle verschillen significant in het voordeel

van het '93 cohort. Alle effectgroottes liggen echter ver onder de .25, de waarde die in het algemeen als een relevante effectgrootte beschouwd wordt. In zijn algemeenheid kunnen we dus concluderen dat de onderwijsopbrengsten sinds de invoering van de Basisvorming (vrijwel) gelijk zijn gebleven. In tabel 3 wordt nagegaan of deze conclusie geldt voor alle onderwijstypen.

De tabel laat zien dat op tekstbegrip de prestaties zijn gedaald in het IVBO, het VBO en het VWO. De verschillen zijn significant in het IVBO en VWO ($p < .005$) en in het VBO ($p < .001$). De effectgroottes variëren van .03 voor het MAVO tot -.12 voor het VWO.

Voor wiskunde zijn de prestaties significant gestegen in het VBO ($p < .005$) en het MAVO ($p < .001$). In het IBO en het HAVO/VWO is een niet-significante stijging van de prestaties waar te nemen. De wiskunde-prestaties in het HAVO en het VWO zijn gedaald, maar de daling is alleen significant in het VWO (bij $p < .001$). De effectgroottes variëren van .01 voor het IVBO tot -.16 voor het VWO.

Tabel 3

Cohortvergelijking per onderwijstype (gemiddelden en resultaten t -toets)

	Tekstbegrip			Wiskunde			Ondpos5		
	'89	'93	t	'89	'93	t	'89	'93	t
IVBO	39.8	38.9	-2.78	32.8	33.8	1.83	5.6	5.6	1.37
VBO	43.9	43.3	-3.59	43.5	44.3	3.42	6.6	6.7	6.29
MAVO	49.5	49.8	1.16	50.4	51.2	4.99	7.7	7.8	4.24
HAVO	53.9	54.5	1.67	55.9	55.6	-1.46	8.6	8.5	-.88
H/V	57.6	58.5	1.36	58.7	58.9	.42	9.1	9.1	-.61
VWO	61.6	60.4	-3.48	61.8	60.2	-7.23	9.7	9.6	-3.18

Vetgedrukte t-waarden significant bij minimaal $p < .005$

Tabel 4

Correlaties tussen advies, intelligentie en score op de entreetoets en de prestaties en onderwijspositie, per cohort

	Advies Ij 1		Intelligentie Ij 1		Entreetoets Ij 1	
	'89	'93	'89	'93	'89	'93
Leerjaar 3						
Tekstbegrip	.65	.59	.20	.28	.67	.61
Wiskunde	.73	.70	.31	.41	.76	.76
Cohortjaar 5						
Onpos5	.73	.74	.24	.33	.68	.70

Wat betreft de onderwijsposities in het vijfde jaar is er een lichte stijging in het IVBO, het VBO en het MAVO en een lichte daling in het HAVO en het VWO. De stijging is significant in het VBO en het MAVO ($p < .001$) evenals de daling in het VWO ($p < .005$). De effectgroottes variëren van .01 in het HAVO tot -.05 in het VWO.

De conclusie tot dusver moet zijn dat er in de specifieke schooltypen geen sprake is van een relevante stijging van de onderwijsopbrengsten. In sommige schooltypen is een lichte stijging waarneembaar, terwijl in andere schooltypen, met name in de hogere, een daling is te constateren. Het voornaamste resultaat is de consistente daling voor VWO-leerlingen en de consistente geringe stijging voor MAVO-leerlingen, waarbij nog niet is gecontroleerd voor aanvangsverschillen.

In tabel 4 staan de correlaties van 1) het advies (onderwijspositie 0), 2) de score op de intelligentietest (PSB3 + PSB8) en 3) de score op de entreetoets (de score op de drie toetsonderdelen samen) met de score op de toetsen wiskunde en tekstbegrip in leerjaar 3 en de onderwijspositie in het vijfde jaar.

Alle correlaties zijn significant ($p < .001$). Er zijn enkele zaken die opvallen. In de eerste plaats is te zien dat de score op de entreetoets de prestaties in leerjaar 3 en de onderwijspositie in het vijfde jaar beter voorspelt dan de score op de intelligentietest. In de tweede plaats voorspelt de score op de entreetoets beter de prestaties in het derde jaar, terwijl het advies beter de onderwijspositie in het vijfde jaar voorspelt. In de derde plaats is er in het '93 cohort een sterker verband tussen de score op de intelligentietest en de onderwijsopbrengsten dan in het '89 cohort.

In het voorgaande is geconstateerd dat er tussen de cohorten verschillen zijn in het beginniveau van de leerlingen (intelligentie en score op de entreetoets), maar nauwelijks in de onderwijsopbrengsten (toetsprestaties in leerjaar 3 en onderwijsposities in het vijfde jaar). Wel is geconstateerd dat in sommige onderwijstypen de opbrengsten iets zijn gestegen, terwijl deze in andere schooltypen iets zijn gedaald. Om deze redenen is met multiple regressie-analyses nagegaan in welke mate het cohort waartoe de leerling behoort en het onderwijstype dat de leerling in het derde jaar volgt voorspellend zijn voor de onderwijsopbrengsten, nadat gecontroleerd is voor het beginniveau van de leerlingen. Hiertoe zijn de variabelen advies, intelligentie en score op de entreetoets 'gedwongen' als co-varianten in de analyses opgenomen. In tabel 5 staan de resultaten. In de analyses is de variabele onderwijstype in leerjaar 3 omgezet naar een score op de leerjarenladder (positie 3), waarvan het scorebereik loopt van 4 tot 8. De analyses met als criteriumvariabele de onderwijspositie in het vijfde jaar zijn twee keer uitgevoerd, een keer met (ondpos5A) en een keer zonder (ondpos5B) de onderwijspositie in het 3e jaar als predictorvariabele. Dit om na te gaan of het voor het cohorteffect veel uitmaakt of deze variabele – die uiteraard sterk samenhangt met de criteriumvariabele – al dan niet als voorspeller voor onderwijspositie in het vijfde jaar wordt opgenomen.

De resultaten laten zien dat er, na controle voor advies, intelligentie en de score op de entreetoets, door de dummyvariabele 'cohort' een significante bijdrage wordt geleverd aan de voorspelling van de onderwijsopbrengsten. Deze bijdrage is negatief voor de score op de toetsen wiskunde en tekstbegrip en positief

Tabel 5

Prestaties en onderwijspositie voorspeld uit beginniveau, onderwijstype en cohort (regressiecoëfficiënten en verklaarde variantie).

	Tekstbegrip	Wiskunde	Ondpos5 (A)	Ondpos5 (B)
Co-variaten				
B Advies	.092	.100	.007	.530
B Intelligentie	.011	.091	.006	.015
B Entreetoets	.354	.450	.109	.295
Predictoren				
B Ondpos3	.281	.264	.795	---
B Cohort (93)	-.023	-.024	.017	.021
Verkl. Variantie (R^2)				
Co-variaten	.445	.618	.578	.620
Predictoren	.397	.516	.783	.001
Totaal	.463	.634	.789	.620

Vetgedrukte regressiecoëfficiënten zijn significant bij $p < .001$

voor de onderwijspositie. Dit betekent dat, hoewel de prestaties in het derde leerjaar absoluut gezien in cohort '93 hoger zijn dan in cohort '89, zoals in tabel 2 is aangetoond, er na controle voor het instroomniveau, sprake is van een daling. Voor de onderwijspositie in het vijfde jaar geldt dat er absoluut gezien geen sprake is van een verschil tussen beide cohorten, maar dat er sinds de invoering van de Basisvorming, na controle voor het instroomniveau, sprake is van een stijging in de onderwijspositie. Behalve naar de voorspelling door 'cohort' is ook gekeken naar de bijdrage die het gevolgde onderwijstype in leerjaar 3 levert aan de voorspelling van de onderwijsopbrengsten. Het hoofdeffect van deze variabele is beduidend groter dan dat van cohort. Voor de voorspelling van de onderwijspositie in het vijfde jaar is het een zeer sterke voorspeller, maar dit ligt ook voor de hand. Wanneer deze variabele wordt weggelaten wordt de bijdrage deels overgenomen door advies en deels door de score op de entreetoets. Het cohorteffect is echter in beide analyses significant positief.

Tot nu toe is aangetoond dat er overall in het '93 cohort sprake is van lagere prestaties en van hogere onderwijsposities in vergelijking met het '89 cohort. Omdat in tabel 3 bleek dat de verschillen in opbrengsten tussen de cohorten verschillen naar onderwijstype, is ook de 'interactie' van cohort en onderwijstype in leerjaar 3 in de voorspelling opgenomen. De percentages verklaarde variantie nemen hierdoor nauwelijks toe. De interactie-effecten zijn significant voor de wiskundescores en de onderwijspositie in het vijfde jaar, en wel beide

bij $p < .001$. De regressiecoëfficiënten zijn respectievelijk $-.188$ en $-.101$. De hoofdeffecten van cohort zijn ook beide significant bij $p < .001$. De regressiecoëfficiënten zijn respectievelijk $.158$ en $.113$. Nadere bestudering van de regressiecoëfficiënten voor wiskunde laat zien dat het negatieve interactie-effect groter is dan het positieve hoofdeffect. Als we de effecten van het cohort voor elk afzonderlijk onderwijstype nu uitschrijven in een regressievergelijking zal blijken dat het cohorteffect voor elk onderwijstype negatief is, en meer negatief is naarmate het onderwijstype hoger is. Voor de onderwijspositie in het vijfde jaar ligt het iets anders. Het positieve hoofdeffect van cohort is iets sterker dan het negatieve-interactie effect. Nadere bestudering van de regressievergelijking laat hier zien dat het cohorteffect voor het laagste onderwijstype vrijwel 0 is en iets toeneemt met het in leerjaar 3 gevolgde onderwijstype.

Verschillen in onderwijskansen tussen cohorten

In de vorige paragraaf is de vraag naar veranderingen in de onderwijsopbrengsten sinds de invoering van de Basisvorming beantwoord. In deze paragraaf gaat het om de verandering in groepsgebonden onderwijsopbrengsten. Allereerst worden in de tabellen 6 tot en met 8 voor beide cohorten de samenhangen van respectievelijk sekse, SES en etniciteit met de toetsscores in leerjaar 3 en de onderwijsposities in het vijfde jaar gepresenteerd. Voor de variabele sekse en etniciteit staan de gemiddelden op de opbrengstvariabelen in de tabel, evenals de t -waarden respectievelijk de F -waarden van de

Tabel 6
Verschillen in onderwijsopbrengsten tussen jongens en meisjes, per cohort

	Tekstbegrip		Wiskunde		Ondpos5	
	'89	'93	'89	'93	'89	'93
Jongens	49.5	48.5	50.3	51.0	7.29	7.39
Meisjes	50.2	51.5	49.1	49.7	7.63	7.70
t-waarde	-5.48	-13.91	6.27	5.88	-16.18	-14.47
Effectgrootte	-.07	-.31	.12	.13	-.23	-.21

Vetgedrukte t-waarden zijn significant bij $p < .001$

Tabel 7
Correlaties tussen SES en de onderwijsopbrengsten, per cohort

	Tekstbegrip		Wiskunde		Ondpos5	
	'89	'93	'89	'93	'89	'93
SES	.30	.33	.36	.37	.42	.40

toetsing van de verschillen tussen de groepen leerlingen en de effectgroottes. Voor de variabele SES gaat het om correlaties met de opbrengstvariabelen.

Het verschil in prestatie tussen jongens en meisjes op de toets tekstbegrip is aanmerkelijk groter geworden, terwijl de verschillen op wiskunde en de onderwijspositie in het vijfde jaar vrijwel gelijk zijn gebleven. In beide cohorten scoren meisjes hoger op tekstbegrip en behalen meisjes een hogere onderwijspositie, terwijl jongens hoger scoren op wiskunde.

Alle correlaties in Tabel 7 zijn significant. ($p < .001$). SES correleert het sterkst met de onderwijspositie in het vijfde jaar en het minst sterk met de scores op de toets tekstbegrip. Vergelijking tussen de cohorten laat zien dat de correlaties met de toetsprestaties iets zijn geste-

gen, terwijl de correlatie met de onderwijspositie in het vijfde jaar iets is gedaald.

Tabel 8 laat zien of de relaties met etniciteit ook zijn veranderd sinds de invoering van de Basisvorming. De tabel bevat de gemiddelde scores van de onderscheiden etnische categorieën in elk cohortjaar, de resultaten van de toetsing van de verschillen tussen deze categorieën in elk cohortjaar en de effectgroottes voor elke etnische groep in vergelijking met de Nederlandse leerlingen.

De effectgroottes in de tabel laten enkele opvallende zaken zien. In de eerste plaats zijn de achterstanden van de Turkse en Marokkaanse leerlingen ten opzichte van Nederlandse leerlingen toegenomen voor de scores op de toets tekstbegrip. De achterstand van de ASA leerlingen is afgenomen. De leerlingen uit de categorie 'overig' hebben een grotere voor-

Tabel 8
Verschillen in onderwijsopbrengsten tussen etnische groepen, per cohort

	Tekstbegrip		Wiskunde		Ondpos5	
	'89	'93	'89	'93	'89	'93
Nederlands	50.3	50.2	50.2	50.8	7.56	7.59
ASA	46.7	47.5	45.4	44.3	6.91	6.95
Marokkaans	44.7	43.7	41.4	42.2	6.61	6.61
Turks	44.2	42.1	42.4	41.2	6.58	6.62
Overig	50.4	50.9	50.5	50.7	7.49	7.59
F-waarde	33.03	35.40	61.64	66.21	84.18	110.22
Eff.gr.ASA	-.37	-.28	-.47	-.67	-.44	-.44
Eff.gr. Marok.	-.58	-.68	-.86	-.89	-.65	-.67
Eff.gr. Turks	-.63	-.84	-.77	-.99	-.67	-.66
Eff.gr. Overig	.01	.07	.03	.01	.05	.00

Vetgedrukte F-waarden significant bij $p < .001$

sprong gekregen. Voor de scores op de toets wiskunde geldt dat zowel de Turkse en de Marokkaanse als ook de ASA leerlingen achteruit zijn gegaan ten opzichte van de Nederlandse leerlingen, terwijl de leerlingen uit de categorie 'overig' een iets kleinere voorsprong hebben gekregen. De achterstand in de onderwijspositie is voor de Marokkaanse leerlingen iets toegenomen, voor de ASA leerlingen gelijk gebleven en voor de Turkse leerlingen iets afgenomen. De voorsprong van de 'overige' leerlingen op de Nederlandse leerlingen is verdwenen. De Turkse en Marokkaanse leerlingen hebben in beide cohorten een vergelijkbare achterstand ten opzichte van Nederlandse leerlingen, de ASA leerlingen hebben in beide cohorten minder achterstand dan de Turkse en Marokkaanse leerlingen en de leerlingen in de categorie 'overig' hebben vrijwel dezelfde onderwijspositie als de Nederlandse leerlingen.

In het voorgaande is geconstateerd dat de relaties tussen SES, etniciteit en sekse en de onderwijsopbrengsten soms zijn gedaald en soms zijn gestegen. Al eerder bleek dat de onderwijsopbrengsten sterk samenhangen met de achtergrondkenmerken intelligentie, advies, score op de entreetoets en het gevolgde onderwijstype. Waar het nu om gaat is vast te stellen of er nog steeds een effect is van respectievelijk SES, etniciteit en sekse, nadat gecontroleerd is voor de achtergrondkenmerken en of dat effect hetzelfde is in beide cohorten. Dit is nagegaan in drie maal drie multiple regressie-analyses. Er zijn drie criteriumvariabelen, namelijk de

scores op de toetsen Nederlands en wiskunde en de onderwijspositie in het vijfde jaar. In elke analyse zijn de achtergrondkenmerken intelligentie, advies en score op de entreetoets als covariaten opgenomen, in de analyses met de toetsscores als criteriumvariabele bovendien ook de onderwijspositie in het derde jaar. Elk van de analyses is drie maal uitgevoerd, namelijk met respectievelijk sekse, SES en etniciteit als predictoren. In deze analyses zijn de variabelen sekse, etniciteit en cohort opgenomen als dummy variabelen (1= respectievelijk meisje, ASA/Turks/Marokkaans en cohort '93). In elk van deze analyses is ook de interactie van de betreffende predictor met cohort opgenomen, omdat we geïnteresseerd zijn in de vraag of het effect van de betreffende predictor hetzelfde is in beide cohorten.

Tabel 9 bevat de resultaten van de regressie analyses met sekse als predictor, tabel 10 bevat de resultaten met SES als predictor en tabel 11 de resultaten met etniciteit als predictor.

De resultaten van de regressie-analyses in Tabel 9 bevestigen het beeld dat ook al in tabel 6 naar voren kwam. Voor tekstbegrip zijn de verschillen tussen jongens en meisjes, in het voordeel van meisjes, in cohort '93 groter dan in cohort '89. Op wiskunde scoren jongens hoger dan meisjes, maar dit is in beide cohorten in dezelfde mate het geval. Op de onderwijsposities in het vijfde jaar scoren meisjes hoger dan jongens en dit verschil is in het '93 cohort wat kleiner dan in het '89 cohort.

Tabel 9

*Voorspelling van prestaties en onderwijspositie uit beginniveau, sekse, cohort en sekse*cohort (regressiecoëfficiënten en verklaarde variantie).*

	Tekstbegrip	Wiskunde	Ondos5 (A)	Ondpos5 (B)
Co-variaten				
B Advies	.088	.111	.004	.514
B Intelligentie	.009	.094	.005	.012
B Entreetoets	.369	.427	.119	.310
B Ondpos3	.267	.282	.787	----
Predictoren				
B Cohort (93)	-.072	-.018	.025	.031
B Sekse (meisje)	.027	-.102	.049	.093
B Cohort*sekse	.080	-.008	-.014	-.018
Verkl. Variantie (R²)				
Co-variaten	.462	.633	.789	.620
Predictoren	.012	.004	.015	.015
Totaal	.470	.645	.790	.627

Vetgedrukte regressiecoëfficiënten significant bij minimaal $p < .005$

Tabel 10

Voorspelling van prestaties en onderwijspositie uit beginniveau, SES, cohort en SES*cohort (regressiecoëfficiënten en verklaarde variantie).

	Tekstbegrip	Wiskunde	Ondos5 (A)	Ondpos5 (B)
Co-variaten				
B Advies	.077	.099	.003	.515
B Intelligentie	.012	.091	.006	.013
B Entreetoets	.035	.448	.098	.274
B Ondpos3	.285	.252	.798	---
Predictoren				
B Cohort (93)	-.020	.005	.028	.001
B SES	.029	.034	.032	.077
B Cohort*SES	-.003	-.029	-.016	.016
Verkl. Variantie (R^2)				
Co-variaten	.458	.629	.794	.621
Predictoren	.100	.135	.174	.174
Totaal	.459	.630	.795	.627

Vetgedrukte regressiecoëfficiënten significant bij minimaal $p < .005$

Het interactie effect is voor geen van de criterium variabelen significant. Het effect van SES op de onderwijsopbrengsten is in beide cohorten hetzelfde, en wel positief: hoe hoger het sociaal milieu, des te hoger de onderwijsopbrengsten. Opvallend is dat het cohort effect nu niet meer of alleen nog bij $p < .005$ significant is.

Het interactie-effect SES * cohort (Tabel 10) is significant ($p < .005$) voor de voorspelling van de onderwijspositie in het vijfde jaar waarbij ook de onderwijspositie in het derde jaar als co-variant is opgenomen (Tabel 11). De interpretatie is dat het verschil tussen Nederlandse en allochtone leerlingen ten nadele van de laatsten in het '93 cohort iets kleiner is geworden. Met andere woorden, allochtone leerlingen in

het '93 cohort hebben hun achterstand in onderwijspositie in het vijfde cohortjaar iets ingelopen in vergelijking met het '89 cohort. Gezien de hoge p -waarde en het grote aantal leerlingen dat in de analyse betrokken is moet dit resultaat effect echter als verwaarloosbaar worden beschouwd.

Conclusies en discussie

De conclusies die uit de gepresenteerde gegevens getrokken kunnen worden spreken eigenlijk grotendeels voor zichzelf. In de eerste plaats blijkt dat de opbrengsten van de Basisvorming nauwelijks beter zijn dan die van het voortgezet onderwijs 'oude stijl'. Enerzijds is de effectiviteit – de mate van doelrealisatie in

Tabel 11

Voorspelling van prestaties en onderwijspositie uit beginniveau, etniciteit, cohort en etniciteit*cohort (regressiecoëfficiënten en verklaarde variantie).

	Tekstbegrip	Wiskunde	Ondos5 (A)	Ondpos5 (B)
Co-variaten				
B Advies	.087	.105	.010	.535
B Intelligentie	.012	.090	.005	.016
B Entreetoets	.346	.445	.100	.291
B Ondpos3	.288	.256	.799	---
Predictoren				
B Cohort (93)	-.023	-.025	.011	.014
B Etniciteit (all)	-.026	-.037	-.026	-.013
B Cohort*etniciteit	.002	-.007	.012	.012
Verkl. Variantie (R^2)				
Co-variaten	.460	.630	.791	.622
Predictoren	.013	.025	.015	.015
Totaal	.461	.632	.791	.622

Vetgedrukte regressiecoëfficiënten significant bij minimaal $p < .005$

termen van geleerde kennis en vaardigheden – wat lager, anderzijds zijn de onderwijsposities iets verbeterd zoals blijkt uit scores op de leerjarenladder die een combinatie zijn van rendement en efficiëntie. Deze effecten zijn wat sterker voor de hogere schooltypen. In HAVO en VWO liggen de toetsprestaties van het '93 cohort lager dan die van het '89 cohort, terwijl voor de positie op de leerjarenladder in het vijfde jaar het omgekeerde het geval is. Leerlingen van het '93 cohort die in het derde leerjaar in het HAVO of VWO zaten hebben in het vijfde jaar een iets hogere score op de leerjarenladder dan dezelfde groepen leerlingen in het '89 cohort hadden. Wellicht wordt dit veroorzaakt door het feit dat in het '93 cohort leerlingen minder vaak zijn blijven zitten dan in het '89 cohort het geval was. Het kan ook zijn dat er meer opstroom heeft plaatsgevonden. Met andere woorden, het is nog niet duidelijk of hier sprake is van verbetering van het rendement dan wel van verhoging van de efficiëntie. Hoe het precies zit zal nog verder worden uitgezocht in analyses van de jaarlijkse cohortgegevens over zittenblijven, opstroom, afstroom en dergelijke. Voor meer informatie hierover wordt verwezen naar het rapport van Van der Werf et al (1999).

In het kader van dit artikel is alleen de constatering van verbetering van de onderwijsposities van belang, niet de wijze waarop de scores op de leerjarenladder in het '93 cohort tot stand zijn gekomen in vergelijking met de scores van het '89 cohort. Het gaat hier namelijk om een opmerkelijk fenomeen: de onderwijsposities zijn verbeterd terwijl de prestaties omlaag zijn gegaan. Dit duidt erop dat in het voortgezet onderwijs prestaties minder een rol spelen bij overgangsbepalingen en (beslissingen tot) keuze van onderwijstypen en dus dat het voortgezet onderwijs minder selectief is geworden op grond van prestatiecriteria. Opmerkelijk is ook dat dit meer het geval is in de hogere dan in de lagere onderwijstypen. Hoewel het om kleine verschillen gaat, is het raadzaam om in de gaten te houden of het hier gaat om een voortschrijdende trend. Als dit het geval is, is de kwaliteit van het onderwijs op de langere termijn in het geding. Het bijbrengen van kennis en vaardigheden als belangrijkste doel van het onderwijs wordt daarmee ongeschikt gemaakt aan het streven naar verho-

ging van het rendement en de efficiëntie van het onderwijs. Anderzijds is het wel zo dat de prestaties van de leerlingen op de in dit onderzoek afgenomen toetsen niet volledig overeen hoeven te komen met hun werkelijke schoolprestaties. De scores op de toetsen hebben voor de leerlingen geen gevolgen voor hun schoolloopbaan, zodat het niet valt uit te sluiten dat ze niet speciaal hun best hebben gedaan om de toets te maken. Daarnaast is het mogelijk dat bij de beslissingen van docenten (en leerlingen zelf) over de schoolloopbaan andere criteria dan scores op traditionele toetsen een rol spelen, zoals bijvoorbeeld de motivatie en inzet van leerlingen of dat docenten zelf toetsen afnemen die meer recht doen aan de doelen die met de Basisvorming worden beoogd (zie ook Roelofs, Vermeulen & Houtveen, 1998).

De tweede conclusie die uit de resultaten getrokken kan worden is dat sinds de invoering van de Basisvorming de mate van groepsgebonden ongelijkheid van kansen niet is afgenomen. De verschillen in prestaties en onderwijspositie, samenhangend met sekse, SES en etniciteit, na controle voor intelligentie, advies en score op de entreetoets, zijn voor leerlingen die Basisvorming hebben gehad even groot als voor leerlingen die geen Basisvorming hebben gehad. Er zijn ook geen tekenen die erop wijzen dat dit soort gelijkheid van kansen in de toekomst zal toenemen. Integendeel, het feit dat hierboven een tendens tot een afnemende mate van selectie op grond van prestaties is geconstateerd, wijst op een verminderd meritocratisch onderwijssysteem in plaats van op een vermeerdering hiervan. En een verminderde meritocratisering van het onderwijs draagt de gevaren in zich van het hanteren van oneigenlijke selectie – door individuen zelf dan wel door anderen – op grond van bijvoorbeeld sekse, SES of etniciteit. In dat geval zal de groepsgebonden ongelijkheid in plaats van kleiner juist groter worden.

We moeten concluderen dat de Basisvorming nog niet tot een verbetering van de onderwijsopbrengsten heeft geleid, althans niet ten aanzien van de definities die in dit artikel zijn gehanteerd. Over andersoortige opbrengsten valt op basis van de hier gepresenteerde gegevens niets te zeggen. Wel moet er voor worden gewaakt dat er geen negatieve effecten – in termen van lagere prestaties in de traditionele

kennis en vaardigheden zullen optreden. Aan de andere kant moeten we ook niet te pessimistisch zijn. Het '93 cohort was de eerste lichting leerlingen die Basisvorming hebben gehad. Deze leerlingen hadden te maken met docenten die voor het eerst moesten lesgeven met een nieuwe methode, met andere – meer vak overstijgende – doelen en een gemoderniseerde didactiek. Voor deze leerlingen was het onderwijsaanbod ook duidelijk zwaarder – in de zin van meer vakken – dan dat voor de eerdere lichtingen leerlingen. Gegeven deze situatie is het wellicht gunstig dat er geen sprake is van een dramatische daling van de onderwijsopbrengsten, zoals heel vaak het geval is bij invoering van onderwijsvernieuwingen. Zoals in de inleiding is opgemerkt gaat het hier om een tussenbalans. De effecten van de Basisvorming kunnen pas echt goed beoordeeld worden bij een nieuwe lichting leerlingen die in het voortgezet onderwijs komen op het moment dat de invoeringsproblemen grotendeels zijn overwonnen. De onderwijsopbrengsten van het binnenkort te starten VOCL '99 cohort zullen moeten uitwijzen of de eerder gevonden resultaten het begin zijn van voortschrijdende daling van de kwaliteit van het onderwijs dan wel het startpunt voor een verbetering.

Literatuur

- Bosker, R.J., Velden, R.K.W. van der & Hofman, W.H.A. (1985). *Een generatie geselecteerd. Deel I: De loopbanen*. Groningen: RION.
- Groot, A.D. de (1983). Is de kwaliteit van het onderwijs te beoordelen? In B. Creemers, W. Hoebe & K. Koops (red.). *De kwaliteit van het onderwijs* (pp.54-73). Groningen: Wolters-Noordhoff.
- Hofstee, W.K.B. (1992). Independent Educational Assessment in the Interest of Educational Quality. In P. Vedder (red.). *Measuring the Quality of Education* (pp. 23-35). Amsterdam/Lisse: Swets & Zeitlinger.
- Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung PSB*. Göttingen: Verlag für Psychologie.
- Peschar, J. (1988). *Evaluatie van de basisvorming. Kader voor het uitvoeringsplan*. Den Haag: SVO.
- Procesmanagement Basisvorming (1993). *Inrichtingsbesluit VWO-HAVO-MAVO-VBO. Besluit kerndoelen en adviesurentabel basisvorming*

1993-1998. Almere: Procesmanagement Basisvorming.

- Roelofs, E.C., Vermeulen, C.J. & Houtveen, A.A.M. (1998). *Basisvorming op weg. Onderzoek naar de mening van docenten over de realisatie van de Basisvorming*. Utrecht: ISOR.
- Verhelst, N.D., Glas, C.A.W. & Verstralen, H.H.F.M. (1995). *One-Parameter Logistic Model (OPLM)*. Arnhem: CITO.
- Werf, M.P.C. van der, Kuyper, H. & Lubbers, M.J. (1999). *Achtergrond- en gezinskenmerken van leerlingen en opbrengsten van het voortgezet onderwijs*. Groningen: GION.
- Werf, M.P.C. van der, Lubbers, M.J. & Kuyper, H. (1999). *De opbrengsten van het voortgezet onderwijs voor de leerlingen uit de VOCL-cohorten 89-1 en 93-1*. Groningen: GION (verschijnt binnenkort).

Manuscript aanvaard 20-6-1999

Auteurs

Greetje van der Werf is senior onderzoeker bij het GION, Rijksuniversiteit Groningen. Zij is sinds 1989 coördinator van de VOCL cohortonderzoeken die bij het GION worden uitgevoerd.

Hans Kuyper is senior onderzoeker bij het GION. Hij is projectleider van diverse onderzoeken die deel uitmaken van de VOCL cohortstudies.

Miranda Lubbers is als aio werkzaam bij het GION. Zij verricht de data analyses van de VOCL cohortonderzoeken en zal op het nieuwe VOCL '99 cohort promoveren.

Correspondentieadres: M.P.C. van der Werf, GION, Rijksuniversiteit Groningen, Postbus 1286, 9701 BG Groningen, tel. 050-363665, email: m.p.c.van.der.werf@ppsw.rug.nl.

Abstract

Educational results and educational opportunities before and after the implementation of the common core curriculum

G. van der Werf, M. Lubbers & H. Kuyper. *Pedagogische Studiën*, 1999, 76, 273-288.

The data sets of the VOCL '89 and VOCL '89 cohorts are analysed to study the degree to which students' attainment in secondary education has been improved and the educational opportunities of disadvantaged students have been increased, since the implementation of the Common Core Curriculum. Overall it appears that the achievements in mathematics and Dutch language comprehension in the third grade and the attained educational position after five years of schooling stayed almost equal. Looking at educational tracks separately however shows a small improvement for (I)VBO and MAVO students and a small decline for HAVO and VWO students. Taking into account initial differences (intelligence, score on an entry test and recommendation of the primary school) between both cohorts, the results show a decrease of achievements and an increase of the educational positions. These changes are larger for the students in the higher tracks of secondary education, than in the lower tracks.

The educational opportunities have not improved since the implementation of the Common Core Curriculum. The differences in achievements and attained educational position related to gender, SES and ethnicity, after taking students' initial capacities into account, are equally large in both cohorts.

The conclusion is that the Common Core Curriculum did not yet have realized the aimed effects in terms of the criteria that are used in this study. But there also is no reason to be pessimistic about the consequences of the Common Core Curriculum for the results of secondary education.

The Annual Meeting of the American Educational Research Association, April, 19-23, 1999, Montreal. On the threshold of the 21st century

Inleiding (H. Dekkers, Katholieke Universiteit Nijmegen/ITS)

Ondanks de niet echt mediterrane ligging van de stad, en hoewel, tegen de gewoonte in, niet georganiseerd in de vrije paasweek, trok de AERA in Montreal een redelijk aantal bezoekers. En zoals andere jaren liepen de meningen over de kwaliteit weer zeer uiteen. Die verschillende waarderingsen worden mijns inziens vooral bepaald door twee achtergronden: AERA-ervaring en gehechtheid aan een mono-discipline. Wat het eerste betreft, er is inderdaad een overstelpende hoeveelheid bijdragen, in de vorm van papers, symposia, ronde tafels, 'invited addresses' et cetera, maar als men de zeer goed gestructureerde programmagids weet te gebruiken, is het goed mogelijk om, bijvoorbeeld op naam van de onderzoeker in combinatie met onderwerp, de interessante sessies te vinden. Dat laat onverlet dat het jammer is dat veel ruimte wordt ingenomen door presentaties van onderzoek dat kwalitatief en/of personeel nog in de kinderschoenen staat, terwijl door de beoordelaars soms hoogstaand onderzoek van minder bekende buitenlanders niet wordt geaccepteerd. Dat er daarnaast veel sessies voornamelijk worden bevolkt door praktijkonderzoekers die de nadruk meer op de praktijk dan op het onderzoek leggen is jammer, maar maakt van de andere kant het congres ook aardig voor niet-wetenschappers. In ieder geval is de afstand tussen onderzoek en praktijk in de Verenigde Staten heel wat kleiner dan in Nederland.

Gehechtheid aan een mono-discipline verklaart mogelijk ook teleurstelling over de AERA. Psychologen of sociologen die gewend

zijn op hun discipline-congressen uit heel toegespitste bijdragen te kunnen kiezen, vinden de veelheid van AERA-onderwerpen wat 'overdone'. Wie echter breed geïnteresseerd is, en de vele aspecten die voor het onderwijs en de onderwijsresearch van belang zijn erkent, en zoals gezegd het kaf van het koren kan scheiden, kan zich met gemak een dag of vier inhoudelijk vermaken. En, laten we wel wezen, met twee of drie interessante nieuwe invalshoeken naar huis gaan is een weekopbrengst die op de eigen werkplek niet iedere week wordt gehaald.

Bij de zorgvuldige selectie hoort ook dat men zich niet beperkt tot slechts een divisie of Special Interest Group (SIG). Vaak worden voor het vakgebied interessante bijdragen juist op andere plaatsen aangeboden dan voor ons voor de hand ligt. Desalniettemin wordt in deze rubriek verslag gedaan per divisie, waarbij een aantal verslaggevers zich terecht niet geremd heeft gevoeld om uitstapjes te maken naar andere divisies of relevante SIG's.

Divisie A: Administration (P. Sleegers, Katholieke Universiteit Nijmegen)

De divisie A van de AERA bevat onderzoek naar scholen en andere onderwijsinstellingen, leiderschap en/of contexten (community, political, professional, social, etc.). In 1997 (Chicago) kende deze divisie nog vier secties, te weten 1) 'organizational contexts; 2) 'school reform and leadership studies' 3) 'policy perspectives in administration' en 4) 'teaching and learning in educational administration'. In 1998 (San Diego) werden er nog maar 2 secties binnen deze divisie onderscheiden: 1) 'organizational studies' en 2) 'leadership studies'. In de bespreking van het AERA-congres van vorig jaar in Pedagogische Studien, werd deze reductie van vier naar twee secties betreurd, omdat het onvoldoende recht zou doen aan de verschillende perspectieven waarmee 'administration' en 'leadership' heden ten dage wordt bestudeerd. In Montreal was er helemaal geen sprake meer van een onderverdeling naar secties: de divisie was sectieloos. Er worden slechts inhoudelijke thema's onderscheiden die binnen de divisie aan bod kunnen komen. Het betreft de volgende vijf thema's: 1) 'leadership', 2) 'school improvement', 3) 'organiza-

tional dimensions of schools', 4) 'community linkages and collaboration' en 5) 'organizational learning and professional development'.

In de vorige twee bijdragen aan de kroniek over het AERA-congres zijn er een aantal trends gesignaleerd die herkenbaar waren in het onderzoek binnen divisie A. Het ging daarbij om de volgende trends: de relatie tussen theorie en praktijk; perspectieven op toekomstig onderzoek; 'accountability' en kwaliteitsbewaking; nieuwe conceptualisering van leiderschap (leadership as an organizational characteristic/leadership as a function/'teacher leaders') en het hanteren van diverse onderzoeksmethodieken en -benaderingen (naast sterk kwantitatief onderzoek, worden ook steeds meer kleinschalige kwalitatieve onderzoeken gerapporteerd). Hoewel deze trends ook in Montreal waren te herkennen, was er ook sprake van duidelijke accentverschillen in vergelijking met de voorgaande jaren. Zo was er dit jaar relatief weinig aandacht voor onderzoek waarbij het 'accountability' perspectief dominant is. En ik ben ook veel minder dan de voorgaande jaren geconfronteerd geweest met de vraag naar het praktisch nut van organisatie- en leiderschapsstudies voor de onderwijspraktijk. Misschien heb ik die sessies wel onbewust overgeslagen naar aanleiding van eerdere ervaringen. Veel meer dan de vorige jaren was er aandacht voor de relatie met de omgeving ('community linkages') en voor 'whole school reform' benaderingen. Daarnaast was er meer dan voorheen aandacht voor reflectie op onderzoek naar schoolorganisatie en leiderschap en perspectieven voor toekomstig onderzoek. Dit laatste was ook wel te verwachten gezien het centrale thema van het congres: 'On the Threshold of the 21st century: Challenges and Opportunities'.

In deze bijdrage aan de kroniek zal ik met name rapporteren over onderzoek naar twee trends die ook dit jaar weer duidelijk zichtbaar waren: nieuwe conceptualisering van leiderschap en perspectieven op toekomstig onderzoek.

Met betrekking tot de eerste trend zien we de laatste jaren dat er steeds meer aandacht is voor, en behoefte bestaat aan, een herconceptualisering van leiderschap. Dit heeft te maken met, en wordt beïnvloed door, de toegenomen

complexiteit van het sturing geven aan onderwijsorganisaties. Een goed voorbeeld van een dergelijke nieuwe benadering van schoolleiderschap wordt door Firestone (1996) gegeven. Firestone conceptualiseert leiderschap in termen van een verzameling van leiderschapsfuncties die verspreid in de organisatie moeten worden uitgevoerd door daarvoor in aanmerking komende organisatieleden, ongeacht of zij tot de formeel benoemde leiding horen. Het gaat er bij onderzoek naar de rol van leiderschap bij schoolverbetering volgens Firestone veel meer om de vraag welke leiderschaps- of sturingsfuncties worden uitgeoefend in de school dan om de vraag wat de schoolleider doet. Maar ook de conceptualisering van leiderschap als een kenmerk van de organisatie zoals door Pounder, Ogawa en Adams (1995) past in deze nieuwe lijn van denken. Door deze ontwikkeling is er binnen divisie A steeds meer aandacht voor onderzoek naar andere bronnen van leiderschap (substitutes for leadership), gedeeld leiderschap (shared leadership opportunities, teacher leadership) en de rol van teams bij schoolverbetering. In dit kader wil ik twee sessies kort bespreken. De eerste sessie waarin onderzoek naar deze thema's werd gepresenteerd betrof de sessie 'Organizational learning and the potential for schoolreform' (Chair: Kathryn Riley). Binnen deze sessie werd onderzoek gerapporteerd naar de relatie tussen leiderschap, organisatieleren en het verbeteren van de klassenpraktijk (Brown & Sheppard), de relatie tussen organisatieleren en het leren van studenten (Louis & Marks) en het leren van teams in organisaties (Leithwood & Steinbach). Centraal bij deze onderzoeken stond steeds de vraag naar de rol en invloed van sturingsmechanismen op het niveau van het team of de school als geheel. Samen met het gedrag van de schoolleider proberen deze onderzoekers meer inzicht te krijgen in de condities die een bijdrage leveren aan het vergroten van individuele en collectieve leerprocessen in en van de organisatie. Deze sessie maakte duidelijk dat een belangrijk begin is gemaakt, maar dat er tevens nog een heleboel conceptuele en empirische vraagstukken onbeantwoord blijven die opheldering behoeven in de nabije toekomst. Een belangrijk punt daarbij is de vraag naar de validiteit en betrouwbaarheid van het meten van begrippen als organisatieleren en teamleren.

De bijdrage van teams aan schoolverbetering stond ook centraal in de sessie over de rol van vaksecties bij schoolverbetering en -vernieuwing. Dit symposium getiteld 'High school departments and whole-school reform' bevatte bijdragen van Canadese (Lynne. M. Hannay), Amerikaanse (Judith Warren Little; Leslie Santee Siskin), Engelse (Nigel Bennett en Alma Harris) en Nederlandse onderzoekers (Jeroen Imants, Peter Slegers en Bob Witziers). Het symposium is een resultaat van de toegenomen belangstelling van onderzoekers naar de rol van vaksecties en vaksectieleiders. Bij alle bijdragen kwam er een centraal punt duidelijk naar voren, te weten de invloed van de vaksectie op de rol van docenten bij schoolvernieuwing en -verbetering. De onderzoeken laten zien dat structuur van vaksecties de ontwikkeling van de school kan blokkeren. Doordat docenten hun professionele identiteit vooral definiëren in termen van hun vak, kunnen zich zogenaamde processen van 'balkanization' voordoen. Hierdoor kunnen vaksecties interdisciplinaire samenwerking en gedeeld leiderschap bemoeilijken, met als gevolg dat er een spanningsveld ontstaat tussen het functioneren van de vaksectie als team en 'whole-school reform'. Gedeeld leiderschap, in deze onderzoeken gepositioneerd op het niveau van de vaksectieleider, hoeft dus niet een positief effect hebben op het functioneren van de schoolorganisatie. Dergelijke spanningsverhoudingen of dilemma's zouden in toekomstig onderzoek meer aandacht moeten krijgen.

Met betrekking tot de tweede trend, perspectieven op toekomstig onderzoek op het terrein van schoolleiderschap, wil ik twee sessies onder de aandacht brengen. De eerste betreft een 'invited lecture' van Joseph Murphy met als titel 'The quest for a center: Notes on the state of the profession of educational leadership'. In zijn lecture schetste Murphy de ontwikkeling van ons denken over schoolleiderschap in theorie en praktijk. Dominante visies op leiderschap die de revue passeerden, waren het scientific management, onderwijskundig gericht ('instructional') 'leadership' en het veranderingsgericht management. Murphy hield een pleidooi voor een herbezinning op het beroep van schoolleider door te benadrukken dat de kern van het beroep niet zozeer ligt in het alleen

beaccentueren van welke soorten van gedragingen en vaardigheden voor het effectief functioneren van schoolleiders van belang zijn. Een dergelijke functionalistische benadering is weliswaar van belang, maar is onvoldoende in staat om verschillen tussen effectieve schoolleiders verklaren. De kern van het schoolleiderschap zou in de ogen van Murphy vooral moeten zitten in het ontwikkelen van een visie op onderwijs, waarbij de relatie met de relevante gemeenschap waarbinnen dat onderwijs functioneert een duidelijke plaats krijgt. Volgens Murphy stelt dit nogal hoge eisen aan schoolleiders, eisen overigens die in zijn ogen ook gesteld moeten worden wil het beroep van schoolleider zich verder professionaliseren. Een vergelijkbare discussie was te bespeuren in een interactief symposium met de titel 'Educational leadership in an era of reform: Retrospects and prospects'. In dit symposium werd door vooraanstaande onderzoekers op het gebied van schoolleiders en schoolverbetering (Leithwood, Louis, Hallinger & Greenfield) gereflecteerd op de opbrengsten van onderzoek naar 'change management', 'transformational leadership', 'instructional leadership' and 'moral leadership'. De presentatie en de discussie waren wat ingehouden, voorzichtig en vooral beleefd. De conclusie was dat meerdere perspectieven op het verschijnsel van schoolleiderschap van belang zijn om meer zicht te krijgen op de rol van leiderschap bij het bevorderen van schooleffectiviteit en schoolverbetering. Evenals Murphy waren deze onderzoekers van mening dat leiderschap in scholen gefundeerd moet zijn op onderwijskundige waarden, opvattingen en visies. Jammer was dat de consequentie van deze gedachte voor toekomstig onderzoek naar schoolleiders niet echt goed werd bediscussieerd.

Tot slot het volgende. Vorig jaar eindigde de kroniek voor de divisie A met de constatering dat de kwaliteit van de presentaties en papers van het AERA-congres in San Diego, evenals in Chicago, sterk wisselde. Tevens werd daarbij opgemerkt dat de bijdrage van niet-Amerikaanse onderzoekers daarop een positieve uitzondering vormde. Helaas was deze teleurstellende trend ook in Montreal aanwezig. Een mooi voorbeeld daarvan was het symposium 'School reform: History and current issues'

met bijdragen van Murphy en Adams over kenmerken van de onderwijsveranderingen tussen 1980-2000, Shields over de relatie tussen ouders, docenten en schoolleiders bij vernieuwingen, Krasnow over praktijklessen van schoolvernieuwing en Klerks over de veranderingscapaciteit van scholen voor voortgezet onderwijs in Nederland. De bijdragen van met name Shields en Krasnow waren beide persoonlijke verhalen over ervaringen van werken met en aan schoolveranderingen. Deze persoonlijke ervaringen werden weliswaar overtuigend en vol compassie gebracht, maar voegden weinig toe aan het vergroten van onze reeds verworven inzichten over schoolverbetering. De gemeenplaatsen en platitudes volgden elkaar op tijdens de presentaties en maakten steeds meer ongestraft deel uit van de aanwezige ruimte. Enige vorm van kritische reflectie op de eigen bevindingen kenden de onderzoekers niet. De bijdrage van Klerks over haar onderzoek naar de veranderingscapaciteit was daarbij een welkome en waardevolle aanvulling, die helaas te weinig aandacht kreeg in de discussie na de presentatie. Op zo'n moment kun je de Nederlandse onderzoekscultuur en traditie weer eens op juiste waarde schatten en waarderen.

Referenties

- Firestone, W. A. (1996). Leadership: Roles or functions? In K. Leithwood et al. (Eds.). *International Handbook of Educational Leadership and Administration*. (pp. 395-418). Boston/Dordrecht: Kluwer Academic Publishers.
- Pounders, D.G., Ogawa, R.T. & Adams, E.A. (1995). Leadership as an organization-wide phenomenon: Its impact on school performance. *Educational Administration Quarterly*, 31, 564-588.

Divisie B: Curriculum (J.J.H. van den Akker, Universiteit Twente)

De a.s. millenniumwissel gaf binnen divisie B (Curriculum Studies) aanleiding tot diverse sessies met pretentieuze titels, waarbij het terugblikken voorop stond. En evenals vorig jaar stemde de inhoud van veel sessies binnen die divisie mij niet bijster vrolijk, met name door de neiging met elkaar vrijwel louter over zichzelf te spreken in plaats van over reële problemen buiten de cirkel van de 'geleerden' zelf.

Een treffend voorbeeld daarvan was de vice-presidential address van Janet Miller met als motto: *What's left in the field...A curriculum memoir*. Zij benadrukte dat er geen sprake kon (en zelfs mocht) zijn van een stabiel domein met gedeelde opvattingen. Controverses waren vruchtbaar (mits gepaard gaand met respect voor elkaars standpunten). Slechts 'situated and contextualized' benaderingen waren mogelijk. Haar betoog was dan ook zeer persoonlijk ingekleurd, vol met anekdotes en poëtische citaten. De kernboodschap luidde dat reconceptualisering van het curriculum-domein nodig was door middel van sociologische, politieke en culturele text-analyse. Benaderingen van curriculumtechnologische aard dienden sterk bekritiseerd en zelfs bestreden te worden. In de haast exclusieve aandacht voor elkaars (kritische) denkbeelden in geschriften leken verbanden met vraagstukken in curriculumbeleid en -praktijk in the 'real world' vrijwel onmogelijk.

Dat niet alle Amerikaanse curriculumcollega's gelukkig zijn met deze overmatig introverte neigingen, bleek niet alleen uit de wandelingen ('division B is still only talking to itself'), maar ook uit een zeer levendig forumdebat tussen een aantal kopstukken dat meer gericht was op uitdagingen voor de toekomst ('Curriculum studies at the threshold of the 21st century: Challenges and opportunities'). O.L. Davis opende met het oordeel dat veel universitaire curriculumtheoretici wel erg ver verwijderd zijn geraakt van de curriculumwerkers in de praktijk. Er is zijns inziens een overmaat aan kritische analyse van teksten (zonder serieuze onderlinge discussie) en te weinig bereidheid de 'handen uit de mouwen te steken' in samenwerking met de onderwijspraktijk. Het echte vuurwerk startte daarna bij Elliot Eisner die een reeks van 15 nijpende curriculumproblemen voorlegde waar curriculumonderzoekers zich over zouden moeten bekommeren. Een kleine selectie: spanning tussen de voortdurend toenemende curriculumstandaarden en de opgave voor docenten daar chocola van te maken; de gebrekkige samenhang tussen schoolvakken; de dikwijls slechte afstemming tussen innovatieve leerdoelen en assessment praktijken; de ervaring dat voor leerlingen het schoolbestaan meer in het teken staat van aanpassen en overleven dan van werkelijke moti-

vatie en stimulansen om te leren; het gegeven dat scholen zeer matige leeromgevingen zijn voor docenten; de overmatige verwarring over (kwaliteits-) ranglijsten tussen scholen. Maxine Greene volgde met een bevlogen, in literaire stijl gegoten voordracht over de relatie tussen school en maatschappij, met een accent op allerlei morele dilemma's. Madeleine Grumet kwam, in een zo mogelijk nog bevlogener stijl, met de boodschap dat het school-curriculum in volle dialoog met de gecompliceerde wereld moest staan. In de daarop volgende discussie ontstond een heftig debat waarin Eisner de beide dames eerst prees om hun welsprekendheid en belezenheid, maar ze vervolgens verweet veel te abstract en vrijblijvend over problemen in de onderwijsrealiteit te praten. Met name Grumet reageerde venijnig ('this was not girl talk') en kritiseerde Eisner om zijn 'technische-rationele' opvattingen. Dat was een verrassende typering, zeker in de ogen van Nederlandse curriculumspecialisten die Eisner nu niet bepaald als een 'hard-nosed' technoloog beschouwden. Tegen de gebruikelijke neiging tot convergentie in, kwam het tijdens het debat niet meer goed tussen de deelnemers, en dat maakte het wel zo boeiend.

Een zeer aardige sessie (buiten divisie B) voor een genuanceerd curriculumtechnoloog als schrijver dezes, had als titel: 'Increasing the impact of research through good "engineering"'. Onder leiding van Hugh Burkhardt (UK) werden levendige presentaties geleverd en discussies gevoerd door een aantal (Amerikaanse, Engelse en Australische) onderzoekers die zich – je zou kunnen zeggen: in de spirit van Eisners pleidooi – zeer concreet bemoeiden met pogingen tot verbeteringen van het wiskunde-onderwijs. Ze kwamen daarbij met sprekende voorbeelden van de worstelingen waarmee curriculumontwikkelaars te kampen hebben. Ze beklemtoonden de noodzaak tot systematische benaderingen met aandacht voor afstemming tussen beleid, curriculumontwikkeling (zowel standaarden als materialen), toets- en examenontwikkeling, en professionele ontwikkeling van docenten. Bovendien kwamen ze met prikkelende uitspraken over en voorbeelden van onderzoeksbenaderingen die dergelijk ontwikkelwerk kunnen ondersteunen. Opvallend in de gepropageerde onderzoeks-aanpak was het accent op de combinatie van

verbeeldingskracht en empirische feedback. Tevens werd benadrukt dat succesvol ontwerpen niet geschiedt door het toepassen van theorie maar door het (re)construeren van kennis door ontwerpers in interactie met practici. Aan het eind van de sessie werd de oprichting van een 'International Society for Design and Development in Education' geannonceerd. Wellicht leidt dat tot meer van dergelijke stimulerende symposia op toekomstige AERA-conferenties.

In bescheiden mate waren er ook nog enige Nederlandse bijdragen op curriculumterrein te bespeuren. Zo was er bijvoorbeeld een symposium over curriculumvernieuwing in Europa onder leiding van Wilmad Kuiper (UT), met daarin, naast bijdragen uit Engeland en Israël, een paper van Jan van den Akker (UT) over de curriculaire verwickelingen in het studiehuis, met een sterk accent op implementatieproblemen van docenten. Discussiant John Olson (Canada) was goed in vorm met zijn analyse van de spanning tussen mondiale trends en verschillen in culturele tradities tussen landen. Hij wees ook op de dikwijls weinig productieve polarisatie in vaak oppervlakkige curriculum-discussies. Een curriculaire getinte bijdrage over het studiehuis werd (in een andere sessie) ook geleverd door Wiel Veugelers en Henk Zijlstra (UvA), die vooral ingingen op de spanning tussen overheidsbeleid, schoolmanagement en autonomie van individuele leraren.

Tot slot een relativerende uitspraak die ik ergens tijdens de conferentie opving: in alle landen ter wereld blijkt men ontevreden over het onderwijs, maar overal weet men niet goed hoe dat te veranderen. Het meest algemene advies dat daaromtrent in allerlei toonaarden door velen geuit werd, luidt dat onderzoekers die aan onderwijsverbetering willen bijdragen er goed aan doen (meer) met practici samen te werken. Dat klinkt als een open deur, maar is voor sommigen wellicht toch nog niet zo vanzelfsprekend.

Divisie C: Leren en Instructie (Frank P.C.M. de Jong, Katholieke Universiteit Nijmegen/LSOP)

In de door mij bezochte presentaties in divisie C bleken veel van de actuele issues in de samenleving terug te komen: terugblik en vooruitblik in de context van de eeuwwisseling, de

enorme opkomst van Internet en WWW en tenslotte de paradigmatische veranderingen binnen leren en instructie als gevolg van veranderende epistemologische ideeën over kennis, leren en onderwijs. In de sessie over 'Hypermedia environments for thinking about teaching and learning' pakte mij de bijdrage van William Kealy waarin deze inging op het gebruik van 'knowledge maps' in Computer Supported Collaborative Learning Environments (CSCL) en hun mogelijke waarde voor de analyse voor CSCL-databases. In een volgende sessie werd meer ingegaan op de invloed van Internet op het leren en onderwijzen in het K12-onderwijs. Mark Windschitl benadrukte daarin dat het gebruik van het Internet nog steeds daar te zien is waar sprake is van vooruitstrevende klassensituaties. De aandacht zou volgens hem voorlopig ook vooral gericht moeten zijn op wat er qua leren op het Web gebeurt en hoe dat leren te bevorderen in plaats van alleen naar de leerresultaten te kijken. Daardoor zouden juist de unieke mogelijkheden van het Web als een gereedschap en medium om te leren beter geëxpliciteerd kunnen worden. Het laatste is hard nodig als we zien dat computers voornamelijk gebruikt worden voor het surfen op het net, tekstverwerking en spelletjes. In zijn onderzoek met betrekking tot de vraag waarom sommige leerkrachten wel en anderen geen gebruik maken van ICT vond Henry Becker de volgende 8 kritische factoren: 1) de aanwezigheid van 'high speed' Internet toegang in het eigen klaslokaal; 2) het hebben en praktiseren van constructivistische ideeën met betrekking tot leren en instructie in de klas; 3) zichzelf computervaardig vinden; 4) het hebben van informeel overleg over ICT met andere leerkrachten in de school; 5) de betrokkenheid bij activiteiten met betrekking tot professioneel leiderschap; 6) deelname aan formele professionaliteit ontwikkeling op ICT-gebied; 7) de leeftijd, hoe jonger een leerkracht des te eerder hij/zij ICT in de klas gebruikt en tenslotte 8) het al of niet een wiskundeleraar zijn. Zogenaamde Web-leerkrachten bleken zich meer dan anderen te oriënteren op constructivistische didactiek (zie ook <http://www.crito.uci.edu/tlc/html/findings.html>). John Mergendoller concludeerde in dezelfde sessie dat een aantal ICT-projecten problematisch zijn omdat leerlingen zowel

declaratieve als metacognitieve kennis-deficiënties hebben in het www-gebruik. Het niet kunnen omgaan met ongestructureerde en moeilijke routes beïnvloedt de betrokkenheid van leerlingen. Ook aspecten als hoe leerlingen te groeperen en hun individuele of groepsinbreng te accrediteren vormen problemen die de inzet van ICT belemmeren. Als oplossing voor de problemen suggereerde hij dat in projecten meer voorbeeldmodellen (video's en werkwijzen van vorige groepen) gegeven zouden kunnen worden met betrekking tot zelfreflectie op het leerproces, strategische planning en metacognitieve sturing. Ook zou meer aandacht aan het proces besteed kunnen worden door met leerlingen aan vragen te werken als bijvoorbeeld: Wat hebben we geleerd?; Hoe hebben we het gedaan?; Hoe gaan we verder en hoe kan het beter? Duidelijkere leerdoelen, inhouden, vaardigheden, samenwerking, groepsdiscussie en eindtoetsing zouden ook bijdragen aan een succesvol www-gebruik in de klas. Thomas Reeves sloot aan bij John's suggestie om leerlingen meer te betrekken in reflecterende activiteiten en het met elkaar delen van kennis. Zoals zo mooi in de reclameslogan 'who is afraid of sharing knowledge' naar voren komt. Reeves stelde verder groot belang in het stimuleren van zeggenschap en verantwoordelijkheid van leerlingen over de leertaak en het creëren van activiteiten om kennis op te bouwen. Barry Fishman concludeerde in zijn presentatie dat het gebruik van www tot meer ongestructureerde leersituaties leidt, waarbij een meer projectmatig georiënteerde didactische benadering beter zou passen. Een benadering waarin leerlingen meer met elkaar overleggen en waarin het Web als een informatiebron gezien wordt. Maar ook hij concludeerde dat het doelgericht zoeken op het Web voor veel leerlingen nog problemen oplevert.

Niet geheel ontevreden over de vorige dag begaf ik mij in de frisse vrieswind en voorjaarszon van Montreal de volgende ochtend naar het symposium dat Michael Jacobson had georganiseerd rond de thematiek van 'complexity and complex systems'. Zoals gebruikelijk vinden goede sessies altijd plaats in veel te kleine zaaltjes. Niet alleen de drukte maar ook de presentatie van Jacobson, Chi en Resnick, gericht op de opkomende perspectieven met betrekking tot cognitie, leren en didactiek deed

me de ochtendkou snel vergeten. Jacobson de-
finieerde het leren en begrijpen van weten-
schappelijke kennis als het reorganiseren van
intuïties waarbij men vaak inconsistenties met
alledaagse ervaring moet overbruggen. Miche-
lene Chi stelde dat studenten wetenschappe-
lijke kennis vaak niet vatten door het gemis aan
relevante wetenschappelijke denkvaardighe-
den, leervaardigheden, en metacognitief bewust-
zijn hoe de wetenschap werkt. Het karakter van
wetenschappelijke inhoud is voor studenten
vaak moeilijk vanwege het gecompliceerde,
complexe, abstracte en het dynamische van
de kennis, de misconcepties daaromtrent van
studenten en hun verkeerde attributies. Het
probleem zit hem vooral in het onevenredige
tussen de feiten en het proces van het weten-
schappelijk kennen. De afzonderlijke proces-
sen en feiten zijn wel begrijpelijk maar het zich
ontwikkende mechanisme versus het oor-
zakelijke mechanisme in dit denken is voor hen
problematisch. Mitchel Resnick sloot in zijn
presentatie hierbij aan door in te gaan op ge-
decentraliseerd denken als tegenpool van de
gecentraliseerde 'mind-set'. Daarbij spelen
nieuwe concepties van 'control' en oorza-
kelijkheid (immergent mechanism als een andere
vorm van oorzakelijkheid). Het gebruik van
computers faciliteert de nodige epistemologi-
sche verandering in het denken over het onder-
wijsleerproces. Vooral de mogelijkheid dat
daardoor de lerende meer in de gelegenheid is
zijn eigen model te creëren door actieve parti-
cipatie (indaling in het kennissysteem door
deelname aan dat systeem in plaats van alleen
ernaar te kijken). De sociale component van
leren wordt versterkt door het delen van inzicht
met anderen, maar ook de persoonlijke compo-
nent wordt versterkt als met het onderwerp
dichter bij alledaagse ervaringen wordt geble-
ven. De Bees-light (participatory simulation)
experimenten van het MIT, zoals de versprei-
ding van infectieziektes, zijn wat dat betreft
exemplarisch voor hoe computer-gebaseerde
simulaties uit de computer zijn te halen en in
dagelijkse ervaringsituaties geplaatst kunnen
worden. Ze stellen de lerende in de gelegenheid
directe denkervaring op te doen in en buiten
dynamische systemen. Daarbij blijken didacti-
sche elementen als leren van eigen ervaringen,
stimuleren van onderzoekmatige onderwijs-
situaties en collaboratieve theorievorming

belangrijk te zijn. Tevens is het belangrijk om
in dergelijke leersituaties gelijkwaardige maar
niet-identieke ervaringen te stimuleren en
coherente regels die consistent zijn met de
wetenschappelijke wereld te hanteren. Onder
de indruk van de presentaties plaatste de dis-
cussie het symposium weer met beide voeten
op de grond door tot de slotconclusie te komen
dat professionals niet altijd op het niveau van
micro kansmodellen denken. Ook zij gebruiken
vaak het niveau dat werkt, maar niet altijd het
meest verklarende is. Je hebt niet altijd de rela-
tiviteitstheorie nodig. Men kan zich afvragen
of het echt kritiek is dat leerlingen en studenten
niet altijd het diepste niveau begrijpen. Het is
wel een belangrijk niveau maar niet het enige.

In de 'state of the art'-sessie van divisie C
kwamen in snel tempo alle zeven aandachts-
gebieden van de divisie aan het woord over
'where have we been and where are we going
to'. De aftrap kwam van Patricia Alexander
voor het aandachtsgebied 'Language and lear-
ning'. In haar ogen is het duidelijk dat informa-
tie, bevolking en media kwantitatief enorm zijn
toegenomen. Niet alleen de kwantitatieve toe-
name heeft een invloed op het leren van taal
gehad maar ook de toename in de diversiteit
aan informatie, bevolking, het met elkaar in
verbinding staan en de manier van kennistrans-
missie. De mentale hard- en software waarmee
we moeten werken in relatie tot al die
veranderingen zijn daarentegen nauwelijks
veranderd. In het overbruggen van die kloof zie
je steeds thema's iteratief terugkomen zoals
bijvoorbeeld fonetische leesmethoden, het cog-
nitieve versus het affectieve, collaboratief en
sociaal leren versus het individuele en zelfstan-
dig leren, directe instructie of ingebedde strate-
gie-instructie, proces of product etc. Ondanks
de paradoxale spanningen die er in de verande-
ringen zitten zoals bijvoorbeeld de rijkdom aan
informatie en tegelijkertijd de schaarste aan
betekenisvolle kennis, noodzaken de verande-
ringen tot verschillende herbezinningen met
betrekking tot taal en leren. De begrippen tekst
en leren van tekst moeten bijvoorbeeld worden
verbreed als zijnde niet alleen meer lineair en te
vinden in boeken. Er is niet alleen een affectie-
ve en cognitieve maar ook een meer holistische
kijk op leren en de lerende nodig. Er tevens
ontbreekt een ontwikkelingstheoretisch per-

spectief ten aanzien van leren lezen en schrijven. Nu gaat de meeste aandacht uit naar het aanvankelijk leren lezen en schrijven. Meer aandacht voor de epistemologie en ontologie van kennis is eveneens noodzakelijk. Daarbij dient fundamenteel te worden ingegaan op het leren binnen niet-lineaire en niet-traditionele tekstomgevingen.

De bijdrage van William Schmidt in relatie tot wiskundeonderzoek richtte zich, naast erkenning dat er veel goed onderzoek is gedaan, vooral op hoe verder te komen op een aantal gebieden. Zijn aandacht ging daarbij vooral uit naar de beperkte waarde van allerlei vergelijkende studies van de wiskundige bekwaamheid van leerlingen in diverse delen van de wereld. Het gehanteerde instrumentarium doet volgens hem geen recht aan de werkelijkheid. Er dienen diagnostische instrumenten te komen die betrouwbare inzichten geven in de aard van de instructie en het curriculum in de klas. Daarbij is niet alleen van belang dat ze inhouden dekken maar vooral de coherentie ertussen. Er moet meer aandacht zijn voor de verschillen in culturen, didactische benaderingen, curricula en mogelijkheden die er voor kinderen zijn om te leren rekenen.

Peter Hewson vertegenwoordigde het 'science'-gebied en pleitte voor het schenken van meer aandacht aan de problematiek van het begrijpen van de science-kennis: Wat weten we?; Wat kunnen we doen met deze kennis?; Hoe weten we dat we weten?; Hoe veranderen we hetgeen we weten?. Behalve dat leerlingen de science-kennis dienen te begrijpen is het ook belangrijk dat zij inzicht krijgen in het proces van wetenschappelijk onderzoek. Als veelbelovende didactische benaderingen voor de toekomst zag hij het argumenteren, modelleren van het ontwikkelen en reviseren van kennis, leerling-gecentreerde instructie waarin zij leren naar ieders perspectieven te luisteren, redenen ter argumentatie van hun eigen ideeën te geven, hun ideeën metacognitief te analyseren op coherentie en consistentie en tot consensus weten te komen over de meest belovende perspectieven. De kracht in deze didactiek ligt in het integreren van het actief wetenschappelijke ingaan op kennis bij de instructie van 'science'-kennis.

Met betrekking tot de sectie leeromgevingen startte Ricky Goldman na aandacht te heb-

ben gevraagd voor de dood van Jan Hawkins met de vraag 'Wat is geen leeromgeving?'. Vervolgens ging zij diep in op het constructivisme vanuit de overtuiging dat de lerende in staat is culturele betekenis te creëren. De technologie moet daarbij meer gelijkheid en gelegenheden bieden voor alle leerlingen. Theorieën die onderzoek in een groter verband plaatsen en clusters die 'co-constructed' en 'co-contextualized' zijn, leiden tot een betere leefbaarheid van de samenleving. Leren en onderwijzen zijn dan geen twee gescheiden functies meer, maar onderwijs wordt dan gerealiseerd door ontwerpteams van lerenden, leerkrachten en onderzoekers die werken als een gemeenschap waarin inzichten en kennis worden gedeeld. Co-vis is bijvoorbeeld zo'n voorbeeld. In deze project-gebaseerde leeromgeving is de sociale constructie van kennis gesitueerd, collaboratief, direct diepgaand en multidisciplinair. 'Communities of practices' leren niet alleen wat al bekend is maar creëren ook nieuwe inzichten in de omringende wereld. De uitdaging voor onderzoekers en ontwerpers van leeromgevingen ligt hem in het realiseren van generieke platforms voor het faciliteren van onderzoeksgemeenschappen. Daarin zijn de kennisbemiddelaars de lerenden in plaats van de leerkrachten met hun voorgestructureerde pakketjes kennis.

Wilbert McKeachie volgde een historische benadering om verleden, heden en toekomst van de sectie 'Cognitive, social development and motivational processes' te bespreken. Werkend vanuit het heersende behaviorisme in '46 toen hij als leerkracht begon, gaande via de benaderingen als 'life span', context, transition, social psychology en motives, belandde hij bij hedendaagse benaderingen als 'hot cognitions' (motivations), beliefs en feelings, metacognitie en zelf-regulerend leren en will-power. Zaken zijn complexer geworden en het onderzoek houdt zich meer en meer bezig met complexiteit en interactie. In dat wetenschappelijke gebeuren komen en gaan theorieën die vurige aanhangers zullen kennen maar McKeachie ziet niet zo snel dat er zich een 'Grand Theory' uit ontwikkelt.

Michele Foster besprak de sectie 'learning, teaching, and assessment in school and community settings'. Zij ging daarbij in op de aspecten van etnische minderheidsgroepen in

relatie tot school- en thuissituaties en de rol van de culturele context. In relatie tot de sectie Technology and Education sprak John Bransford de verwachting uit dat technologie flink zal veranderen wat betreft verschijningsvorm en het leren dat deze zal bevorderen. Technologie wordt steeds kleiner en meer geïntegreerd in onze leermiddelen en leersituaties. Ten tweede zullen er meer data komen die de positieve invloed van technologie op het leren laten zien. Vooral omdat technologen en onderwijskundigen meer en meer zullen samenwerken in het creëren van creatieve onderwijsinterventies die effect hebben. Daardoor ontstaat systeemintegratie die niet alleen krachtige software omvat maar een implementatie en inbedding in een compleet onderwijssysteem met een ontwerp van leeromgevingen die zowel toetsing als kennisbronnen omvatten welke leerlingen nodig hebben om hun leerproces beter te realiseren. Nieuwe vormen van evaluatie waarin niet alleen reproductie getoetst wordt of wat men zonder hulp naar andere situaties kan transformeren, maar die inzicht geven in wat men effectief kan leren. Ten derde komen er belangrijke veranderingen in hoe we werken. Meer collaboratief, afstandleren en meer modulaire software die te gebruiken is door de individuele leerling dragen daar aan bij.

Ook in de resterende AREA-dagen komen veel aspecten terug: Hoe leerinhouden en kennis meer toegankelijk te maken; Het denken en leren meer zichtbaar krijgen; Leerlingen en studenten helpen van elkaar en anderen te leren; Stimuleren van 'life-long'-leren; Onderzoekend bezig zijn met authentieke vragen voortkomend uit de ervaringen van leerlingen als de centrale strategie voor onderwijzen; Van leerling-gecentreerd naar leerling-gesitueerd leren; Leerling gegenereerde discussies. Bij alle vraagstukken draait het steeds om centrale vragen als: Hoe (technologische) innovaties veranderingen in de cultuur van het klassikaal onderwijs stimuleren? Hoe veranderen als gevolg van (technologische) innovaties de interactiepatronen tussen student en docent? Welke nieuwe uitdaging(en), dynamiek wordt er gerealiseerd als gevolg van de innovaties? Voor mijzelf kwamen veel vragen en onderwerpen bijeen in de door John van Aalst georganiseerde sessie 'Knowledge building in diverse contexts: Challenges for teaching,

design en teacher education'. Niet alleen vanwege het sterk internationale karakter waar ik zelf ook aan bij mocht dragen, maar vooral doordat de sessie en inhoud van informatie naar collaboratie en terug ging. Tijdens de AERA-presentaties viel mij op dat er verschillende keren naar Europees en in het bijzonder Nederlands onderwijskundig onderzoek werd verwezen. Na afloop van de AERA heb ik voor mijzelf geconcludeerd dat er in het nieuwe millennium een pedagogisch didactische, epistemologische en technologische agenda ligt te wachten. Agenda's waar leerkrachten en onderzoekers samen een centrale rol spelen in het identificeren van problemen en analyseren van data van de in ontwikkeling zijnde technologische leeromgevingen en de integratie daarvan in de dagelijkse ervaringen. Het belangrijkste doel blijft daarbij de lerenden te helpen diepere niveaus van inzicht te bereiken.

Divisie D: Measurement and Research Methodology (C.A.W Glas, R.R Meijer, Universiteit Twente)

Dit jaar was het thema van de AERA 'On the threshold of the 21st Century: Challenges and Opportunities'. Als we met het programmaboekje van de AERA en de 'Annual meeting of the NCME (National Council on Measurement in Education)' terugkijken op het laatste jaar van de 20ste eeuw, valt op dat item response theorie (IRT) het gebied van onderwijskundig meten bepaalt. Presentaties over klassieke testtheorie en generaliseerbaarheidstheorie zijn zeldzaam geworden. Geheel tegen de verwachting is ook het aantal bijdragen over multilevel modellen erg klein.

Als we de onderwerpen van symposia van divisie D van de AERA en van de NCME inventariseren, valt op dat computer-adaptief toetsen (CAT) de belangrijkste uitdaging voor de 21ste eeuw lijkt te worden. Verder lijkt ook 'differential item functioning (DIF)' een onderwerp waarover de discussie in de 21ste eeuw nog wel even door zal gaan. Voordat we de presentaties over deze twee onderwerpen wat nader bekijken, eerst iets over een van de weinige echt nieuwe dingen die te beluisteren waren.

Het evalueren van beoordelaar-betrouwbaarheid is sinds jaar en dag een van onderwerpen die nog steeds binnen het kader van de

generaliseerbaarheidstheorie wordt afgedaan. Binnen de IRT is er tot nog toe geen acceptabele oplossing voor dit probleem; de enige bijdrage was tot voor kort het multifacet model (FACETS) van Wright en Linacre. Tegen deze benadering zijn echter fundamentele bezwaren aan te voeren, die erop neerkomen dat er geen rekening wordt gehouden met de afhankelijkheid van de oordelen van de beoordelaars van het antwoord van de respondent. De oplossing voor dit probleem is onafhankelijk door drie groepen onderzoekers gevonden: Wilson en Hoskens, Junker en Patz en Bock en Muraki. Hun oplossing is dezelfde: een multilevel IRT model waaraan parameters zijn toegevoegd voor de interactie tussen beoordelaars en items. De drie groepen gaven een gezamenlijk symposium waar ze de overeenkomst in hun modellen ruiterlijk toegaven en vervolgens ingingen op de verschillen tussen hun benaderingen: de specifieke methoden die ze gebruikten om de parameters in het model te schatten.

Een snelle telling leert dat er tenminste vijf sessies aan differential item functioning (DIF) besteed werden, iets minder dan vorig jaar, maar nog steeds een respectabel aantal. De interessantste sessie was een symposium onder de titel 'Has DIF made a difference', waar coryfeeën Holland, Swaminathan en Zwick zowel achteruit als vooruit keken. Holland en Thayer introduceerden in 1982 de Mantel-Haenszel-procedure voor het opsporen van DIF. Holland stelde dat hij ten tijde van de introductie van de procedure zeker beseftte dat er geavanceerdere, op IRT gebaseerde procedures denkbaar waren, maar dat hij een procedure wilde ontwikkelen die zeer eenvoudig en goedkoop (10 cent per item) zou zijn. Door deze keuze is het evalueren van DIF tot het standaardrepertoire van de psychometricus gaan behoren. Swaminathan ging juist op die meer geavanceerde procedures in, en hij presenteerde een logistische regressie-procedure die het mogelijk moet maken inzicht te krijgen in de oorzaken van DIF. Zwick, tenslotte, gaf een overzicht van inhoudelijk onderzoek naar de oorzaken van DIF, en ze kwam tot de conclusie dat we nog niet veel weten dat het niveau van het evidente (voetbal-items voor meisjes) overstijgt.

Wat betreft CAT kan gezegd worden dat de praktijk weerbarstig is. Ongeveer 25 procent

van alle presentaties binnen de NCME ging dit jaar over CAT. Het construeren van een toets op basis van een geschatte vaardigheid heeft sinds de eerste ontwikkelingen begin jaren 70 duidelijk aan aantrekkelijkheid gewonnen, mede door de ontwikkeling van snelle computers. Efficiëntie en klantvriendelijkheid zijn de voornaamste voordelen van het adaptief testen. Verschillende onderzoeksafdelingen zoals die van Educational Testing Service (ETS), American College Testing (ACT), Defense Manpower Center (DMC) alsmede de afdeling Onderwijskundige Meetmethoden en Data-Analyse (OMD) van de Universiteit Twente presenteerden hun resultaten van onderzoek. Om adaptief te kunnen testen wordt gebruik gemaakt van een bank met gekalibreerde items. Een van de lastige problemen is dat deze itembank veelal bestaat uit een relatief beperkt aantal items (300-500) waarvan een gedeelte niet wordt gebruikt omdat deze niet passen bij de vaardigheden die het meeste voorkomen. Als gevolg hiervan is een itembank gemakkelijk te 'kraken' door bijvoorbeeld een test meerdere malen te maken of door met een groep respondenten af te spreken ieder een groepje items te memoriseren. Remedies voor deze problemen werden op een aantal presentaties gepresenteerd. Zo kan men het algoritme dat de items selecteert zo specificeren dat de items die al vele malen zijn gebruikt niet in de onderhavige test worden opgenomen. Andere methoden houden zich bezig met het opsporen van scorepatronen die het resultaat zijn van voorkennis van items.

Opmerkelijk was dat er een trend te bespeuren viel naar het niet psychometrisch optimaal kiezen van de items uit een itembank, maar terug te keren naar eerder voorgestelde simpelere vormen van adaptieve testen zoals two-stage testing. Niet hetgeen psychometrisch optimaal is bepaald is dan de keuze van het item maar hetgeen inhoudelijk optimaal is. DMC presenteerde papers over on-line item kalibratie, grafische item-analyse, het minimaliseren van item exposure en toepassingen van multidimensioneel adaptief testen. In het ACT symposium was de aandacht gericht op het construeren van een CAT die in hoge mate vergelijkbaar was met papier-en-potlood tests, waarbij de aandacht vooral gericht was op het garanderen van multidimensionaliteit en waar-

bij de inhoud werd gespecificeerd door middel van doel informatie functies. Op het symposium van de afdeling OMD (Universiteit Twente) kwamen onderwerpen aan bod als het equaten van testcores van adaptieve tests met papier-en-potlood tests, methoden voor het vergelijken van item kalibratie resultaten van pre-test en on-line afnamen en het bestuderen van optimale beslissingsregels om master's en niet master's te classificeren. Tot slot was er het panel dat de inhoud van de papers besprak en richting wilde geven aan het vervolgonderzoek. Interessant was de bijdrage van ETS waarbij het adaptief testen per computer tot nu toe heeft geleid tot aanzienlijke problemen met name door het eerder geschetste probleem van het bekend worden van items, waardoor extra kosten moesten worden gemaakt voor het schrijven van items.

Een interessant geval dat wederom duidelijk maakte dat de eersten niet altijd de succesvolsten zijn: anderen kunnen leren van de fouten en problemen die pioniers ondervinden en hun product tegen geringere kosten op de markt zetten.

Divisie G: Social Context of Education (D. Uerz en H. Dekkers, ITS, Katholieke Universiteit Nijmegen)

In divisie G staat van oudsher de relatie tussen het onderwijs en haar sociale, economische en culturele context centraal. In een groot aantal lezingen worden de onderwijsloopbanen van leerlingen in relatie tot hun sociaal-economische achtergrond, sekse en etnische herkomst besproken. Evenals in voorgaande conferenties bleef daarbij een integrale benadering van het onderwerp veelal achterwege. Wanneer in de diverse sessies de invloed van de genoemde achtergrondkenmerken van leerlingen al in onderlinge samenhang werd bestudeerd, bleef dit beperkt tot een beschrijving van de onderwijspositie van zeer specifieke groepen. Zoals bijvoorbeeld in het door de organisatie geformuleerde divisie-highlight 'The invisibility of race in classrooms: The experiences of middle class black students' of het paper 'A portrait of a high achieving, low-income, minority, first-generation college student'. In het merendeel van de sessies bleef de combinatie van achtergrondfactoren echter achterwege en werden

onderwijsloopbanen van de afzonderlijke groepen leerlingen beschreven ('Examining literacy practices for Latino students; new challenges', 'Creating partnerships to help Alaska native students succeed in two worlds'). In beide gevallen is het moeilijk om een vergelijking te maken met de problematiek van culturele minderheden in Nederland en de rest van Europa. Methodologisch ging daarbij, zoals ook uit de titels van de genoemde papers blijkt, de voorkeur uit naar het kwalitatieve onderzoek; met name case-studies, biografisch onderzoek en 'single case' onderzoek kwamen binnen deze divisie aan bod. Verschillende bijdragen beperkten zich zelfs tot het meedelen van 'best practices', of het bieden van een reflectie op eigen ervaringen met compenserende programma's voor achterstandsleerlingen. Voor meer generaliseerbare resultaten en kwalitatief-empirisch onderzoek op het gebied van onderwijs en ongelijkheid moest daarom, evenals in de voorgaande jaren, voornamelijk naar andere divisies en Special Interest Groups (SIG's) worden uitgeweken.

Een vaak terugkerend thema dit jaar was de relatie tussen school en gezin. Zowel binnen divisie G als in de SIG Sociology of Education werd in een aantal papers ingegaan op de rol van het in het gezin aanwezige sociaal kapitaal als mogelijke verklaring van sociaal-economische verschillen in schoolloopbanen. Opmerkelijk was dat in alle gevallen de betrokkenheid van ouders bij het onderwijs van hun kind als een belangrijke maatstaf voor het aanwezige sociaal kapitaal werd beschouwd. De uiteindelijke definitie van deze indicator (ouderlijke betrokkenheid) bleek echter per paper te verschillen. Zo werd in het paper 'The hidden component of college choice: parent willingness to support education' (Chang, Stevenson & Schneider) ouderlijke betrokkenheid gemeenten als de bereidheid van ouders om een financiële bijdrage te leveren aan het vervolgonderwijs van hun kind. Zoals te verwachten bleek dit positief samen te hangen met de uiteindelijke keuze voor vervolgonderwijs. In het paper van Haghighat, 'What are the compositional effects of school socio-economic status on students' academic achievement', werd juist ingegaan op de betrokkenheid van de ouders bij activiteiten binnen de school. Hij stelt vast dat

wanneer scholen ouders meer bij het onderwijs betrekken dit een positief effect heeft op de prestaties van leerlingen, waardoor sociaal-economische verschillen afnemen. Zeer interessant was de bijdrage 'Parental involvement and parental investment in an Asian educational system' van Sui-chu Ho, waarin werd ingegaan op de relatie tussen culturele en economische hulpbronnen in het gezin (parental investment), de betrokkenheid van ouders bij het onderwijs (parental involvement) en het zelfvertrouwen van leerlingen. De ouderlijke betrokkenheid werd in dit paper gemeten aan de hand van twee dimensies: betrokkenheid binnen het gezin (zoals het helpen met huiswerk, voorlezen, etc.) en betrokkenheid naar de school (bijwonen ouderavonden, contacten met leerkrachten en vrijwilligerswerk binnen de school). Geconcludeerd werd dat economische en culturele hulpbronnen, in overeenstemming met de theorie van Bourdieu, het zelfvertrouwen van leerlingen weliswaar beïnvloeden, maar dat de invloed van het sociaal kapitaal in het gezin, uitgedrukt in de ouderlijke betrokkenheid bij het onderwijs, een nog grotere bijdrage hieraan levert. Helaas werd in dit paper niet ingegaan op de relatie hiervan met de uiteindelijke prestaties van leerlingen, maar het biedt desondanks aanknopingspunten voor verder onderzoek waarin de samenhang tussen het cultureel en economisch kapitaal van Bourdieu met het sociaal kapitaal van Coleman verder in kaart worden gebracht.

Ook in divisie L 'Educational policy and politics' werd aandacht besteed aan sociaal kapitaal, zij het vanuit een andere invalshoek. In het symposium 'Social capital in six high schools: theoretical and practical implications' werd ingegaan op het creëren van sociaal kapitaal in het voortgezet onderwijs als beleidsinstrument om kansen van achterstandsleerlingen te vergroten. Zeer interessant was de bijdrage van Helen Marks waarin werd ingegaan op de vraag in hoeverre scholen in staat zijn om het sociaal kapitaal binnen de school te vergroten. Zij stelt dat de basis hiervoor is gelegen in een systeem van normen en regels binnen de school, dat gedragen wordt door zowel de leerkrachten als de ouders. Socialisatie door oudere groepsgenoten zorgt ervoor dat dit stelsel van normen en regels geïnternaliseerd wordt in

nieuwe leerlingen, waardoor het eigen normatieve systeem van de school versterkt wordt. Het sociaal kapitaal dat op deze manier wordt gecreëerd wordt dan een onderdeel van de schoolorganisatie en maakt de school tot een gemeenschap waarin leerlingen worden ondersteund en leren wordt gestimuleerd. Spreekster Valerie Lee maakte echter duidelijk dat maatregelen om het sociaal kapitaal te stimuleren tegelijkertijd ook negatieve gevolgen kunnen hebben. In haar bijdrage besprak zij hoe keuzevrijheid in het onderwijs enerzijds het sociaal kapitaal binnen een school kan stimuleren, maar tegelijkertijd onderwijsongelijkheid kan vergroten. Vastgesteld werd dat wanneer ouders vrij zijn om een school voor hun kind te kiezen dit weliswaar leidt tot een hogere betrokkenheid van ouders bij de school en daarmee het sociaal kapitaal binnen een school stimuleert, maar tegelijkertijd ontstaat zo echter ook sociale stratificatie; de betere leerlingen komen terecht op de zogenaamde 'schools of choice', terwijl leerlingen met minder hoge prestaties worden afgewezen en doorstromen naar de public schools. Hetzelfde geldt voor de keuze van richtingen binnen een school (schools-within-schools). Wanneer leerlingen binnen scholen zelf mogen bepalen naar welke 'sub-unit' zij willen gaan bevordert dit enerzijds de motivatie en betrokkenheid van de leerling. Anderzijds werkt dit segregatie in de hand; zo worden met name in de zogenaamde 'special programs' alleen de betere leerlingen geselecteerd. Geconcludeerd werd dan ook dat wanneer wordt gesproken over het vergroten van de keuzevrijheid van ouders en leerlingen een afweging moet worden gemaakt tussen zowel de positieve als de negatieve gevolgen hiervan.

Opmerkelijk was dat wederom een groot aantal studies binnen de verschillende divisies gebruik maakten van veelal dezelfde nationale Amerikaanse bestanden (NAEP, NELS). Cohorten die niet in eerste instantie in het leven zijn geroepen om achterstandenbeleid te evalueren (zoals in Nederland het geval is), maar waarin door oversampling van bepaalde groepen en districten toch analyses mogelijk zijn om die achterstanden te bepalen. Zeer interessant was in dit kader, het symposium 'Using national data sets to inform policy: influences for equity'. In een discussie tussen statistici van

het NCES, wetenschappelijke onderzoekers en beleidsmakers werden de mogelijkheden en belemmeringen van de huidige bestanden besproken; duidelijk werd dat een betere afstemming tussen beleid en dataverzameling noodzakelijk is.

Het door de SIG Advanced Studies of National Databases georganiseerde symposium over de identificatie en ervaringen van high school drop-outs bood de mogelijkheid tot internationale vergelijkingen. Duidelijk werd dat ook in andere werelddelen (VS, Australië) wordt geworsteld met het gebrek aan eenduidige definities en meetmethoden voor voortijdig schoolverlaten, zoals dat in Europa het geval is. Dat laat onverlet dat resultaten werden gepresenteerd die aanleiding gaven tot interessante discussies, en die relevant zijn voor verder onderzoek.

Tenslotte verdient binnen het thema Social Context of Education de Invited Presentation 'The significance of minority status' van John Ogbu vermelding. Met name de resultaten van zijn vergelijkende studie naar al dan niet succesvolle minderheden in verschillende immigratielanden, en de achtergronden en verklaringen daarvoor (bijvoorbeeld in settler vs non settler societies, en in al dan niet vrijwillige redenen van immigratie) bieden een goed kader voor nationale studies naar (onderwijs en) minderheden.

Divisie H: School evaluation and program development (R.J. Bosker, Universiteit Twente)

Deze divisie biedt binnen de AERA de ontmoetingsplaats voor degenen die zich bezig houden met de ontwikkeling, implementatie en evaluatie van interventieprogramma's op het niveau van de school, het district of de staat. Dat leidt ertoe dat, zoals Creemers zes jaar geleden in de AERA-kroniek vaststelde, men vooral wetenschappelijk opgeleide practici en praktisch georiënteerde wetenschappers bij de symposia en papersessies van deze divisie kan tegenkomen. Die praktisch georiënteerde wetenschappers zijn dan meestal evaluatoren, de wetenschappelijk opgeleide practici programma-ontwikkelaars. Wie echt geïnteresseerd is in de ontwikkelingen op het gebied van evaluatie doet er verstandiger aan zijn inspira-

tie elders te zoeken, want de 'state of the art' sessie voor divisie H, met de veelbelovende titel 'future directions for educational program evaluation' bood weinig nieuws. Dat elders is overigens het eveneens niet bepaald kleinschalige congres van de American Evaluation Association. Binnen de divisie komen ook praktische zaken aan de orde, variërend van onderwerpen als wat te doen met het 'weet niet' antwoordalternatief in vragenlijsten, hoe leren we practici beschikbare gegevens te gebruiken, hoe kun je het IRT-model handzaam toepassen, hoe zet je het hiërarchische lineair (multilevel) model in, en hoe meet je toegevoegde waarde? Naast dit alles kan men (beginnende) evaluatoren aan het werk zien, die resultaten presteren van een uitgevoerd evaluatie-onderzoek naar bijvoorbeeld de effecten van een summerschool program, of van 'lunch buddies', of van een actie-onderzoek 'to empower teachers'.

De interessantere thema's betreffen de grote stimuleringsprogramma's vergelijkbaar met wat hier het onderwijsvoorrrangsbeleid genoemd werd, thans zijnde het gemeentelijke onderwijsachterstandenbeleid, de doorwrochte interventieprogramma's (onder het label 'school restructuring'), de grootschalige initiatieven op het gebied van klassenverkleining en 'school choice'.

Een belangwekkende sessie had betrekking op Title 1, de opvolger van Chapter 1 waarin grootschalig, als uitvloeisel van de 'War on poverty' uit de jaren zestig, het probleem van onderwijsachterstanden wordt aangepakt. Wat leerde de evaluatie van het eerste hoofdstuk ons? Om te beginnen bleek de implementatie van het oude programma gebrekkig te zijn geweest. Leerlingen werden er niet beter van, vermoedelijk omdat er zich in de kwaliteit van de instructie geen verbeteringen voordeden. De middelen verdwenen goeddeels aan ongekwalificeerde klassen-assistenten. De opvolger, Title 1, heeft veel meer een op standaarden (in plaats van een alleen op inputfinanciering) gebaseerde beleidsfilosofie. De evaluatie is echter lastig (zoals bij ons ook die van het OVB was) vanwege het gegeven dat het programma volledig dekkend is (en er dus geen controlegroepen meer zijn). In het algemeen blijken nog steeds veel van de middelen gebruikt te

worden voor de inzet van klassen-assistenten, en er lijkt een tendens dat het met de doelgroepen beter gaat. Wat onverminderd blijft is echter het contextuele effect: slechts 32% van de leerlingen in 'zwarte' scholen haalt de gestelde onderwijsstandaard. Reflecterend op deze bevindingen, stelde Stringfield dat er vier condities vervuld moeten zijn wil een aanpak een zekere slaagkans hebben:

- a) de aanpak moet bewezen effectief zijn (zie verderop);
- b) er moet sprake zijn van stabiel leiderschap (na een periode met een visionaire schoolleider is er iemand nodig die het werk afmaakt in plaats van het roer omgooit);
- c) de nagestreefde verandering moet door de school passend bij de eigen situatie worden vormgegeven;
- d) er moet een blijvende ondersteuning zijn op alle niveaus (voor de leerlingen, voor de leerkracht, de school en het district).

Het bekendste 'school restructuring' programma is Slavin's Roots & Wings, de opvolger van 'Success for All'. Basiselementen in het programma 'Success for All': 1-op-1 onderwijs, coöperatief leren, 2-maandelijke voortgangstoetsen, gezinsondersteuning, een intern begeleider en teamgerichte aanpak. De uitbreiding die met Roots & Wings gerealiseerd is zit in twee elementen. Naast de taalgerichte aanpak is er nu ook een specifieke rekenaanpak ontwikkeld, die sterk doet denken aan wat we hier ten lande het realistisch rekenen noemen. Bovendien hebben Slavin c.s. het World Lab ontwikkeld, een vorm van projectonderwijs waarbij de toepassing en het leren in context voor de wat oudere leerlingen (10 jaar en ouder) aan bod komen. De evaluatieresultaten zien er veelbelovend uit (de overdraagbaarheid van het programma naar andere landen stond op één van de vorige AERA's al eens geagendeerd), al maak ik hierbij graag de aantekening dat de resultaten wellicht enigszins vertekend zouden kunnen zijn door het verschijnsel 'testen-voor-de-test': de evaluatieresultaten zijn voor een deel gebaseerd op de resultaten van de toetsen die ook voor de voortgangscntrole worden gebruikt. Slavin's conclusie blijft dus onverminderd dat men het geld verkregen voor Title 1 of anderszins het beste kan benutten om het kostbare Roots & Wings te implementeren.

Die andere gelden zijn dan onder meer de met de klassenverkleining gemoeide financiën. Er is een federaal initiatief (waarmee zo'n 20 miljard gulden is gemoeid) om de leerkracht-leerling ratio in het Amerikaanse basisonderwijs terug te brengen naar 1 op 18. Er waren dan ook drie grote sessies gewijd aan dit thema. Allereerst werd verslag gedaan van de follow-up-resultaten van het STAR-klassenverkleiningsexperiment, dat zo'n 10 jaar geleden in Tennessee werd uitgevoerd. De gunstige resultaten voor de leerlingen die de kleinere klassen bezochten blijken (acht jaar nadien) te beklijven. De eerste resultaten van dat experiment (die halverwege dit decennium naar buiten kwamen) vormden voor Californië de aanleiding tot een eigen 'billion dollar initiative' (een grootscheepse klassenverkleiningsoperatie in de eerste leerjaren van het primair onderwijs) dat door een groot consortium wordt geëvalueerd. De eerste evaluatieresultaten verschijnen in de zomer van dit jaar, en in het symposium werden de evaluatie-opzet en enkele bijzonderheden van de implementatie gepresenteerd. Die laatste resultaten zijn opvallend: de instructie en de klassenorganisatie in verkleinde en normale klassen vertonen amper verschillen, met als enige uitzondering dat bij groepsgewijs werken in de kleine klassen de groepsgrootte met gemiddeld 3 leerlingen gunstig afsteekt tegen de 5 leerlingen in de grote klassen. Stasz meldde voorts dat er in de kleine klassen door de leerkracht meer tijd wordt vrij gemaakt voor de leerlingen die in het leesonderwijs achterblijven.

De summatieve evaluatie-opzet werd gepresenteerd door vader en zoon Mitchell, en tot verbijstering bleek de evaluatie-opzet niet te voorzien in een voortoets. (Vader) Mitchell bedacht in het verleden een aantal alternatieve verklaringsmodellen voor het optreden van eventuele klassenverkleiningseffecten. Eén van die verklaringen heeft betrekking op de afname van het aanvankelijk verschil in prestatieniveau binnen de kleine klas, zodat de groep voor de leerkracht makkelijker hanteerbaar wordt. De beide Mitchells hypothetiseren voorts een reductie van het Mattheus-effect: de verschillen tussen aanvankelijk laag en hoog presterende leerlingen worden in de loop der tijd in kleine klassen minder groot dan in grote klassen. Zowel het genoemde verklaringsmo-

del als de hypothese over het Mattheus-effect kunnen alleen onderzocht worden als er een voortoets is afgenomen. Daar komt bij dat Goldstein al vaker gewezen heeft op het belang van een voortoets ook bij echte experimenten (dat wil zeggen met aselecte toewijzing van leerlingen en leerkrachten aan de experimentele en controleconditie) in het onderwijs. Een voor de hand liggende vraag, die ik dus ook maar stelde, was waarom het onderzoeksonderwerp daarin niet voorzag. Het antwoord liet weer eens zien hoe beleidsmakers met geld omgaan: 2 miljard voor klassenverkleining, maar geen geld en middelen voor een sterke evaluatie-opzet met een voortoets. Het positieve nieuws was, dat de evaluatie-opzet de nodige nuancering van het begrip klassengrootte bevatte: de operationalisering is zodanig gekozen dat uitgesplitst naar activiteit en tijdperiode de omvang van de groep wordt bepaald waarin de leerling het onderwijs geniet, een variabel in plaats van een constant gegeven per leerling dus. Bleef ook hier natuurlijk de vraag: wat is nu eigenlijk de meest kosten-effectieve maatregel als we het onderwijsniveau van jonge leerlingen willen verhogen?

De laatste vermeldenswaardige sessie betrof 'school choice', iets wat we hier ten lande als een gegevenheid zien, maar dat elders wordt binnengehaald als een noviteit om de kwaliteitsimpuls die van de marktwerking in het onderwijs zou uitgaan te versterken. In Amerika betreft dit met name dan de zogenaamde 'charter schools'. Helaas alleen veel retoriek en weinig zinnige gedachten, laat staan deugdelijk empirisch bewijs dat die marktwerking inderdaad kwaliteitsverhogend werkt. Een interessant thema voor onderzoek in de toekomst, zeker waar we in Nederland het decentralisatiebeleid hebben ingezet, en waar ook andere landen de autonomie van scholen en de keuzemogelijkheden van ouders vergroten.

Divisie K: Teaching and Teacher Education
(D. Beijaard en A.M. Uhlenbeck, ICLON, Universiteit Leiden)

De ontwikkeling van curriculumstandaarden voor het K-12 onderwijs, standaarden voor zowel beginnende als ervaren docenten, standaarden voor lerarenopleidingen alsmede hun

onderlinge afstemming vormden de setting van een groot aantal bijdragen in de Divisie K. De nationale 'standard's movement' heeft ten doel het leraarsberoep te versterken en de kwaliteit van het lesgeven te verhogen en daarmee de kwaliteit van het leren van leerlingen te verbeteren. De wenselijkheid van standaarden, de interpretaties van deze standaarden door de verschillende betrokkenen en de mogelijke negatieve consequenties van het gebruik van standaarden kwamen in verschillende sessies aan bod. G. Delandshere, bijvoorbeeld, rapporteerde over onderzoek naar de wijze waarop in Indiana en Vermont de INTASC standaarden voor beginnende docenten werden geïnterpreteerd en opgenomen in het curriculum van de lerarenopleidingen. Uit deze studie bleek dat lerarenopleiders met weinig expliciete opvattingen over belangrijke thema's in de lerarenopleiding de INTASC standaarden gemakkelijk en kritiekloos overnamen zonder de onderliggende visie op onderwijs en opleiden te begrijpen en zonder het curriculum van de opleiding te wijzigen. Lerarenopleiders met een uitgesproken visie op opleiden en onderwijs stonden daarentegen kritisch tegenover de standaarden met als gevolg dat opleiders diepgaand met elkaar in gesprek gingen over de betekenis van de standaarden en de mogelijke consequenties voor de opleiding. De 'standard's movement' maakt overigens opnieuw duidelijk dat de implementatie van (grootschalige) innovaties niet zonder docenten en opleiders kan geschieden.

In een aantal symposia over het leren van leraren werd telkens de koppeling gelegd tussen het leren van leerlingen en de professionele ontwikkeling van leraren. Als veel werk van het leren van leerlingen wordt gemaakt, hetgeen momenteel alom gebeurt, moet ook veel werk gemaakt worden van het leren van leraren. De aandacht ging vooral uit naar de mogelijkheden voor leraren om in scholen te leren en naar de manieren waarop daarvoor betere voorwaarden kunnen worden gecreëerd. Gepleit werd voor het bestuderen van deze mogelijkheden en de mechanismen die daarbij een rol spelen. Veel waarde wordt in dit verband gehecht aan samenwerkend leren c.q. lerarenteams die in een school samenwerken aan verandering van curricula. Als men in staat is om

op productieve wijze samen te werken, dan blijkt ook veelal sprake te zijn van leren van leraren. Essentieel is dat er een bindmiddel is (het curriculum), dat er de mogelijkheid is tot uitwisseling van kennis en ervaringen (waarvoor het portfolio een geschikt instrument lijkt te zijn) en dat er sprake is van voldoende tijd. Met name de beschikbare tijd is doorslaggevend voor het welslagen van samenwerkend leren. Uit een van de presentaties bleek dat het overigens niet vanzelfsprekend is dat alle leraren de relatie leggen tussen samenwerking met collega's en hun eigen manieren van onderwijzen, ook al wordt aan alle voorwaarden voor een goede samenwerking voldaan.

De aloude tegenstelling tussen onderzoekers en leraren vormde de achtergrond van twee 'highlight' sessies van Divisie K c.q. trends in het onderzoek waarvan men vindt dat die bijzondere aandacht behoeven. In de eerste sessie stond de vraag centraal wie de vragen stelt die in de verschillende onderzoeksprogramma's aan de orde komen, wat als kennis telt en wie dat bepaalt. Vooraanstaande onderzoekers als V. Richardson en J. Brophy stelden dat het belangrijk is ruimte te bieden aan allerlei paradigma's in het onderzoek naar leraren: er is geen dominante manier waarop het onderzoek zou moeten worden verricht. Verder werd met nadruk gepleit voor de koppeling van lesgeven (instructie) en het curriculum in het onderzoek: tot nu zijn beide teveel afzonderlijk bestudeerd. Ook werd gewezen op het belang om kennis van mensen in de praktijk nadrukkelijk te betrekken bij het onderzoek naar wat leraren zouden moeten weten en kunnen. Op dit moment gebeurt dat (in de V.S.) niet of onvoldoende. In de tweede sessie ging het om de zeggenschap over wat als kennis geldt, waarbij de productie van kennis door overwegend blanke, goed opgeleide onderzoekers ter discussie werd gesteld. P. Grimmer en E. McWilliam wezen op de ethische dimensie in het onderzoek naar leraren. Het is volgens hen onethisch van onderzoekers om zich te verschuilen achter allerlei methodologische procedures, ook in kwalitatief onderzoek, en leraren te zien als objecten bij wie data worden verzameld. Scholen moeten in hun ogen niet alleen worden gezien als 'data plantages'. Zij vinden dat onderzoekers meer dan nu deel moeten uitma-

ken van de onderwijsgemeenschap. Onderzoekers hebben volgens hen tevens de plicht hun manieren van werken in deze gemeenschap te expliciteren en hun onderzoeksresultaten ten overstaan van deze gemeenschap openbaar te maken.

In de geest van deze 'divisional highlights' werd in een ander symposium, voorgezeten door L. Shulman, op bovengenoemde thematiek voortgeborduurd. De 'scholarship' van onderzoekers werd in dit symposium omschreven in termen van naar buiten gerichte en op elkaar voortbouwende activiteiten. De 'scholarship' van leraren werd daarentegen getypeerd als bestaande uit niet voor het publiek toegankelijke activiteiten, die niet op elkaar voortbouwen en bovendien niet worden geëvalueerd. Interessant waren de rapportages in dit symposium over effecten van speciale ('advanced') programma's voor leraren. Door deze programma's slaagden leraren erin een vergelijkbare 'scholarship' als onderzoekers te realiseren, namelijk door hen te leren hun kennis en ervaring te expliciteren en daarover te publiceren. M. Cochran-Smith, Vice President van Divisie K, noemde in dit verband een groot aantal leraren die erin zijn geslaagd in allerlei wetenschappelijke tijdschriften en handboeken te publiceren. Het toegenomen aantal leraren dat in staat is hun lokale kennis algemeen te maken ('making local knowledge public'), werd door haar als een bijzonder positieve ontwikkeling gekenschetst. Het levende bewijs van deze ontwikkeling waren enkele leraren die onderzoek rapporteerden van hun eigen onderwijspraktijk en daarover boeken hadden geschreven. De overheersende mening was dat leraren zeer goed kunnen laten zien dat hun werk 'scholarship' is. Het op grote schaal realiseren hiervan werd mede gezien als een belangrijke taak van de vakbonden voor leraren.

Een van de eerste symposia in divisie K ging over de kwaliteit van de stageplaats voor leraren in opleiding. Een goede stageplaats kenmerkt zich vooral door een goede interactie tussen de mentor in de school en de stagiaire. Als deze interactie niet goed is, hebben stagiaires significant andere reflecties op hun leerproces dan stagiaires in situaties waarbij deze

interactie wel goed is. Van belang is dat stagiaires programma-principes van hun eigen opleiding in de school herkennen, deze in de praktijk kunnen uitproberen en erop reflecteren. K. Zeichner, discussiant bij dit symposium, wees er met nadruk op dat een goede mentor veel meer is dan iemand die zijn of haar klas toegankelijk maakt voor een leraar in opleiding. De voorwaarden om een goede 'school-based' lerarenopleider te zijn vereisen volgens hem veel meer aandacht dan tot nu toe het geval is. Uit onderzoek onder allerlei betrokkenen, onder wie ook mentoren in de scholen zelf, blijkt dat mentoren vooral gezien worden als personen die verantwoordelijk zijn voor klassen, die onderwijs organiseren en verzorgen enz. Pas in laatste instantie wordt gedacht aan hun rol als mentor ten behoeve van het opleiden van leraren.

Minder dan in voorgaande jaren werd gepresenteerd over kennis en opvattingen van leraren. Een zeer interessant symposium over deze materie was georganiseerd door onderzoekers van de universiteit van Utah. Vanuit verschillende invalshoeken werd in dit symposium het 'teacher thinking' onderzoek geëvalueerd. Aan de basis hiervan lag de bestudering van ruim 250 studies. Uit deze evaluatie bleek onder meer dat kennis en opvattingen van leraren ('knowledge and beliefs') bepalend zijn voor wat iemand leert en hoe een leraar tegen onderwijsverandering aankijkt. Met name 'beliefs' van leraren, die vaak een biografische oorsprong hebben, vormen nogal eens blokkades voor verandering. Toch kunnen 'beliefs' van leraren door middel van scholing veranderen. Het moet dan niet gaan om 'single shot' interventies, zoals een workshop van een of enkele dagen: uit veel studies komt duidelijk naar voren dat leraren niet reageren op kortdurende interventies die zijn gericht op verandering van hun 'beliefs'. Verandering van 'beliefs' vindt overigens ook lang niet altijd plaats bij interventies van lange duur. Belangrijker dan de duur van de interventies is het type activiteiten

dat wordt verricht tijdens die interventies. Leraren gedurende een bepaalde tijd bijvoorbeeld confronteren met nieuwe leermaterialen is niet toereikend voor verandering van 'beliefs'; interactie met collega's over deze materialen en een ondersteunende omgeving om te kunnen experimenteren zijn daarbij minstens zo belangrijk. Op basis van de evaluatie van de 'teacher thinking' studies, tenslotte, werd gepleit voor het doen van onderzoek naar cognitieve processen die nodig zijn voor het bewerkstellingen van conceptuele verandering bij leraren en naar typen interventies die daarbij behulpzaam kunnen zijn. Dit onderzoek zou methodologisch beter moeten worden opgezet dan tot op heden is gebeurd bij veel 'teacher thinking' onderzoek. In veel onderzoek naar kennis en 'beliefs' van leraren blijken zich nogal eens problemen voor te doen, zoals een gebrekkige theoretische onderbouwing, de kleinschaligheid en het overwegend descriptieve karakter van het onderzoek alsmede het feit dat conclusies vaak niet te herleiden zijn tot de data. Ook werd als probleem gesignaleerd dat veel onderzoekers (lerarenopleiders) hun eigen praktijk onderzoeken en dus een relatie hebben met hun leraren (in opleiding), hetgeen een bedreiging vormt voor de validiteit van het onderzoek.

Ter afsluiting constateren we dat het symposium 'A closer look at the professional development of teachers', met bijdragen uit Leiden, Nijmegen en Utrecht, goed werd bezocht en leidde tot een levendige discussie. Het betrof een symposium verzorgd door leden van de Divisie Lerarenopleiding en Leraarsgedrag van de Nederlandse VOR dat was georganiseerd in samenwerking met de AERA.

Aan deze kroniek werkten mee: *J. van den Akker, D. Beijgaard, R. Bosker, H. Dekkers, C. Glas, F. de Jong R. Meijer, P. Slegers, D. Uerz en A. Uhlenbeck.*

De eindredactie was in handen van *H. Dekkers.*

Mededelingen

Inhoud andere tijdschriften

Tijdschrift voor Orthopedagogiek
Jaargang 38 nummer 6, 1999

'Zover' - afscheidscollage door *PA. de Ruyter*
Hoe nu verder? Commentaar op 'Zover' door
PM. Schoon

Het belang van auditief studeren voor blinde
universiteitsstudenten; een exploratief onder-
zoek - door *J. Laurijssen, P. Ghesquière &*
A.J.J.M. Ruijsenaars

Pedagogisch Tijdschrift
Jaargang 24, nummer 1, 1999

Onderwijsvrijheid: voor en tegen (Red. *A.W.*
van Haften & G.L.M. Snik)

Onderwijsvrijheid: voor en tegen door *A.W.*
van Haften & G.L.M. Snik

Onderwijsvrijheid: een groot pedagogisch
goed door *S. Miedema & D. de Ruyter*

Onderwijsvrijheid en overheidszorg opnieuw
afgewogen door *P. van der Ploeg*

Vrijheid van onderwijs en maatschappelijke
cohesie door *H. Woldring*

De vrijheid tot segregatie door *S. Blom*

Naar een richtingvrije planning? door *B. Ver-
meulen*

Onderwijsvrijheid en marktdenken door *J. de
Jong & G.L.M. Snik*

Grondslagen van liberale visies op onderwijs-
vrijheid door *G.L.M. Snik*

*Meerdink, J., Weet u hoe het in een groep is? Kin-
deren over de kwaliteit van het primair proces
van de (semi)residentiële hulpverlening. Deel 3.,
SWP, Utrecht, 1999.*

*Meerdink, J., Weet u waarom ik hier ben? Kinderen
over de kwaliteit van het primair proces van de
(semi)residentiële hulpverlening. Deel 1., SWP,
Utrecht, 1999.*

*Meijer, W., Ploeg, P. van der, & Thoomes, D. (red.),
Pedagogiek als tijdrede, Intro, Baarn, 1999,
f 29,90.*

*Petri, J.M., Bij haar kan ik niet leren. Een andere
kijk op de aanpak van 'probleemkinderen' op
school., Elsevier/De Tijdstroom, Maarssen, 1999.*

*Ploeg, J. van der, & Scholte, E., ADHD in kort
bestek. Achtergronden, diagnostiek en hulpver-
lening, SWP, Utrecht, 1999.*

*Raad voor maatschappelijk ontwikkeling, Arbeid
en zorg. Reactie op de kabinetsnota Op weg
naar een nieuw evenwicht tussen arbeid en
zorg., Raad voor maatschappelijk ontwikkeling,
Den Haag, 1999.*

*Waalwijk, K., Janus Korczak over klein zijn en groot
worden, SWP, Utrecht, 1999, f 27,50.*

Ontvangen boeken

*Essen, M. van, & Imelman, J.D., Historische pedago-
giek. Verlichting, Romantiek en ontwikkelingen
in Nederland na 1800., Intro, Baarn, 1999,
f 29,90.*

*Malschaert, H., & Traas, M., Werkboek Opvoeding.
Theorie en praktijk., Intro, Baarn, 1999, f 29,90.*

*Meerdink, J., Weet u wat een hulpverlener moet
doen? Kinderen over de kwaliteit van het pri-
mair proces van de (semi)residentiële hulpverle-
ning. Deel 2., SWP, Utrecht, 1999.*

Schoolproblemen van basisschoolleerlingen en de uiteenlopende perceptie van ouders en leerkrachten

A. Kassenberg en R. Bosman*

Samenvatting

De vraag welke leerlingen binnen het basisonderwijs probleemleerlingen zijn hangt voor een belangrijk deel af van de definiëring van een probleemleerling door leerkrachten en ouders. De waarneming van schoolproblemen en de vraag onder welke condities ouders en leerkrachten verschillen in hun perceptie van de problemen van een leerling vormen het onderwerp van deze studie. Uit een steekproef van 1347 leerlingen uit groep 4 van de basisschool komt naar voren dat ouders en leerkrachten niet altijd goed in staat zijn de schoolproblemen van leerlingen te beoordelen en daardoor niet altijd dezelfde leerlingen aanmerken als probleemleerling. Via beschrijvende statistische methodes en multiple regressie-analyse wordt aangetoond dat er tussen leerkrachten en ouders geringe consensus bestaat wat betreft hun perceptie van schoolproblemen. Met name wat betreft gedragsproblemen verschillen de meningen van ouders en leerkrachten. Een iets grotere consensus bestaat er ten aanzien van de prestaties van een leerling.

Inleiding

Sinds een aantal jaren is er meer aandacht voor leerlingen met problemen op school. Aanleiding voor deze belangstelling van onder meer de overheid en sociale wetenschappers vormt het 'Weer Samen Naar School'-akkoord van april 1991 (Ministerie O & W, 1991). Daarmee is een nog voortdurend beleidsproces in gang gezet om de dóórgaande groei van het speciaal onderwijs (LOM, MLK en IOBK) een halt toe te roepen (Doornbos & Stevens, 1987). Het beleid beoogt door middel van het aanbieden

van adaptief onderwijs meer leerlingen binnen het basisonderwijs te houden.

Onderzoek van Doornbos (1991) wijst uit dat veel leerlingen met problemen die op de basisschool blijven zich niet of nauwelijks op relevante punten onderscheiden van leerlingen die wel worden doorverwezen naar het speciaal onderwijs. De geringe verschillen tussen deze leerlingen worden mede veroorzaakt door het feit dat verscheidene ouders zich verzetten tegen de verwijzing. In ons land verzet naar schatting 30 à 40% van de ouders zich als het doorverwijzing naar LOM- of MLK scholen betreft. Vaak wordt dan (op aandringen van de ouders) uitgeweken naar een andere basisschool (Doornbos, 1991). Andere ouders lijken het gesegregeerde systeem voor speciaal onderwijs te steunen: 'Segregatie en discriminatie hebben voor deze ouders niet veel met elkaar te maken.' (De Wit, 1996, p. 35). In tegenstelling tot veel ouders zien leerkrachten in het verwijzen van leerlingen met schoolproblemen een goede maatregel. Met name storend gedrag dat het lesgeven bemoeilijkt en de sfeer op school bederft, en andere moeilijk te hanteren omgangsproblemen in de groep worden als een verzwaring van de werklust beschouwd. In combinatie met leerproblemen leiden deze gedragsmoeilijkheden bij de leerkracht al gauw tot het advies speciaal onderwijs (Doornbos, 1991). De mate waarin ouders en leerkrachten verwijzing naar het speciaal onderwijs als een oplossing zien, hangt voor een belangrijk deel samen met belangenverschillen tussen leerkrachten en ouders. Niet onafhankelijk van de milieufactor is het voor sommige ouders onacceptabel dat hun kind op een school voor speciaal onderwijs wordt geplaatst. Leerkrachten daarentegen zijn vaak gebaat bij een verwijzing, met name wanneer het leerlingen met gedragsproblemen betreft. Het prestatieniveau en de orde in de klas blijven dan gehandhaafd.

In onderwijssociologisch onderzoek worden

* Met dank aan R. Popping voor zijn opmerkingen bij een eerdere versie van dit artikel.

prestaties en problemen van leerlingen veelal beschouwd vanuit het perspectief van de leerkracht. Vanuit het feit dat leerkrachten veelal een andere visie hebben dan de ouders roept dit de vraag op hoe de ouders van leerlingen met schoolproblemen tegen het kind aankijken. Leerlingen met problemen op school zijn gebaat bij een zo vroeg mogelijke aanpassing van het leerproces, ondersteund door zowel de leerkracht als de ouders. Daarom is het van groot belang dat leerkracht en ouders hetzelfde inzicht hebben in de gedrags- en/of leerproblematiek. Immers, wanneer de probleempercepties van beiden verschillen dan zullen de opvattingen over de te treffen maatregelen ook verschillen. Voor een effectieve remediëring is het dan ook noodzakelijk dat er tussen leerkracht en ouders consensus bestaat over de problematiek en de aanpak daarvan. Dat maakt het interessant beide perspectieven te combineren en de vraag te stellen naar consensus tussen ouders en leerkrachten over hun perceptie van schoolproblemen. De centrale vraag van dit onderzoek is daarmee geschetst en als volgt geformuleerd:

1. *Op grond van welke criteria worden leerlingen als probleemleerling aangemerkt door de ouders en de leerkracht?*
2. *Welke factoren zijn van invloed op de consensus tussen de leerkracht en de ouders wat betreft hun perceptie van de schoolproblemen?*

1 Theoretisch kader

1.1 De probleemleerling

Probleemleerlingen zijn om verschillende redenen niet eenvoudig te definiëren. In de eerste plaats is een objectief criterium niet voorhanden. Ten tweede is met de opkomst van scholen voor speciaal onderwijs de norm voor probleemleerlingen veranderd. Leerlingen die vroeger op school bleven worden tegenwoordig vroegtijdig doorverwezen naar het LOM of MLK onderwijs. Ten derde zijn de problemen van leerlingen moeilijk tussen scholen te vergelijken aangezien per school de norm verschilt. Daarnaast wordt de inhoud van gedrag en prestaties van een leerling door de belangrijkste waarnemers, de ouders en de leerkracht, verschillend gepercipieerd. Omdat het oordeel

van de leerkracht en de ouders zich bovendien wijzigt onder invloed van maatschappelijke ontwikkelingen is de omschrijving van een probleemleerling tijd- en plaatsgebonden.

Bij schoolproblemen gaat vaak de meeste aandacht uit naar de prestatieproblemen van een leerling (Maas, 1992). Daarnaast vormen ook gedragsproblemen een belangrijke factor in de bestudering van schoolproblemen. Tussen het vóórkomen van beide typen problemen bestaat vaak een nauwe samenhang. Gedragsproblemen in de klas kunnen lagere leerprestaties tot gevolg hebben. Het omgekeerde effect van zwak presteren op schoolgedrag is, hoewel minder vaak onderwerp van studie, een aantal jaren geleden onderzocht door Van Liere (1990).

'Gedrag en prestaties, dat kan niemand ontkennen zijn gebonden aan biologische wortels. Dat bepaald gedrag problematisch heet, laat zich echter niet biologisch verklaren; dat betreft een maatschappelijk of wel sociologisch fenomeen.' (Bakker, 1984, p. 19). Datgene wat als een probleem wordt gezien kan worden opgevat als de uitkomst van een proces van sociale definiëring (Geelen & De Boom-Van Duin, 1983). Het basisonderwijs wordt gedomineerd door normen en waarden van de middenklasse (vgl. Jungbluth, 1996; Peschar & Wesselingh, 1995). Dit impliceert dat de 'speelruimte' die leerlingen hebben om problemen te vertonen afhankelijk is van de afstand tussen het thuis- en schoolmilieu. Leerlingen uit hogere en midden milieus krijgen van huis uit eigenschappen mee die overeenstemmen met wat er op school gewenst is, waardoor zij in het algemeen al gauw voldoen aan de schoolse normen. Deze leerlingen hebben zo meer 'speelruimte' voordat zich in de visie van de leerkracht problemen openbaren. Kinderen uit lagere milieus hebben veelal niet de vaardigheden die op school vereist zijn en wijken daardoor al snel in negatieve zin af van de norm. Voor hen is de afstand tot de probleemstatus veel kleiner, en iedere afwijking van hun 'normale' schoolhouding wordt dan al gauw als probleem gezien. In het kader van dit onderzoek worden schoolproblemen gedefinieerd als prestaties en gedrag van een leerling die in 'conflict' zijn met wat de omgeving 'normaal' acht, in die zin dat een bepaalde grens wordt overschreden.

1.2 De leerkracht

Leerkrachten hebben in de schoolpraktijk vaak een intensieve omgang met leerlingen. Omdat zij bovendien formeel tot oordelen bevoegd zijn, spelen zij een belangrijke rol bij het beoordelen van gedrag en prestaties, en worden zij gezien als de centrale figuur bij de onderkenning en behandeling van problemen (Rispen, 1990; Meijer, 1986). Onderzoek wijst uit dat in 80% van de verwijzingen van leerlingen naar het speciaal onderwijs de leerkracht het verwijzingsproces op gang brengt (Rispen, 1990; Meijer, 1986).

Leerkrachten signaleren problemen niet alleen op grond van feitelijke toetsprestaties, maar ook op grond van de positie die de leerling volgens de leerkracht inneemt in de klas (Meijnen, 1984). De klascontext, gekleurd door schoolse normen en waarden, waarbinnen verschillende leerlingen onderling met elkaar worden vergeleken vormt het *referentiekader* van de leerkracht. Leerlingen blijken in die praktijk vaak gespiegeld te worden aan een fictieve 'ideale' leerling waarin standaarden met betrekking tot het milieu, de werkhouding, moreel en sociaal gedrag liggen opgesloten (Hargreaves, Hester & Mellor, 1975).

De leerkracht signaleert de problemen bovendien niet als een objectieve buitenstaander, maar is een direct betrokkene die zich een *beeld* van leerlingen vormt en tegen die achtergrond tot een beoordeling van prestaties en gedrag komt. Dikwijls spelen daarbij op het oog niet relevante factoren mee; in het bijzonder milieu en sekse van de leerling beïnvloeden de waarneming. Over het algemeen zijn leerkrachten van mening dat kansrijke kinderen beter kunnen presteren, maar ook meer worden gesteund door hun ouders. Kinderen uit lagere milieus worden verondersteld lager te presteren. Dat leerlingen daarnaast ook seksetypische verschillen vertonen staat voor docenten buiten kijf. Van meisjes hebben leerkrachten met name een positiever beeld in termen van sociaal gedrag en werkhouding. Zij hebben in de ogen van de leerkracht over het algemeen minder vaak individuele hulp nodig, maar presteren gemiddeld iets lager. Jongens worden veel meer gezien als druk, geïnteresseerd en goede prestaties leverend (Van der Kleij, 1983).

De beeldvorming rond een leerling vormt het uitgangspunt voor de *verwachtingen* die

leerkrachten koesteren ten aanzien van de toekomstige prestaties en het gedrag van een leerling. Door het onderzoek van Rosenthal en Jacobson in 1968, dat bekend is geworden als het 'Pygmalion-experiment' is meer duidelijkheid verkregen over de gevolgen van de leerkrachtverwachtingen voor de prestaties van het kind. De uitkomsten van het onderzoek laten zien dat wanneer een leerkracht vermoedt dat een leerling beter kan presteren, deze zich tegenover de leerling zo zal gedragen dat er werkelijk een prestatieverbetering optreedt, ook al bestaat er geen objectieve grond voor deze vermoedens (Van der Hoeven, 1990).

De perceptie van de schoolproblemen hangt direct samen met de verwachtingen die de leerkracht ten aanzien van leerlingen koestert. Gedrag en prestaties die in negatieve zin niet aan de verwachtingen voldoen, worden als *problemen* gepercipieerd.

1.3 De ouders

Bodin (1996) komt in haar onderzoek tot de conclusie dat in ongeveer een derde van de gevallen de ouders schoolproblemen als eerste opmerken. De gezinsomgeving met eventuele andere kinderen vormt daarbij een belangrijk *referentiekader* voor ouders. Deze context met de daar geldende normen en waarden zijn bepalend voor ouders wanneer zij afwijkend gedrag signaleren. Normerend zijn in het algemeen demografische kenmerken zoals opleiding, werk, inkomen, de buurt en de beleving van de eigen opvoeding (Peeters & Hermans, 1996). Zo vormen ouders zich op grond van de in hun sociale omgeving geldende normen een *beeld* van gewenst gedrag en het prestatieniveau.

Verschiedende onderzoeken tonen aan dat niet kan worden geconcludeerd dat bepaalde groepen ouders onderwijs minder belangrijk vinden, wel dat hun *verwachtingen* verschillen (Van Eck, Ledoux & Veen, 1994; Van der Hoeven, 1994; Van Liere, 1990). Op basis van hun eigen ervaringen stellen ouders uit lagere milieus hun verwachtingen ten aanzien van hun kinderen laag en twijfelen vaker dan andere ouders aan de capaciteiten van hun kind (Driessen, 1996). Een contrast hiermee vormen allochtone ouders met een lage opleiding. Hoewel hun kinderen gemiddeld drie maal zo vaak blijven zitten in vergelijking tot autochtone

leeftijdgenootjes, zijn de ouders behoorlijk optimistisch over de capaciteiten van hun kinderen en over hun toekomst in het voortgezet onderwijs (Driessen, 1996). Ouders uit de midden- en hogere klassen ontleen hun sociale positie aan een hogere (beroeps)opleiding. Kennis en diploma's zijn voor deze klassen vaak de enige kapitaalbron die zij door kunnen geven aan hun nageslacht. Onderwijs dient vooral als defensieve strategie om de eigen positie te consolideren en het risico op sociale daling tegen te gaan (Waslander & Bosman, 1997). Deze ouders hebben dus een buitengewoon groot belang bij onderwijs en hebben om die reden hoge verwachtingen van hun kind.

Ouderlijke verwachtingen verschillen vaak ook wat betreft zonen en dochters. In de praktijk betekent dat, hoewel ouders onderwijs voor ieder even belangrijk vinden, dat jongens en meisjes seksespecifiek gesocialiseerd worden (Jungbluth, 1996). Ouders blijken jongens over het algemeen meer competitief en agressief op te voeden dan meisjes. Daarnaast worden van jongens in het algemeen goede schoolresultaten verwacht, waardoor ze met meer prestatieuitdaging worden benaderd. De einddoelen die ouders voor meisjes voor ogen hebben liggen dikwijls anders. Meisjes worden meer beschermd behandeld en niet bijzonder uitgedaagd beter te gaan presteren. Van hen wordt meer dan van jongens verwacht dat zij zich conform de schoolregels gedragen (Van Eck et al., 1994).

De seksespecifieke verwachtingen van ouders hangen direct samen met de sociale klasse waartoe het gezin behoort; er is sprake van een interactie-effect. In de lagere sociale klassen wordt het laag presteren van meisjes als minder problematisch gezien dan het laag presteren van jongens. Ten aanzien van gedrag geldt in deze lagere sociale klassen dat alhoewel jongens over het algemeen strikter worden opgevoed dan meisjes, zij in de regel uiteindelijk toleranter beoordeeld worden. Ouders in de lagere klassen reageren vooral op het direct waar te nemen gedrag, ouders in de hogere klassen meer op de intentie dan op het effect (Dekkers & Smeets, 1982).

Onafhankelijk van de hoogte van de verwachtingen geldt voor de ouders hetzelfde als voor de leerkracht: het niet voldoen aan de verwachtingen in negatieve zin betekent voor hen

dat die prestaties en gedrag als *probleem* worden *gepercipieerd*.

1.4 Perceptieverschillen

Vanuit de beschrijving van de leerkracht- en ouderperceptie van schoolproblemen kunnen nu mogelijke verschillen in de probleemperceptie van ouders en leerkrachten begrepen worden. Het aanmerken van een probleemleerling wordt gezien als een proces van sociale definiëring, waarbij de uitkomst wordt bepaald door de (sub)culturele normen en waarden die de betrokkenen aanhangen (Merton, 1971). Het referentiekader speelt een belangrijke rol als context waarbinnen gedrag en prestaties worden waargenomen en beoordeeld. Juist het feit dat leerlingen door de leerkracht en de ouders vanuit verschillende referentiekaders, met mogelijk verschillende normen, worden waargenomen wordt hier gezien als oorzaak voor het bestaan van perceptieverschillen tussen beiden. Leerkrachten nemen een leerling waar door interactie in de *klas*, ouders zien een kind in de *dagelijkse leefomgeving*. De informatie die leerkrachten en ouders over een leerling hebben is contextgebonden en verschilt daarvoor van elkaar. Daarnaast speelt de afstand tussen de thuis- en de schoolcultuur een belangrijke rol. Naarmate de sociale achtergrond van de ouders met de daarbij geldende normen en waarden dichter de schoolomgeving met de daar geldende normen en waarden nadert, hebben ouders en leerkrachten meer dezelfde verwachtingen van een leerling wat betreft diens prestaties en gedrag. Het niet voldoen aan die verwachtingen geldt voor beide op hetzelfde prestatie- en gedragsniveau als problematisch, waardoor de perceptieverschillen tussen deze ouders en de leerkracht dan ook kleiner worden verondersteld. Ouders die een andere norm hanteren voor hun kind zijn, afhankelijk van de hoogte van hun verwachtingen, sneller of minder snel teleurgesteld in die verwachtingen. Ouders die hun verwachtingen lager stellen dan de leerkracht signaleren (nog) geen problemen bij hun kind op een punt waar de leerkracht deze wel waarneemt. Omgekeerd zullen ouders met hogere verwachtingen eerder en bij een hoger prestatieniveau van hun kind niveau problemen signaleren bij hun kind dan de leerkracht. De factoren sekse van de leerling en gezinsmilieu zijn bepalend voor de hoogte

van de verwachtingen en de probleempceptie van ouders en de verschillen daarin met die van de leerkracht. Een niet in deze studie getoetste veronderstelling is tot slot de verwachting dat de mate waarin er consensus bestaat tussen de leerkracht en de ouders wat betreft de schoolproblemen van een leerling van invloed is op de te nemen maatregelen of het behandelingsplan. In Figuur 1 zijn de veronderstelde theoretische relaties schematisch weergegeven.



Figuur 1. Theoretisch model

2 Methode

2.1 Onderzoekspopulatie

De gegevens voor dit onderzoek zijn ontleend aan twee andere onderzoeken. Beide onderzoeken zijn uitgevoerd door het ITS in samenwerking met het SCO-Kohnstamm Instituut. Het eerste en het belangrijkste deel van de data is afkomstig uit het PRIMA-cohort onderzoek 1994/95 dat representatief is voor alle Nederlandse basisscholen (Jungbluth, Peetsma & Roeleveld, 1996). Het tweede deel van de data betreft het WSNS-onderzoek waarin op een selectie van 54 PRIMA-basisscholen aanvullend een meer specifieke studie is verricht naar 'risicoleerlingen in het basisonderwijs' (Jungbluth & Van Langen, 1997). In het onderhavige onderzoek worden gegevens uit beide onderzoeken gecombineerd. De noodzaak te kunnen beschikken over zowel de probleeminschatting van de ouders als die van de leerkracht heeft de steekproef ingeperkt tot alle

leerlingen van groep 4 op een selectie van 54 PRIMA-basisscholen (N=1347).

De gegevens zijn verzameld bij drie groepen respondenten. In de eerste plaats zijn bij de leerlingen eenmalig gegevens verzameld wat betreft hun werkelijke prestaties via de CITO taal- en rekentoetsen. In de tweede plaats is via leerkrachtvragenlijsten informatie verkregen over hun inschatting van de prestaties, het gedrag en de behoefte van leerlingen aan hulp of maatregelen. Tenslotte is middels de ouder-vragenlijst inzicht verkregen in de inschatting die zij hebben van de prestaties, het gedrag en de behoefte van hun kind aan hulp of maatregelen.

De onderzoekspopulatie kan ten opzichte van de landelijk representatieve steekproef (PRIMA 1994/95) op een aantal belangrijke punten nader worden beschreven. Uitgesplitst naar sekse als één van de *leerlingkenmerken*, blijkt dat er 52,3% jongens zijn en 47% meisjes. Van 9 leerlingen (0,7%) was de sekse niet ingevuld. Ten aanzien van de variabele milieu, blijkt dat leerlingen van allochtone afkomst en ouders uit een lager milieu in deze studie in hoge mate oververtegenwoordigd zijn (24,6% ten opzichte van 6,9% in de representatieve steekproef). Als gevolg daarvan zijn er relatief minder leerlingen uit de midden en hogere milieus dan op grond van de representatieve steekproef zou worden verwacht. Als gevolg van de soms geringe bereidwilligheid van scholen (met vooral achterstandsscholen) achtergrondgegevens van leerlingen te verstrekken, is van een relatief groot aantal (23,9%) leerlingen de sociaal-etnische achtergrond onbekend. Wat betreft de *ouderkenmerken* is door 70% van de ouders gerespondeerd, in voldoende mate verdeeld over de onderscheiden milieucategorieën. Uit de *schoolkenmerken* kan worden opgemaakt dat de oververtegenwoordiging van het aantal allochtone leerlingen verklaard wordt door de oververtegenwoordiging van het aantal scholen met een overwegend allochtoon publiek, en het grote aantal stadsscholen in het midden-westen van het land op openbare grondslag. De oververtegenwoordiging van allochtone leerlingen wordt niet als een belemmering voor het onderzoek gezien, omdat de aandacht in dit onderzoek met name uitgaat naar leerlingen met problemen op school en deze zich vaak onder

allochtone leerlingen bevinden. Daarnaast blijkt tevens uit de ouderrespons dat ondanks het hoge percentage leerlingen waarvan de sociaal-etnische achtergrond ontbreekt, elke milieucategorie voldoende vertegenwoordigd is.

2.2 Variabelen

Voor dit onderzoek is het noodzakelijk het theoretische concept 'probleemleerlingen' nader te operationaliseren. Om nu deze categorie leerlingen te kunnen aanwijzen is gebruik gemaakt van drie methoden. De eerste in dit onderzoek belangrijke methode is het vaststellen van problemen op basis van de *leerkrachten ouderinschatting*. Om probleemleerlingen op deze wijze te kunnen identificeren zijn de items uit de leerkracht- en oudervragenlijsten van het PRIMA en het WSNS onderzoek opnieuw geschaald (Kassenberg, 1997). Via factoranalyse zijn 3 nieuwe variabelen verkregen. De betrouwbaarheid is achter iedere schaal weergegeven voor respectievelijk de leerkracht- en de oudervragenlijst¹. De eerste variabele betreft de inschatting van de prestaties van een leerling ($\alpha=.90$ en $\alpha=.73$) en de tweede variabele de inschatting van het gedrag ($\alpha=.81$ en $\alpha=.63$). Ten derde is een variabele verkregen die aangeeft in hoeverre de leerkracht dan wel de ouders hulp/maatregelen noodzakelijk achten voor een leerling ($\alpha=.88$ en $\alpha=.69$).

De tweede, meest gangbare, methode is het op basis van de toetsprestaties vaststellen van leerlingen die (landelijk gezien) het laagst scoren. Deze leerlingen worden *absolute* laagstpresteerders genoemd. Daarnaast is uit de theorie naar voren gekomen dat ouders en leerkrachten via vergelijking met andere kinderen tot een beoordeling van prestaties en gedrag komen en niet in een vergelijking met een landelijk gemiddelde. Daardoor is het vaak ondoenlijk voor hen de prestaties van leerlingen in het landelijke kader te beoordelen. Bepalend is de positie die een leerling inneemt in de klas. Zo worden probleemleerlingen uiteindelijk vaak aangemerkt op het laag presteren binnen de jaargroep. Om die reden is per school en per groep de relatieve rangpositie op grond van de taal- en rekenscores van iedere leerling berekend. Op deze derde wijze worden er tenslotte binnen iedere school *relatieve* laagstpresteerders aangemerkt.

Om de categorie probleemleerlingen in de drie benaderingen te kunnen afbakenen is er steeds vanuit gegaan dat de leerlingen met de laagste scores als problematisch gelden. Zowel in de ouder- en leerkrachtinschatting als in de taal- en rekentoetsscores (absoluut en relatief) zijn categorieën aangebracht in de scores, waarbij het feit dat iedere categorie door een voldoende aantal leerlingen moet worden gerepresenteerd als uitgangspunt heeft gediend. Er is gekozen voor een vijfdeling waarbij in iedere categorie evenveel leerlingen vallen. De op deze wijze geconstrueerde groepen omvatten steeds 20% van de populatie en worden hier 'kwintielen' genoemd.

Om het verschil in perceptie tussen ouders en leerkracht meetbaar te maken is voor de schalen prestatie-, gedrags- en hulpinschatting steeds de inschatting van de ouders (waarde 1-5) in mindering gebracht op de inschatting van de leerkracht (waarde 1-5). Zo zijn drie nieuwe variabelen (prestatie-, gedrags- en hulpperceptieverschil) ontstaan met een bereik dat loopt van -4 tot en met +4. Beide uitersten worden bereikt wanneer de leerkracht en de ouders het geheel oneens zijn met elkaar. De waarde 0 wordt daaruit volgend verkregen wanneer er sprake is van consensus tussen leerkracht en ouders over de inschatting van problemen: de leerling had in de ogen van beiden dezelfde score.

Via deze werkwijze is het te verwachten dat er bij probleemleerlingen op basis van de tweede en de derde definitie verschillen optreden wanneer leerlingen als absolute, dan wel als relatieve laagstpresteerder (categorie-1) worden aangemerkt, aangezien per school het gemiddelde prestatieniveau verschilt. Op bepaalde scholen zoals 'elite scholen' zullen als het om de absolute indeling gaat geen leerlingen in categorie-1 vóórkomen, terwijl ze op basis van de relatieve score-indeling wel 20% laagstpresteerders hebben (zie ook Jungbluth & Van Langen, 1997). Leerlingen die relatief (binnen de eigen klas) dus tot de laagstpresteerders behoren kunnen absoluut (landelijk) in een aanmerkelijk hogere categorie terecht komen.

De variabele 'milieu' wordt via de typering sociaal-etnische achtergrond opgebouwd uit twee kenmerken van de ouders van leerlingen: hun geboorteland en hun opleidingsniveau (Jungbluth & Van Langen, 1997). De zesdeling

Tabel 1

Percentages door ouders en leerkrachten aangemerkte prestatie-probleemleerlingen per absolute en relatieve taal- en rekencategorie (vetgedrukt zijn correcte inschattingen)

	Werkelijke prestatiecategorie	Absoluut		Relatief	
		Rekenen	Taal	Rekenen	Taal
Leerkracht	1(laagstpresteerder)	40,4	28,6	42,2	28,3
	2	24,5	24,5	20,4	24,9
	3	9,0	17,3	7,3	11,9
	4	11,0	10,9	4,1	15,4
	5(hoogstpresteerder)	5,6	10,7	14,3	15,4
Ouders	1(laagstpresteerder)	28,9	20,5	30,4	18,2
	2	17,9	16,7	16,8	18,2
	3	10,3	14,6	2,1	12,6
	4	5,5	10,9	6,2	6,4
	5(hoogstpresteerder)	3,7	4,4	0,0	7,7

die hieruit is opgebouwd heeft de categorieën: 'LBO-Turks/Marokkaans', 'LBO-overig allochtoon', 'LBO-autochtoon', 'MBO', 'HBO/ WO', 'onbekend'.

Tenslotte wordt ook de variabele sekse van de leerling in het onderzoek betrokken.

2.3 Analyse

De onderzoeksopzet bestaat globaal uit twee delen. In de eerste plaats wordt er via verschillende beschrijvende methodes een overzicht geschetst van het aantal leerlingen dat volgens de leerkracht, volgens de ouders en in objectieve zin (blijkend uit de toetsen) problemen heeft op school. Inzicht in de juistheid van de inschatting van de ouders en de leerkracht wordt daarna verkregen door het aantal leerlingen die in hun respectievelijke visie prestatieproblemen hebben af te zetten tegen de werkelijke (objectieve) taal- en rekenniveaus. Via regressie-analyse is vervolgens het effect van sekse en milieu op de inschatting van de leerkracht en de ouders gemeten.

In de tweede stap zijn de perceptieverschillen tussen ouders en leerkrachten nader bestudeerd. De variabelen perceptieverschil wat betreft gedrag, prestaties en hulpinschatting zijn in drie regressie-analyses ingevoerd als afhankelijke variabele. Als predictoren zijn de mate van objectieve prestatieproblemen, de individuele leerlingkenmerken sekse en milieu en het bestaan van een eventueel behandelingsplan of maatregelen ingevoerd. Met behulp van een gepaarde t-toets is nagegaan of de leerkracht- en ouderpercepties statistisch significant van elkaar verschillen.

3 Resultaten

3.1 Waarnemen en perceptie van prestaties en gedrag

De uitkomsten van de eerste analyses geven inzicht in de waarneming van het gedrag en de prestaties van leerlingen door ouders en leerkrachten. Tabel 1 laat zien dat slechts een deel van de ouders (gemiddeld 25%) en de leerkrachten (gemiddeld 35%) die leerlingen als probleemleerling waarneemt die dat ook op grond van hun werkelijke prestaties (taal en rekenscores) zijn.

De overige ouders en leerkrachten merken leerlingen als probleemleerling aan terwijl zij op grond van hun prestaties tot de goede of zeer goede leerlingen behoren (percentages in categorie 2-5 geplaatst). Tevens blijkt uit de cijfers dat leerkrachten 40,4% van de probleemleerlingen en ouders 28,9% van de probleemleerlingen ook als zodanig op basis van hun absolute rekenprestaties goed inschatten. Het blijkt dat probleemleerlingen op basis van rekenprestaties vaker goed worden ingeschat dan op basis van hun taalprestaties. Leerkrachten schatten 28,6% van de probleemleerlingen en ouders slechts 20,5% van de probleemleerlingen op basis van taalprestaties goed in. Theoretisch werd verondersteld dat leerkrachten en ouders de prestaties van een leerling beter kunnen inschatten ten opzichte van de rest van de klas (relatieve prestaties) dan ten opzichte van de landelijke populatie leerlingen (absolute prestaties). Deze veronderstelling wordt niet bevestigd. De percentages die conform de wer-

Tabel 2

Opbouw van de probleemleerlingen (categorie-1) wat betreft prestaties in absolute en relatieve zin, en wat betreft prestaties en gedrag volgens de leerkracht- en de ouderinschatting, naar sekse en sociaal-etnische achtergrond

	Prestaties		Prestatie-inschatting		Gedragsinschatting	
	absoluut	relatief	ouders	leerkracht	ouders	leerkracht
Jongens	44,5	46,3	54,0	52,3	58,7	62,9
Meisjes	55,5	53,7	46,0	47,7	41,3	37,1
LBO, Tur/Mar	38,2	18,4	10,9	16,8	16,4	12,5
LBO, ov.all	4,6	5,9	5,1	8,1	6,6	7,1
LBO, autocht.	16,8	24,3	29,1	26,6	20,4	22,0
MBO	10,7	22,4	23,4	19,1	23,7	28,0
HBO/WO	4,6	5,3	7,4	7,5	9,2	10,1
onbekend	25,2	23,7	24,0	22,0	23,7	20,2

kelijke prestaties in categorie-1 worden ingeschat liggen bij de relatieve prestaties niet hoger dan bij de absolute prestaties. Wat betreft het gedrag van een leerling is het niet mogelijk te bepalen in hoeverre ouders en leerkrachten in staat zijn dit correct in te schatten. De vergelijking met het werkelijke gedrag kan niet worden gemaakt omdat er geen objectieve toets voorhanden is.

In Tabel 2 wordt vervolgens uiteengezet wat de kenmerken zijn van de leerlingen die door hun ouders en de leerkracht worden aangemerkt als probleemleerling (categorie-1) en welke leerlingen dat ook daadwerkelijk zijn (absoluut en relatief).

Hoewel meisjes over het algemeen iets lager presteren dan jongens, worden zij minder vaak als prestatie-probleemleerling beschouwd (meisjes gem. 47% ten opzichte van jongens gem. 53%). Dit is waarschijnlijk het gevolg van het feit dat zij aanzienlijk minder vaak als gedrags-probleemleerling worden gezien (meisjes gem. 39% ten opzichte van jongens gem. 61%). Jongens worden, naar blijkt uit de cijfers, vaker als probleemleerling aangemerkt. Soms is dat alleen vanwege hun lage prestaties, maar vooral in combinatie met moeilijk gedrag. Verondersteld wordt dat er sprake is van een samenspel tussen prestaties en gedrag in de waarneming daarvan door de ouders en de leerkracht. Bepaalde lage prestaties worden als minder problematisch gepercipieerd wanneer dit samengaat met norm-conform gedrag. Soms, niet noodzakelijk extreem lage prestaties kunnen daarentegen als problematisch worden gezien wanneer dit samengaat met moeilijk schoolgedrag.

Leerlingen uit LBO-autochtone milieus en MBO milieus worden het vaakst als laagstpresterder én gedrags-probleemleerling gezien door ouders en leerkrachten (gem. 25% wat betreft prestaties, gem. 24% wat betreft gedrag), hoewel hun werkelijke prestaties daar geen directe aanleiding toe geven (gem. 14% is landelijk gezien laagstpresterder). Deze leerlingen worden ook vaker dan andere leerlingen aangemerkt als leerlingen die extra hulp/maatregelen nodig hebben. De groep allochtone leerlingen (met name Turks/Marokkaanse leerlingen) die op grond van hun werkelijke prestaties vooral tot de categorie laagstpresterders behoren (landelijk 38%), worden door de ouders en de leerkracht niet in overeenkomstige mate als prestatie-probleemleerling aangemerkt (gem. slechts 14%). Ook nemen beide partijen opvallend weinig gedragsproblemen bij de allochtone leerlingen waar (gem. 14%). Deze leerlingen komen volgens hun inschatting dan ook niet in aanmerking voor een behandelingsplan of maatregelen. Verondersteld wordt dat zowel de leerkrachten van de allochtone leerlingen die vaak op 'zwarte scholen' lesgeven, als de ouders minder goed kunnen refereren aan andere leerlingen; het valt hun dus minder gauw op dat deze leerlingen beneden gemiddeld scoren. Daarnaast hebben leerkrachten soms een dermate lage verwachting van Turks/Marokkaanse leerlingen dat maatregelen om hun prestaties te verbeteren niet nodig worden geacht (vgl. Driessen, 1996).

3.2 De invloed van sekse en milieu op de probleemperceptie

De invloed van sekse en het milieu van de leerling op de waarneming van de prestaties en het gedrag door de ouders en de leerkracht is via regressie-analyse getoetst. Uit de analyses voor zowel leerkrachten als ouders, waarin hun inschatting van prestaties en gedrag zijn opgenomen als afhankelijke variabele, blijkt de invloed van sekse en milieu gering wanneer de verklaarde variantie in ogenschouw wordt genomen. In het geval van de ouders wordt hun inschatting van de prestaties voor 2% verklaard door sekse en milieu. De gedragsinschatting van ouders wordt ook voor 2% verklaard door sekse en milieu. De inschatting van de leerkrachten van prestaties wordt voor bijna 5% verklaard door sekse en milieu. Sekse en milieu verklaren eveneens 5% in de variantie van de gedragsinschattingen van leerkrachten. Ondanks de lage verklaarde variantie is er een tendens waar te nemen. In het algemeen wordt de waarneming van prestaties door ouders en leerkrachten beïnvloed door het milieu van de leerling (en de ouders); terwijl de waarneming van gedrag daarentegen juist wordt beïnvloed door de sekse van een leerling.

In antwoord op de eerste onderzoeksvraag kan worden vastgesteld dat de kenmerken van probleemleerlingen verschillen per definitie die men hanteert voor deze leerlingen. Ouders en leerkrachten merken in hun perceptie van de gedrags- en prestatieproblemen andere leerlingen als probleemcategorie dan op grond van de werkelijke (absolute en relatieve) prestaties zou worden verwacht. Vermoedelijk is er sprake van een samenspel tussen prestaties en gedrag in de waarneming daarvan door ouders en de leerkracht. Een duidelijke invloed van

sekse en milieu bij de beoordeling van leerlingen door ouders en leerkrachten kon niet worden aangetoond.

3.3 Verschillen en consensus in probleemperceptie

Het tweede deel van de probleemstelling gaat specifiek in op de vraag onder welke omstandigheden er al dan niet sprake is van consensus tussen ouders en de leerkracht wat betreft hun perceptie van de problemen. Uit Tabel 3 blijkt dat de meeste consensus bestaat over de prestaties van leerlingen (35%). Gezien het feit dat dit het onderwerp is waar zowel de leerkracht als de ouders via schoolrapporten en toetsen het meest zicht op hebben is dit niet verwonderlijk. Tevens zijn de prestaties meestal het onderwerp van gesprek tijdens de zogenaamde 'tien minuten gesprekjes' op school tussen leerkracht en ouders, wat de percepties dichter bij elkaar kan brengen. De geringste consensus (23%) bestaat er ten aanzien van de gedragsproblemen van een leerling. Dat lijkt aannemelijk omdat voor gedrag de normen minder duidelijk zijn, en de wijze waarop gedrag wordt waargenomen sterk afhankelijk is van de context waarbinnen het wordt waargenomen. Vaak zullen ouders vanuit de thuissituatie anders tegen schoolgedrag aankijken dan de leerkracht vanuit de klas. De school stelt specifieke en andere eisen aan het gedrag van een leerling dan de ouders. Wat betreft de inschatting van de hulpbehoefte verschillen ouders en leerkrachten in 28% van de gevallen van mening.

De t-toets wijst uit dat de prestatieperceptie van ouders en leerkrachten significant van elkaar verschilt ($p < .05$). Via regressie-analyse is nagegaan wat het effect van eventueel getroffen maatregelen, de sekse van de leerling

Tabel 3

Mate perceptieverschillen tussen ouders en leerkrachten wat betreft prestaties, gedrag en hulpinschatting (in procenten)

Verschil waarden	Prestatieperceptie	Gedragsperceptie	Hulpinschatting
-4 (lk 1/ ou 5)	1,4	2,8	0,6
-3	3,4	5,3	3,1
-2	7,4	10,4	8,1
-1	17,7	17,8	21,0
0 (Consensus)	34,7	23,1	28,4
1	17,9	19,3	18,7
2	11,2	10,7	11,8
3	5,3	8,6	5,9
4 (ou 1/ lk 5)	1,0	2,1	2,4

en het milieu van de leerling en ouders op de perceptieverschillen is. Deze effecten zijn vrijwel afwezig gezien de verklaarde variantie van 2%. Slechts het bestaan van een behandelingsplan (maatregel) oefent een geringe invloed uit op de mate van consensus. Wanneer een dergelijk plan bestaat, of maatregelen zijn getroffen, dan zijn de perceptieverschillen, zoals verondersteld, kleiner. Een effect van sekse en milieu kon niet worden aangetoond.

Ten aanzien van het gedrag wijst de t-toets uit dat de ouder- en leerkrachtpercepties niet significant van elkaar verschillen ($p > .10$). Dit blijkt een gevolg te zijn van de wijze waarop de perceptieverschillen zijn geoperationaliseerd. Doordat zowel het aantal gevallen waarin de ouders positiever zijn dan de leerkracht en, omgekeerd, de gevallen waarin de leerkrachten positiever zijn dan de ouders ongeveer aan elkaar gelijk zijn, worden deze waarden tegen elkaar weg gemiddeld, waardoor verschillen niet meer inzichtelijk zijn. Uit tabel 3 is echter gebleken dat deze verschillen ten aanzien van gedrag wel degelijk aanwezig zijn. De invloed van sekse en milieu van een leerling, als ook van de eventueel getroffen maatregelen is blijkens de regressie-analyse gering en bovendien niet significant ($R^2=1\%$).

Ouders en leerkrachten verschillen ook significant van elkaar in de mate waarin zij hulp nodig achten voor de leerlingen ($p < .01$). Een effect van sekse, milieu en de eventueel reeds getroffen maatregelen op de perceptieverschillen wordt via de regressie-analyses niet bewezen ($R^2=2\%$).

In antwoord op de tweede onderzoeksvraag blijkt dat er aanmerkelijke verschillen bestaan tussen de perceptie van de leerkracht en de ouders als het gaat om schoolproblemen. Met name wat betreft hun gedragsperceptie bestaat er vrij weinig consensus. De van belang veronderstelde factoren sekse, milieu en maatregelen verklaren weinig variantie in de perceptieverschillen.

objectieve scores op taal- en rekentoetsen zowel leerkrachten als ouders een aanzienlijke groep leerlingen ten onrechte *niet* als probleemleerling identificeren. Wat taal en rekenen betreft wordt gemiddeld 72% respectievelijk 59% van de leerlingen niet als probleemleerling (h)erkend door de leerkracht. Voor de ouders zijn deze percentages 81% en 70%. Ook het omgekeerde komt voor: leerlingen die als probleemleerling worden bestempeld, die dat op grond van de scores op taal- en rekentoetsen niet zijn (vgl. Tabel 2). De discrepantie met de objectieve scores is bovendien niet gelijk voor leerkrachten en ouders; zij verschillen in aanzienlijke mate in hun definiëring van wat een probleemleerling is (respectievelijk wat betreft gedrag, hulpinschatting en prestaties in 77%, 72% en 65% van de gevallen).

Bij deze constatering moet een tweetal opmerkingen gemaakt worden. De eerste betreft onze aanname dat de taal- en rekenscores van leerlingen in dit onderzoek een maat zijn voor de 'werkelijke' problemen van een leerling. Deze aanname lijkt gerechtvaardigd omdat het door het CITO gevalideerde toetsen betreffen, en de scores in de praktijk ook als indicator voor leersucces gebruikt worden. De tweede opmerking behelst een relativisering; alhoewel taal- en rekenscores valide indicatoren zijn en in dit onderzoek een zinvol onderscheid aanbrengen, wordt er in de praktijk juist ook veel waarde gehecht aan het oordeel (c.q. advies) van de leerkracht. Verondersteld mag worden dat de leerkracht naast de prestatiebeoordeling ook de beoordeling van het gedrag laat meewegen in zijn of haar oordeel. Dit zou althans een deel van het verschil tussen prestaties en oordeel kunnen verklaren, zeker waar het de discrepantie tussen 'goede' scores en een definitie als 'probleemleerling' betreft. Het is aannemelijk dat ook ouders in hun beoordeling van de prestaties het gedrag meewegen. Jongens blijken zo vaker als probleemleerling te worden aangemerkt dan meisjes. Daarnaast worden leerlingen van ouders met een LBO-autochtone of MBO achtergrond veelal als probleemleerling gezien, terwijl volgens de objectieve prestatietoetsen de LBO-Turks/Marokkaanse leerlingen hier meer aanleiding toe geven. Hieruit wordt duidelijk dat de definiëring van een leerling als probleemleerling niet eenduidig is. Ouders en leerkrachten

4 Conclusies en discussie

De resultaten van dit onderzoek laten zien dat veel ouders en leerkrachten leerlingen met schoolproblemen niet als zodanig inschatten. Dat wil zeggen dat naar de maatstaf van de

beoordelen leerlingen vanuit hun eigen context heel verschillend en kunnen op grond daarvan tot een geheel verschillende conclusie komen. Dit heeft tot gevolg dat er in veel gevallen onderling verschillen bestaan tussen ouders en leerkrachten in hun perceptie van de schoolproblemen. Wat betreft de perceptie van prestaties, zijn ouders en leerkrachten het in ruim een derde van de gevallen met elkaar eens, echter wat betreft de perceptie van gedragsproblemen bestaat er met nauwelijks 25% overeenstemming de geringste consensus. Een effect van sekse en milieu van de leerling op de perceptie van prestaties, gedrag en de hulpbehoefte van een leerling, als ook de verschillen in de perceptie op deze vlakken, kon niet worden aangetoond. Dit is vanuit theoretisch oogpunt op zijn minst opmerkelijk en verdient in de toekomst nadere bestudering.

Gestuurd door het WSNS-beleid maken scholen vorderingen op het gebied van integrale leerlingzorg en adaptief onderwijs (onderwijs op maat); die vorderingen zijn echter beperkt te noemen (Notebomer, 1998; Peschar & Meijer, 1997). Hoewel leerkrachten beseffen dat zij lesgeven aan leerlingen die onderling sterk van elkaar verschillen wordt nog vaak via klassikale lesmethoden gewerkt. Dat betekent dat alle leerlingen in de praktijk via één algemene aanpak worden onderwezen, ongeacht hun onderlinge verschillen. Daarnaast zijn veel leerkrachten planmatiger gaan werken, hoewel op veel punten nog verbetering mogelijk is. Niet voor alle probleemleerlingen blijken behandelingsplannen te worden opgesteld en bovendien wordt vaak niet gecheckt of de aanpak effect sorteert. Per school en per leerkracht worden op deze manier zeer verschillende werkwijzen gehanteerd.

Adaptief onderwijs en eventuele maatregelen, afgestemd op de individuele leerling, vergen in de eerste plaats betrouwbare kennis omtrent welke leerlingen bepaalde problemen hebben en zo ja, wat daar vervolgens aan moet gebeuren. Leerkrachten en ouders blijken aanzienlijk van mening te verschillen als het gaat om de definiëring van probleemleerlingen. Afhankelijk van de context waarin de leerling wordt waargenomen, thuis of op school, en de normen die daarbinnen gelden worden per school en door de leerkracht en de ouders ver-

schillende leerlingen als probleemleerling aangemerkt. Voor een gericht beleid is daarom meer eenduidigheid in de mate waarin leerlingen als probleemleerling worden aangemerkt noodzakelijk. Dit komt ook de uiteindelijke remediëring van de schoolproblemen ten goede. Het lijkt aannemelijk dat als gevolg van de soms zeer uiteenlopende probleemperceptie van ouders en leerkrachten ook hun visie op de eventueel te nemen maatregelen verschilt. Zo heeft dit onderzoek aangetoond dat met name de consensus wat betreft gedragsproblemen erg gering is. Aangezien gedragsproblemen voor leerkrachten de grootste problemen opleveren in verband met de sfeer in de klas en de ordehandhaving, zijn het niet voor niets deze leerlingen die in combinatie met lage prestaties naar het speciaal onderwijs verwezen worden. Hoewel ouders het vaak niet met deze oplossing eens zijn, blijken leerkrachten uiteindelijk vaak de belangrijkste stem te hebben in de uiteindelijke beslissing wat betreft de maatregelen.

Tegenwoordig lijkt de invloed van ouders te worden versterkt met de invoering van het leerlinggebonden budget, het 'rugzak-model'. Deze vorm van financiering geeft ouders van (gehandicapte) leerlingen die aanvullend op WSNS speciale zorg nodig hebben meer zeggenschap over de inhoud en inrichting van de extra zorg voor hun kind. In overleg met de school wordt een 'plan van aanpak' opgesteld en kunnen ouders het budget besteden binnen het regulier of het speciaal onderwijs. De inbreng van de ouders is daarbij van wezenlijk belang. Afsluitend kan dan ook worden gesteld dat consensus tussen ouders en leerkrachten wat betreft de schoolproblemen van leerlingen én de aanpak daarvan als uitgangspunt moet dienen om onderwijs te bewerkstelligen dat werkelijk 'op maat' is.

Noten

- 1 De volgende items zijn in dezelfde bewoording aan zowel de ouders als de leerkracht gesteld en vormen samen de schaal 'prestatie-inschatting' ($\alpha=.90$ voor leerkrachten, $\alpha=.73$ voor ouders). De items met daarachter (-) zijn voor het berekenen van de betrouwbaarheid gespiegeld.
(Bij) mijn kind/ deze leerling:

- moet de leerstof beperkt worden tot wat echt nodig is (-)
- zijn de prestaties op school nog onvoldoende (-)
- kan goed leren
- kan moeilijk meekomen op school (-)

De volgende items behoren bij de schaal 'gedragsinschatting' ($\alpha=.81$ voor leerkrachten, $\alpha=.63$ voor ouders).

Mijn kind/ deze leerling:

- probeert altijd de eigen zin door te drijven (-)
- is vaak brutaal (-)
- maakt nooit ruzie
- denkt al gauw dat zijn/haar werk af is (-)
- verveelt zich op school (-)
- voelt zich thuis op school
- voelt zich bij de leerkracht goed op zijn/haar gemak
- kan met de leerkracht maar moeilijk opschieten (-)

De volgende items behoren bij de schaal 'hulpinschatting' ($\alpha=.88$ voor leerkrachten, $\alpha=.69$ voor ouders).

Bij mijn kind/ deze leerling:

- is systematisch extra hulp nodig (-)
- moet de leerstof beperkt blijven tot wat echt nodig is (-)
- is aandacht van deskundigen op school vereist (-)
- wordt 'remedial teaching' aangeboden (-)

Literatuur

- Bakker, J. (1984). *De leerkracht en de slechte leerling; een studie naar inhoudelijke en structurele aspecten van de oordeelsvorming van leerkrachten*. Nijmegen: Katholieke Universiteit.
- Bodin-Baarends, C. (1996). *Ouders en leerkrachten over leermoeilijkheden*. Amsterdam: Universiteit van Amsterdam.
- Dekkers, H. & Smeets, M. (1982). *Sekse-ongelijkheid op school*. Nijmegen: ITS.
- Doornbos, K. & Stevens, L.M. (1987). *De groei van het speciaal onderwijs. Deel A: Analyse van historie en onderzoek*. 's-Gravenhage: Staatsuitgeverij.
- Doornbos, K. (1991). School- en leermoeilijkheden: epidemie of systeemeffect? In: K. Doornbos (Ed.), *Samen naar school; aangepast onderwijs in gewone scholen* (pp. 11-30). Nijkerk: Intro.
- Driessen, G. (1996). Allochtone ouders te optimistisch over onderwijskansen eigen kind. *Didaktief*, 9a, 20-21.
- Eck, E. van, Ledoux, G. & Veen, A. (1994). Schoolloopbaan en strijdige verwachtingen. *Pedagogische Studiën*, 71 (1), 16-34.
- Geelen, K. & de Boom-van Duin, F. (1983). *Visies op hulp bij psycho-sociale problemen*. Tilburg: IVA-ITIBA.
- Hargreaves, D.H., Hester, S.K. & Mellor, F.J. (1975). *Deviance in classrooms*. London: Routledge and Keagan Paul.
- Hoeven-van Doornum, A.A. van der (1990). *Effecten van leerlingbeelden en streefniveaus op schoolloopbanen. Een studie naar leerkrachtverwachtingen, leerkrachteffectiviteit en school-effectiviteit*. Nijmegen: ITS/OOMO.
- Hoeven-van Doornum, A.A. van der (1994). Effecten van streefniveaus in het basisonderwijs op schoolloopbanen in het voortgezet onderwijs. *Tijdschrift voor Onderwijs Research*, 19 (3), 204-213.
- Jungbluth, P. (1996). Ouders en de onderwijskansen van hun kinderen. In: K. Autar (Ed.), *Ouders en school* (pp. 49-64). Alphen aan den Rijn: Samson H.D. Tjeenk Willink.
- Jungbluth, P., Peetsma, T. & Roeleveld, J. (1996). *Leerlingprestaties en leerlinggedrag in het primair onderwijs*. Ubbergen: Tandem Felix.
- Jungbluth, P. & Langen A. van (1997). *Risicoleerlingen in het basisonderwijs*. Nijmegen: ITS.
- Kassenberg, A. (1997). *Dubbelfocus op probleemleerlingen. Een studie naar de verschillen tussen ouders en leerkrachten in hun perceptie van schoolproblemen van leerlingen in het basisonderwijs*. Doctoraal scriptie. Groningen: Vakgroep Sociologie.
- Kleij, P. van der (1983). *Zeg na, jij! Over regels, patronen, beurten, selectie en reproductie in het lager onderwijs*. Purmerend: Muusses.
- Liere, C. van (1990). *Lastige leerlingen; een empirisch onderzoek naar sociale oorzaken van probleemgedrag op basisscholen*. Groningen: Rijksuniversiteit Groningen.
- Maas, C.J.M. (1992). *Probleemleerlingen in het basisonderwijs*. Amsterdam, Thesis (ICS).
- Meijer, C.J.W. (1986). *Verwijzing en terugplaatsing. Eerste eindverslag*. Groningen: RION.
- Meijnen, G.W. (1984). *Van zes tot twaalf*. Harlingen: Flevodruk.
- Merton, R.K. (1971). Social problems and social theory. In: R.K. Merton & R. Nisbett (Eds.), *Con-*

- temporary social problems (pp. 793-847). New York: Harcourt Brace Jovanovich, inc.
- Ministerie van Onderwijs en Wetenschappen (1991). *Instellingsbeschikking stuurgroep Evaluatie Weer Samen Naar School*. d.d. 15 april 1991.
- Notebomer, M. (1998). Deelname aan speciaal onderwijs daalt. *Didaktief & School*, 1, 8-10.
- Peeters, M. & Hermanns, J. (1996). *Opvoedingsvoorlichting in de jeugdgezondheidszorg: Het perspectief van de ouders: Een onderzoek naar verwachtingen en wensen van ouders*. Utrecht: SWP.
- Peschar, J.L. & Meijer, C.J.W. (1997). *WSNS op weg. De evaluatie van het 'Weer Samen Naar School'-beleid*. Groningen: Wolters-Noordhoff.
- Peschar, J.L. & Wesselingh, A. (1995). *Onderwijs-sociologie*. Groningen: Wolters-Noordhoff.
- Rispens, J. (1990). *Kinderen met leerproblemen*. Onderwijskundig Lexicon Editie II, A 4200-1A 4200-21.
- Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the classroom: teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart and Winston.
- Waslander, S. & Bosman, R. (1997). Sociale daling en onderwijs. *Beleid en Maatschappij*, 2, 74-84.
- Wit, C. de (1996). *Weer Samen Naar School: een wereld van verschil*. Den Haag: procesmanagement WSNS.

Auteurs

A. Kassenberg is als onderzoeker in opleiding verbonden aan het Interuniversitair Centrum voor Sociaal-wetenschappelijke theorievorming en methodenontwikkeling (ICS) in Groningen.

R. Bosman is docent onderwijssociologie aan de vakgroep Sociologie van de Rijksuniversiteit Groningen en is daarnaast als beleidsmedewerker onderzoek werkzaam bij het ICS.

Correspondentie-adres: A. Kassenberg, ICS/ Vakgroep Sociologie, Grote Rozenstraat 31, 9712 TG, Groningen, e-mail: a.kassenberg@ppsw.rug.nl.

Abstract

School problems of pupils in primary education and the divergent perception of parents and teachers

A. Kassenberg and R. Bosman, *Pedagogische Studiën*, 1999, 76, 307-319.

The question which pupils are problem pupils in school depends to a large extent on the definition of problem pupils of teachers and parents. This study discusses the observation of problems and the question relating to under which conditions parents and teachers differ in their perception of the problems that a pupil has. A sample of 1347 pupils in the fourth grade of primary school showed that parents and teachers are not always capable of judging the problems of pupils and, accordingly, do not consider the very same pupils as problem pupil. Via descriptive statistical methods and regression-analysis it is shown that between teachers and parents little consensus exists with regard to their perception of the problems in school. In particular, parents and teachers differ in their perception of behavioural problems. Somewhat more consensus has been found in the perception of academic achievement.

Leerkrachtbeoordelingen van emotionele en gedragsproblemen bij kinderen van een Pedologisch Instituutsschool

J. H. van Huizen en J. W. Veerman

Samenvatting

Met de Teacher's Report Form (TRF) werd gekeken naar de ernst en aard van emotionele en gedragsproblemen van leerlingen bij opname op de interne school van het Paedologisch Instituut (PI). Van de 332 leerlingen die in de periode 1989-1997 werden opgenomen vulden zowel de PI-leerkrachten (N=215) als de verwijzende leerkrachten (N=223) een TRF in. Blijkens hun scores beoordeelden zij de ernst en diversiteit van de problemen als zeer aanzienlijk vergeleken met normgroepen. Voor meisjes werden op externaliserend gebied ernstiger problemen aangegeven. Leeftijd en type behandeling hadden geen invloed op de ernst van de problemen. De correlatie tussen de scores van verwijzende leerkrachten en PI-leerkrachten bleek redelijk tot goed voor totale problemen en externaliserend gedrag. Over meisjes bestond meer overeenstemming dan over jongens. Verwijzende leerkrachten gaven vooral wat betreft externaliserend gedrag ernstiger problemen aan. In de discussie worden mogelijke verklaringen voor de gevonden resultaten gegeven en wordt ook ingegaan op de bijdrage van dit soort onderzoek bij beslissingen over het toelaten van leerlingen tot het speciaal onderwijs.

1 Inleiding

In dit artikel wordt inzicht gegeven in de prevalentie van emotionele en gedragsproblemen zoals leerkrachten die waarnemen bij kinderen die geplaatst worden op een Pedologisch Instituutsschool. Een Pedologisch Instituutsschool (in het vervolg afgekort met PI-school) is één van de schoolsoorten voor speciaal onderwijs die ons land kent. In de nieuwe indeling die door wijzigingen in de wetgeving van het speciaal onderwijs in de periode 1998-2002 van

kracht wordt, behoort een PI-school tot het zogeheten cluster-4 onderwijs. Naast de PI-scholen maken de scholen voor Zeer Moeilijk Opvoedbare Kinderen (ZMOK) en de scholen voor Langdurig Zieke Kinderen (LZK) deel uit van dit cluster (De Rugzak, 1996; zie ook Van Rijswijk, 1996; en voor de laatste stand van zaken Hover & Haperink, 1998)¹. Het onderwijs binnen dit cluster richt zich op kinderen met ernstige gedragsstoornissen en/of psychiatrische problemen. Hiermee wordt niet alleen bedoeld op agressief of delinquent gedrag, maar ook op druk, angstig of teruggetrokken gedrag. Ter aanduiding van dit brede spectrum hanteren we in dit artikel de term emotionele en gedragsproblemen die in epidemiologisch onderzoek gebruikelijk is (Verhulst, Van der Ende & Koot, 1997).

Aan de hand van eigen onderzoek en dat van anderen laten Pameijer & Zijlmans (1994) zien dat emotionele en gedragsproblemen voorkomen bij zo'n 36 tot 86% van de kinderen die naar een PI-school worden verwezen (zie ook Van den Bergh, 1997; Ten Brink, 1998; Kroes & Van der Doef, 1994; voor naderhand gepubliceerde gegevens). De variatie in deze cijfers kan voor een deel verklaard worden door het gegeven dat sommige PI-scholen zich, vanwege hun taak en functie binnen regionale of landelijke verwijzingskanalen, richten op kinderen met ernstiger problemen dan andere PI-scholen. Andere verklaringen voor deze variatie kunnen liggen in het moment waarop de gegevens werden vastgelegd (bij aanmelding of op een later tijdstip) en de manier waarop de gegevens werden verzameld. Zo werden de meeste gegevens uit de genoemde onderzoeken verzameld via gedragsvragenlijsten die door direct betrokkenen (zoals ouders of leerkrachten) ingevuld konden worden.

In dit onderzoek baseren we ons op gegevens die bij aanmelding door leerkrachten zijn

verstrekt. Het moment van aanmelding lijkt het meest geschikt om inzicht te krijgen in de prevalentie van emotionele en gedragsproblemen. Op dat moment zullen deze problemen het meest zichtbaar zijn. Na toelating op de PI-school is de aanpak er immers op gericht de problemen te laten afnemen. De informatie van de leerkracht is van belang omdat kinderen een groot deel van de dag op school, het tweede milieu, verblijven. De school vertegenwoordigt een belangrijk gebied van functioneren waarbinnen kinderen problemen kunnen hebben die daarbuiten niet opvallen. Verhulst et al. (1997) stellen dat de leerkracht het sociale functioneren van het kind kan observeren en vergelijken met dat van leeftijdgenoten en dat hij daarbij, in vergelijking tot de ouders, niet wordt gehinderd door gezinsinteracties. Daarnaast geven zij aan dat de leerkracht vaak een rol speelt bij verwijzing naar speciaal onderwijs en overige vormen van professionele hulpverlening voor leer- en gedragsproblemen. Het onderhavige onderzoek kan gepositioneerd worden binnen dit verwijzingsproces.

Bij het inwinnen van gedragsbeoordelingen door leerkrachten in het kader van verwijzingen naar speciaal onderwijs dienen zich twee mogelijkheden aan: men kan zich baseren op informatie van de leerkracht van de verwijzende scholen, of op die van de ontvangende scholen (in het dit artikel de PI-school). Onderzoek heeft laten zien dat de keuze van de informant niet om het even is. Uit een meta-analyse van Achenbach, McConaughy en Howell (1987) bleek dat de correlatie tussen informanten die het kind in eenzelfde situatie kunnen observeren redelijk is (rond de 0.60), maar dat de correlatie drastisch daalt als er sprake is van wisselende situaties met wisselende informanten. Zo werden de laagste correlaties (gemiddeld rond de 0.22) gevonden tussen de beoordelingen van kinderen over zichzelf en die van hun ouders of leerkrachten. Deze matige correlaties werden niet toegeschreven aan de onbetrouwbaarheid van het instrument, maar aan de rol van situatie- en beoordelaarskenmerken die inherent zijn aan het beoordelen van gedrag (zie ook Offord, Boyle, Racine, Szatmari, Fleming, Sanford & Lipman, 1996).

Naast verschillen in de rangorde van individuen die verschillende informanten binnen groepen aanbrengen (uitgedrukt in correlatie-

coëfficiënten), kunnen er ook verschillen optreden in het niveau van de beoordelingen (uitgedrukt in gemiddelde scores van verschillende informanten). Ten aanzien hiervan is bekend dat leerkrachten over het algemeen minder problemen signaleren dan ouders (Meijer & Veerman, 1989; Kroes & Van der Doef, 1994; Stanger & Lewis, 1993). Over de relatie (zowel qua rangorde als qua niveau) tussen gedragsbeoordelingen van verwijzende en ontvangende leerkrachten is ons geen onderzoek bekend. In de meta-analyse van Achenbach et al. (1987) bleken de correlaties tussen leerkrachten gemiddeld 0.64. Een recent onderzoek op een PI-school liet onder leerkrachten die in een deeltijdfunctie samen een klas hadden, afhankelijk van het soort problematiek, een overeenstemming van 24% tot 47% zien (Van den Bergh, 1997). In beide publicaties gaat het echter om leerkrachten die kinderen in eenzelfde situatie zien. In de situatie van 'verwijzende' en 'ontvangende' leerkrachten gaat het om leerkrachten die het kind in verschillende situaties zien. Om meer zicht te krijgen op de onderlinge relatie tussen beide soorten beoordelingen zijn beide dan ook in het onderhavige onderzoek ingewonnen.

Meer specifiek worden in dit artikel de volgende onderzoeksvragen gesteld:

1. Wat is de prevalentie van emotionele en gedragsproblemen van de door een PI-school opgenomen kinderen zoals waargenomen door PI-leerkrachten en verwijzende leerkrachten?
2. Zijn er ten aanzien van de voorgaande vraag leeftijd- en sekse verschillen, en verschillen tussen kinderen die verschillende vormen van behandeling krijgen?
3. Hoe verhouden de gedragsbeoordelingen van PI-leerkrachten en verwijzende leerkrachten zich tot elkaar?

2 Methode

2.1 Setting en deelnemers

Het onderzoek vond plaats in de interne school van het Paedologisch Instituut (PI) te Duivendrecht/Amsterdam. Deze school maakt deel uit van dit centrum voor kinder- en jeugdpsychiatrie, waar de zorg zich richt op kinderen van 4-14 en 12-19 jaar en hun ouders. Per jaar worden

52 kinderen van 4-14 jaar met emotionele stoornissen, gedragsstoornissen en ontwikkelingsstoornissen residentieel behandeld. Voor 20 kinderen in die leeftijd is er dagbehandeling. Jaarlijks worden ook ongeveer 300 jeugdigen poliklinisch behandeld. Alle 72 kinderen in residentiële of dagbehandeling (in het vervolg de kliniek genoemd) volgen onderwijs aan de interne PI-school. Daarnaast heeft deze school plaats voor een kleine 30 externe leerlingen. De dataverzameling van het onderhavige onderzoek vond plaats in het kader van het lopende follow-up onderzoek in de PI-kliniek (Veerman, 1990; 1993; Ten Brink, 1998). Om die reden zijn de gegevens van de externe leerlingen niet bij het onderzoek betrokken. Van de leerlingen die naar de PI-school worden verwezen komt ongeveer de helft uit het reguliere basisonderwijs, de andere helft uit het speciaal onderwijs.

In de periode juli 1989 tot en met december 1997 werden er in de kliniek van het Paedologisch Instituut te Duivendrecht/Amsterdam in totaal 332 kinderen opgenomen. In deze periode werd er voor 215 van de 332 kinderen gemiddeld acht weken na opname door de leerkracht van de interne PI-school een Teacher's Report Form (TRF, zie voor een beschrijving par. 2.2) ingevuld. Om de prevalentie van gedragsproblemen te bepalen is het van belang dat de onderzoeksgroep representatief is voor de gehele populatie die in de periode juli 1989 tot en met december 1997 is opgenomen. Voor de representativiteit is in de eerste plaats een goede respons van belang. Deze is met 67% redelijk te noemen. Daarbij moet worden vermeld dat non-respons niet het gevolg is geweest van selectiviteit, maar van een verminderde toelevering en administratieve verwerking van de schoolvragenlijsten in de jaren 1995 en 1997. Daarnaast is het belangrijk dat de groep (N=215) wat betreft de samenstelling op het gebied van leeftijd en de verhouding tussen de seksen en behandeltypen overeen komt met die van de totale populatie (N=332). Er werden op genoemde variabelen tussen de totale populatie van de PI-school en de door de PI-leerkrachten beoordeelde groep gevonden ($p < 0.10$; zie voor een uitgebreide beschrijving Van Huizen, 1998).

Tabel 1
Demografische gegevens van de onderzoeksgroepen

	PI-leerkrachten	Verwijzende leerkrachten
	N=215	N=223
jongens	75%	72%
meisjes	25%	28%
dag	39%	37%
resi	61%	63%
gem. leeftijd	8,4 jr.	9,0 jr
range	3-14 jr.	3-14 jr

Ook door de leerkrachten van de verwijzende scholen werd een TRF ingevuld (zie Tabel 1). Van de 332 kinderen in de kliniek waren dat er 223. Gemiddeld werd door deze leerkrachten 10 weken voor opname een TRF ingevuld. Deze groep bleek eveneens representatief voor de gehele populatie wat betreft sekse, leeftijd en behandeltype. Ook hier was de respons redelijk (65%).

2.2 Meetinstrument

De Teacher's Report Form (TRF; Achenbach, 1991; Verhulst, et al., 1997) wordt door de leerkracht ingevuld voor kinderen van 4 tot en met 18 jaar. De vragenlijst bestaat uit een 'schoolprestatiedeel', dat ook wel 'vaardigheden- of competentiedeel' wordt genoemd, en een 'gedragsproblemedeel'. Alleen dit laatste deel wordt in dit onderzoek gebruikt. Het gedragsproblemedeel bestaat uit 118 items die beantwoord kunnen worden op een driepuntsschaal (helemaal niet, een beetje of duidelijk of vaak van toepassing) en twee open vragen. De items gaan over concrete gedragingen zoals 'is ongehoorzaam op school' en 'neemt regels te strikt, is overaangepast'. In de eerste plaats wordt van het 'gedragsproblemedeel' van de TRF over alle items een totaalscore berekend. Deze geeft een indicatie van de mate waarin het kind emotionele en gedragsproblemen heeft. Daarnaast wordt aangegeven of deze problemen overwegend internaliserend, ook wel overgecontroleerd genoemd, of externaliserend, ook wel ondergecontroleerd genoemd, van aard zijn. Verder worden de volgende gedragsgebieden als kernsyndromen onderscheiden: Teruggetrokken, Somatische klachten, Angstig/Depressief, Sociale Problemen, Denksteurnissen, Aandachtsproblemen,

Delinquent Gedrag en Agressief Gedrag. De acht syndromen kunnen worden weergegeven in een gedragsprofiel. Tenslotte kan er op item-niveau worden gekeken naar specifieke gedragingen.

Om de scores op de TRF te kunnen interpreteren wordt de ruwe score van een kind vergeleken met die van de normgroep. Er zijn vier normgroepen: jongens en meisjes van 4-11 en van 12-18 jaar. Voor dit onderzoek worden de scores uitgedrukt in norm-gerelateerde deviatiescores ('Normative Effect Sizes', *NES*, Ten Brink & Veerman, 1997; zie par 2.4). Voor de interpretatie van de scores zijn er middels afkappunten drie gebieden te onderscheiden. Het klinische gebied, het niet-klinische gebied en het grensgebied. De scores van kinderen die in het klinische gebied vallen komen overeen met kinderen van dezelfde leeftijdsgroep en sekse die naar een instelling voor de Geestelijke Gezondheidszorg zijn verwezen. Op Totale Problemen en op de dimensies Internaliseren en Externaliseren vallen scores boven het 89ste percentiel van de normale populatie in het klinische gebied. Op de syndroomschalen vallen scores boven het 97ste percentiel in dit gebied. De scores van kinderen die niet in het klinische gebied vallen komen daarentegen overeen met leeftijdgenoten van dezelfde sekse die niet zijn verwezen naar een instelling voor geestelijke gezondheidszorg. Als in het grensgebied, dat net onder het klinische gebied ligt, wordt gescoord, geeft dat aan dat er problemen zijn die aandacht behoeven. Op Totale Problemen en op de dimensies Internaliseren en Externaliseren vallen scores boven het 80ste percentiel in het grensgebied. Op de syndroomschalen vallen scores boven het 90ste percentiel in dit gebied. Het grensgebied en het klinische gebied duiden we in dit artikel tezamen aan met probleemgebied.

Uit zowel Amerikaans als Nederlands onderzoek is gebleken dat de psychometrische kwaliteiten van de TRF voldoende zijn. Bij een Nederlandse steekproef bleek er, bij een interval van zes weken, een test-hertest correlatie van 0.90 voor Totale Problemen, van $r=0.81$ voor Internaliseren, van 0.95 voor Externaliseren, en waren er test-hertest correlaties van 0.56 tot 0.94 voor de syndroomschalen. De validiteit van Totale Problemen en Internaliseren en Externaliseren is goed, die van de syn-

droom schalen variërend van voldoende voor Sociale Problemen, Aandachtsproblemen en Agressief Gedrag tot matig voor Denkstoringen, Somatische Klachten en Delinquent Gedrag (zie de Nederlandse handleiding van Verhulst et al., 1997, voor een uitvoerige bespreking van de betrouwbaarheid en de validiteit).

2.3 Procedure

Wanneer een kind wordt aangemeld voor opname op het PI wordt bij de verwijzende school standaard schriftelijke informatie over het kind aangevraagd. De TRF, ook wel schoolvragenlijst genoemd, is hier een onderdeel van. Door de verwijzende leerkrachten ($N=223$) werd gemiddeld 10 weken voor opname een TRF ingevuld. Gemiddeld acht weken na opname werd volgens protocol door de PI-leerkracht een TRF ingevuld ($N=215$).

2.4 Statistische analyses

Ten eerste worden voor het analyseren van de data de norm-gerelateerde deviatiescores ('normative effect sizes', afgekort *NES*) voor de verschillende schalen van de TRF berekend (Ten Brink & Veerman, 1997; Albrecht, Veerman & Ten Brink, 1997). Bij het berekenen van deze score wordt van de ruwe schaal score van een individueel kind de gemiddelde ruwe schaal score van de normgroep afgetrokken en gedeeld door de standaarddeviatie van de normgroep. Door de *NES* te berekenen wordt aangegeven hoeveel standaarddeviaties de score afwijkt van de norm. Daarmee is de *NES* een maat voor de ernst van de emotionele en gedragsproblemen en kunnen groepen worden vergeleken. Een gemiddelde *NES* van nul betekent dat de TRF-score van een bepaalde groep exact overeenkomt met de norm. Bij een gemiddelde *NES* van 1 heeft 84% van de normale populatie een lagere score, bij een *NES* van 2 heeft 97.5% een lagere score en bij een *NES* van 3 is dat 99.9%.

Nadat per schaal de *NES* is berekend wordt met behulp van de *z*-statistic getoetst of deze afwijkt van nul (toets voor één gemiddelde). De invloed van sekse, leeftijd en behandeltype op de hoogte van de *NES* zal met behulp van een variantie-analyse worden bekeken.

Omdat steeds voor de 11 schalen van de TRF zal worden getoetst, wordt over 'deze familie van toetsen' een significantieniveau

van 10% aanvaardbaar geacht. Een Bonferroni-correctie zal worden toegepast voor de bepaling van het significantieniveau per toets. Daarbij zal het overall significantieniveau van 0.10 gedeeld worden door het aantal toetsen, in dit onderzoek steeds 11. Het significantieniveau waartegen wordt getoetst zal dus telkens $0.10/11 = 0.009$ zijn.

Voor een indicatie van de aard van de problemen in termen van diversiteit zal worden gekeken naar de prevalentie van scores in het probleemgebied op de Totale Problemen-schaal, de hoofddimensies internaliseren en externaliseren en de acht kernsyndromen van de TRF binnen de populatie van het PI. De prevalentie wordt vermeld middels percentages scores in het probleemgebied. Als cut-off score voor het probleemgebied wordt conform de handleiding van de TRF (Verhulst et al., 1997) het afkappunt normaal-grensgebied gebruikt (zie ook par. 2.2).

Tenslotte zullen op basis van de NES per schaal de correlaties worden berekend tussen de leerkrachten van de verwijzende scholen en de PI-leerkrachten.

3 Resultaten

Eerst zal door het weergeven van de NES per schaal een gedragsprofiel worden weergegeven op basis van de beoordelingen van PI-leerkrachten ($N=215$). Hierdoor ontstaat zicht op de ernst van de problemen. Vervolgens wordt gekeken naar de diversiteit van de problema-

tiek van deze groep, waarbij wordt gekeken of scores op de verschillende schalen van de TRF wel of niet in het zogenaamde probleemgebied vallen. Op deze manier kan worden nagegaan welke probleemgebieden voorkomen binnen de populatie. Ook wordt gekeken naar de invloed van leeftijd, sekse, en type behandeling op de ernst van de problemen. Dezelfde analyses zullen gedaan worden voor de gedragsbeoordelingen van de verwijzende leerkrachten ($N=223$).

3.1 Beoordelingen door PI-leerkrachten

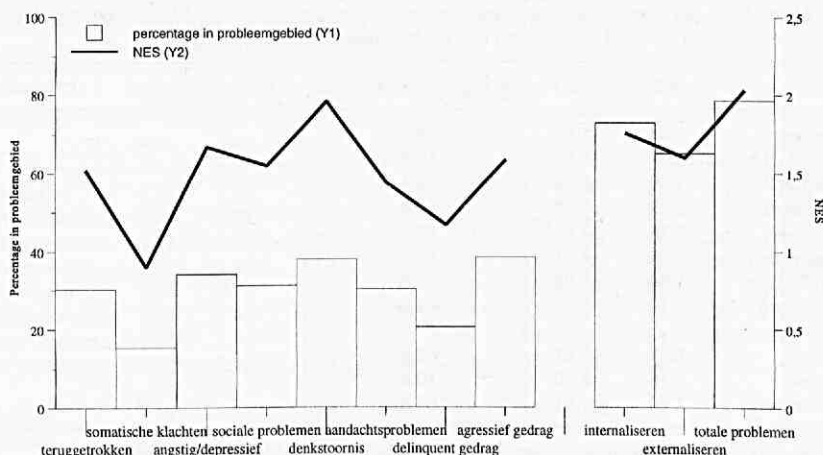
Voor de 11 schalen van de TRF wordt in Tabel 2 de NES weergegeven en worden de percentages scores in het probleemgebied vermeld (zie ook Figuur 1). De NES wijkt op elke schaal significant af van nul ($z > 2.36$; $p < 0.009$). Volgens de PI-leerkrachten hebben de kinderen in de kliniek van het PI dus op alle middels de TRF gemeten gedragsgebieden gemiddeld meer problemen dan kinderen in de normpopulatie. De NES varieert van 0.90 voor Somatische Klachten tot 2.03 voor Totale Problemen. Daar bij een NES van 1.5 door ruim 93% van de normale populatie lager wordt gescoord betekent dit dat de meeste gemiddelde scores van de groep in de bovenste 7% van de verdeling van de scores van de normgroep vallen. Hieruit kan worden opgemaakt dat in vergelijking met de normpopulatie de problematiek van de kinderen uit de PI-kliniek over de hele linie zeer ernstig is. Dit geldt echter in mindere mate voor Somatische Klachten (0.90) en Delinquent Gedrag (1.17).

Tabel 2

Normative Effect Sizes (NES) met standaarddeviaties en percentages scores in het probleemgebied van de PI-leerkrachten ($N=215$) en effecten van leeftijd, sekse en behandeltype

Schaal	NES	SD	% in probleemgebied	lft ¹	sekse ¹	behandeltype ¹
Teruggetrokken	1.52	1.39	30.2	-	-	-
Somatische Klachten	0.90	1.96	15.4	-	-	-
Angstig/Depressief	1.67	1.53	34.4	-	-	-
Sociale Problemen	1.55	1.43	31.2	-	-	-
Denkstoornissen	1.96	2.60	38.1	-	-	-
Aandachtsproblemen	1.44	1.22	30.2	-	-	-
Delinquent Gedrag	1.17	1.73	20.5	-	52	-
Agressief Gedrag	1.59	1.84	38.6	-	72	-
Internaliseren	1.76	1.49	73.0	-	-	-
Externaliseren	1.60	1.88	65.1	-	72	-
Totale Problemen	2.03	1.49	78.6	-	52	-

¹ Percentages verklaarde variantie voor significante verschillen in ANOVA'S ($p < 0.009$)



Figuur 1. Normative Effect Sizes (NES) voor de 11 schalen van de TRF en percentages scores in het probleemgebied voor de groep PI-leerkrachten (N=215)

Als wordt gekeken naar percentages scores in het probleemgebied blijkt dat rond drie kwart van de kinderen ernstige problemen heeft over de hele linie en op externaliserend gebied. Twee derde heeft internaliserende problemen. Met uitzondering van Somatische Klachten en Delinquent Gedrag heeft steeds één derde van de kinderen ernstige problemen op Teruggetrokken, Angstig/Depressief, Sociale Problemen, Denkstoorissen, Aandachtsproblemen en Agressief gedrag. Op grond van deze resultaten kan gezegd worden dat de problematiek zeer divers van aard is.

Uit de ANOVA's (zie Tabel 2) blijkt alleen een effect van sekse op de NES van Delinquent Gedrag ($F(1,213)=9.51$, $p<0.002$), Agressief Gedrag ($F(1,21)=15.80$, $p<0.001$), Externaliseren ($F(1,213)=17.11$, $p<0.001$) en Totale Problemen ($F(1,213)=10.64$, $p<0.001$). Voor meisjes werd op deze schalen hoger gescoord dan voor jongens. De percentages verklaarde variantie liggen tussen de vijf en zeven hetgeen volgens de vuistregels van Cohen (1988, p. 413) betekent dat het effect van sekse op de NES van voornoemde schalen klein is. Leeftijd (t/m 8 jr. en 9 jr. en ouder) en behandeltype (residentiële of dagbehandeling) hebben op geen van de schalen een effect op de NES.

3.2 Beoordelingen door verwijzende leerkrachten

Als naar de beoordeling van de emotionele en gedragsproblemen door leerkrachten van de

verwijzende scholen wordt gekeken (N=223, zie Tabel 3 en Figuur 2) blijkt ook dat de NES op elke schaal significant van nul afwijkt ($z>2.36$; $p<0.009$). De NES varieert van 0.77 voor Somatische Klachten tot 2.40 voor Totale problemen. Ook uit deze resultaten blijkt, met uitzondering van Somatische Klachten, de ernst van de problematiek. Uit de percentages scores in het probleemgebied blijkt dat vier van de vijf aangemelde kinderen ernstige problemen hebben over de hele linie. De helft heeft ernstige problemen op sociaal gebied en agressief gebied. Ook uit deze resultaten blijkt de grote diversiteit aan probleemgebieden binnen de PI-populatie. Er werd een effect van sekse gevonden op Aandachtsproblemen ($F(1,221)=6.92$, $p<0.009$), op Delinquent Gedrag ($F(1,221)=10.61$, $p<0.001$) en op Totale Problemen ($F(1,221)=17.82$, $p<0.009$). Voor meisjes werd op deze schalen hoger gescoord dan voor jongens. De percentages verklaarde variantie liggen tussen de drie en vijf, het effect van sekse is derhalve klein (Cohen, 1988). Ook hier werd geen effect van leeftijd en behandeltype op de NES gevonden.

3.3 Beoordelingen van verwijzende en PI-leerkrachten vergeleken

Tenslotte werd om de beoordelingen van verwijzende en PI-leerkrachten met elkaar te kunnen vergelijken een combinatiegroep gevormd (N=152). Deze groep bestaat uit 152 van de 332 kinderen uit de kliniek waarvoor zowel

Tabel 3

Normative Effect Sizes (NES) met standaarddeviaties en percentages scores in het probleemgebied van de verwijzende leerkrachten (N=223) en effecten van leeftijd, sekse en behandeltype

Schaal	NES	SD	% in probleemgebied	lft ¹	sekse ¹	behandeltype ¹
Teruggetrokken	1.44	1.54	35.5	-	-	-
Somatische Klachten	0.77	1.82	15.3	-	-	-
Angstig/Depressief	1.49	1.58	34.1	-	-	-
Sociale Problemen	2.23	1.70	53.4	-	-	-
Denkstoornissen	2.16	2.65	45.3	-	-	-
Aandachtsproblemen	1.85	1.41	45.7	-	3 ²	-
Delinquent Gedrag	1.64	1.98	37.8	-	5 ²	-
Agressief Gedrag	2.14	2.10	49.3	-	-	-
Internaliseren	1.57	1.47	65.5	-	-	-
Externaliseren	2.16	2.10	70.4	-	-	-
Totale Problemen	2.40	1.61	84.3	-	3 ²	-

¹ Percentages verklaarde variantie voor significante verschillen in ANOVA'S ($p < 0.009$)

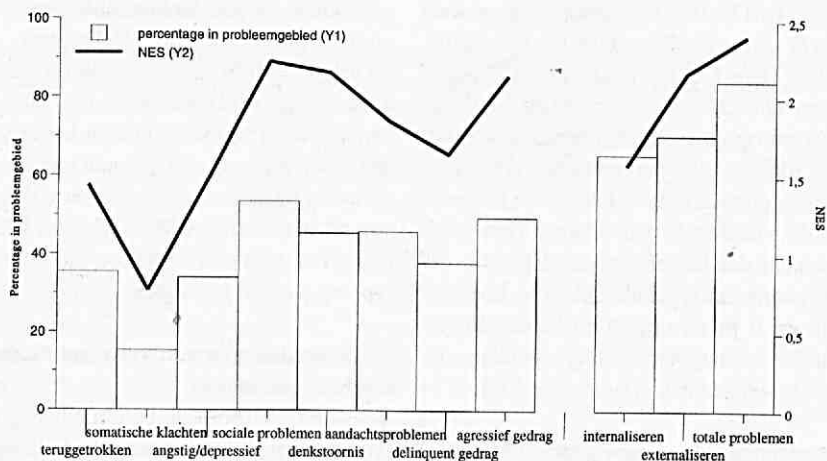
² Voor meisjes werd hoger gescoord dan voor jongens

voor de verwijzende leerkracht van de buitenschool als door de PI-leerkracht van de interne PI-school een TRF werd ingevuld. Ook deze groep bleek representatief voor de populatie in de kliniek in de periode 1989-1997. De respons was 46%, de gemiddelde tijd tussen de afnames was ruim vier maanden.

Om de mate van overeenstemming in kaart te brengen werden de correlaties berekend tussen verwijzende leerkrachten en PI-leerkrachten op basis van de NES (zie Tabel 4). De overeenstemming op Totale Problemen is volgens de vuistregels van Cohen (1988) redelijk tot goed te noemen, die op Externaliseren ronduit goed. Op Internaliseren is de overeenstemming

bescheiden tot matig. Het is opvallend dat over meisjes meer overeenstemming bestaat dan over jongens, waarbij leeftijd een ondergeschikte rol lijkt te spelen. Dit geldt met name voor Agressief Gedrag, Aandachtsproblemen en Sociale Problemen. Over kinderen in residentiële behandeling lijkt meer overeenstemming te bestaan dan over kinderen in dagbehandeling.

Mogelijke verschillen qua niveau van de scores werden getoetst met de paired samples t-toets. Uit de resultaten blijkt dat op de schalen Sociale Problemen, Aandachtsproblemen, Delinquent Gedrag, Agressief Gedrag, Externaliseren en Totale Problemen door verwijzen-



Figuur 2. Normative Effect Sizes (NES) voor de 11 schalen van de TRF en percentages scores in het probleemgebied voor de groep verwijzende leerkrachten (N=223)

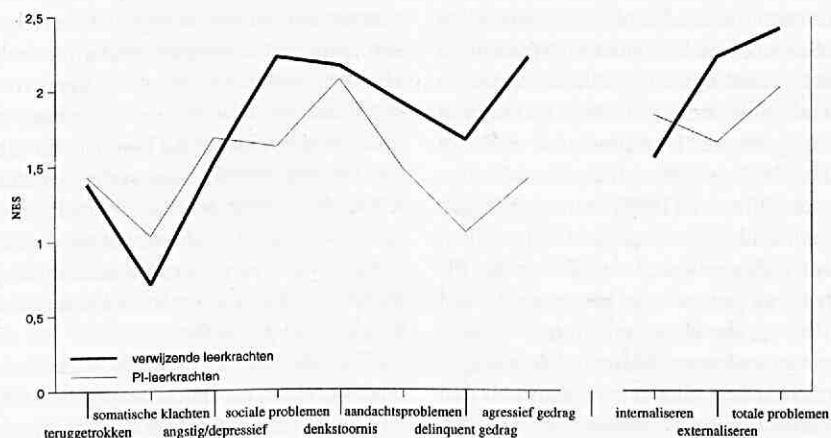
Tabel 4

Normative Effect Sizes (NES) met standaarddeviaties van de PI-leerkrachten en de verwijzende leerkrachten (N=152) en correlaties tussen leerkrachten op basis van NES

Schaal	Verwijzende lk.		PI-lk.		Correlaties ²						
	NES	SD	NES	SD	tot.	j	m	d	r	jk	ok
Teruggetrokken	1.38	1.50	1.43	1.79	.25	.16	.49	.27	.24	.29	.17
Somatische Klachten	0.71	1.81	1.03	2.15	.24	.22	.25	.23	.26	.07	.31
Angstig/Depressief	1.52	1.53	1.68	1.56	.19	.16	.28	.16	.20	.18	.18
Sociale Problemen	2.22 ¹	1.71	1.62	1.52	.43	.33	.57	.32	.48	.39	.47
Denkstoornissen	2.16	2.80	2.07	2.71	.41	.38	.47	.05	.55	.54	.33
Aandachtsproblemen	1.91 ¹	1.45	1.48	1.28	.57	.42	.77	.55	.58	.54	.60
Delinquent Gedrag	1.67 ¹	1.93	1.06	1.58	.40	.34	.43	.34	.43	.36	.44
Agressief Gedrag	2.22 ¹	2.17	1.42	1.56	.69	.58	.80	.62	.74	.72	.68
Internaliseren	1.56	1.45	1.84	1.59	.23	.14	.44	.16	.25	.22	.18
Externaliseren	2.23 ¹	2.15	1.66	1.79	.68	.57	.78	.61	.73	.70	.67
Totale Problemen	2.42 ¹	1.64	2.03	1.56	.52	.33	.73	.40	.57	.53	.54

¹ Significant hogere NES. Paired samples t-toets, $p < 0.009$.

² Tot.: totale groep, j: jongens, m: meisjes, d: dag, r: residentieel, jk: jongere kk., ok: oudere kk.



Figuur 3. Normative Effect Sizes (NES) voor de 11 schalen van de TRF van de groep PI-leerkrachten en de groep verwijzende leerkrachten binnen (N=152)

de leerkrachten een significant hogere NES werd aangegeven ($p < 0.009$; zie Tabel 4 en Figuur 3).

4 Discussie

Wat betreft de beantwoording van de eerste onderzoeksvraag naar de prevalentie van emotionele en gedragsproblemen bij kinderen van een PI-school blijken zowel de leerkrachten van de PI-school als de leerkrachten van de verwijzende scholen bij aanmelding op de interne PI-school een diversiteit aan ernstige

emotionele en gedragsproblemen waar te nemen. Gelet op de Totale Probleemscore als algemene indicator voor het aanwezig zijn van emotionele en gedragsproblemen hebben naar het oordeel van leerkrachten 79% (PI-leerkrachten) of 84% (verwijzende leerkrachten) van de aangemelde leerlingen te kampen met dit soort problemen. De prevalentie van internaliserende problematiek varieert van 66% (verwijzende leerkrachten) tot 73% (PI-leerkrachten), de prevalentie van externaliserende problemen ligt tussen de 65% (PI-leerkrachten) en 70% (verwijzende leerkrachten). De prevalentie van problemen op de acht kernsyndro-

men varieerde van 15% tot 53%. Problemen op de schaal Somatische Klachten kwamen het minst voor, problemen op de schaal Sociale Problemen het meest.

Deze resultaten duiden er op dat het PI de beoogde doelgroep van cluster-4 (kinderen met gedragsstoornissen) in belangrijke mate bedient. Daarbij dient echter de kanttekening gemaakt te worden dat, hoewel de gegevens in dit onderzoek redelijk representatief lijken voor de interne PI-school in Duivendrecht, men zich af kan vragen hoe representatief de gegevens van de PI-school in het onderhavige onderzoek representatief zijn voor 'de' PI-school in het algemeen. Alleen al afgaande op de aard van deze scholen, waarvan er in Nederland acht zijn, liggen populatieverschillen voor de hand. Er zijn Pedologische Instituten die vanuit hun geschiedenis gerelateerd zijn aan de jeugdhulpverlening of de jeugd-geestelijke gezondheidszorg. Andere Pedologische Instituten zijn meer ingebed in het onderwijssysteem en leggen het accent meer op orthodidactiek en (indien nodig) de ambulante begeleiding van kinderen en gezinnen. De variatie aan cijfers in het in de inleiding genoemde overzicht van Pameijer & Zijlmans (1994) is mogelijk een uiting van de hier gesignaleerde verschillen. Een recent onderzoek met de TRF in de PI-school in Nijmegen, waarin aangetoond werd dat ook daar sprake is van ernstige en diverse emotionele en gedragsproblemen, liet ook zien dat deze problemen minder ernstig waren dan die in de PI-school in Duivendrecht (Van den Bergh, 1997).

Aangaande de tweede onderzoeksvraag naar mogelijke verschillen in leeftijd, sekse en behandelvorm werden alleen sekse-effecten gevonden. Door PI-leerkrachten werd voor meisjes hoger gescoord op de syndromen Delinquent Gedrag en Agressief Gedrag en op Externaliseren en Totale Problemen. Door verwijzende leerkrachten werd hoger gescoord voor meisjes op de syndromen Aandachtsproblemen en Delinquent Gedrag en op Totale Problemen. Hoewel de aard van deze syndromen voor de beide typen beoordelaars niet geheel gelijk is lijkt de nadruk te liggen op druk, agressief gedrag.

Een verklaring voor deze bevinding zou kunnen zijn dat, om verwezen te worden, de (relatieve) ernst van deze problemen bij meis-

jes groter moet zijn dan die van jongens. Om zodanig op te vallen dat verwijzing naar een PI-school noodzakelijk wordt, moeten meisjes op externaliserend gebied relatief meer problemen vertonen dan jongens. Dit betekent niet dat meisjes in absolute zin meer druk en agressief gedrag vertonen dan jongens. Daar op de TRF meisjes met een normgroep van meisjes en jongens met een normgroep van jongens worden vergeleken, kan alleen worden gesproken van een afwijking ten opzichte van wat normaal is voor de sekse (en leeftijd). Om verwezen te worden moeten meisjes blijkbaar ten opzichte van hun seksegenoten meer van normaal afwijken dan jongens. Daarnaast is dit externaliserende, drukke gedrag vanuit het maatschappelijke stereotype niet 'des meisjes': meisjes zouden eerder internaliserende problemen hebben. Het zou kunnen dat vooral meisjes bij de PI-school worden aangemeld voor wie andere vormen van hulpverlening niet toereikend bleken. Juist het externaliserende, niet sekse-stereotype gedrag zou de reden kunnen zijn dat deze meisjes moeilijk behandelbaar zijn. Er zou dus mogelijk sprake kunnen zijn van selectie, waarbij meisjes met ernstige externaliserende problemen worden doorverwezen naar de PI-school, of waarbij -vanuit de PI-school gezien- meisjes met externaliserende problematiek eerder worden toegelaten dan meisjes zonder deze problemen.

Ten aanzien van de derde onderzoeksvraag naar de relatie tussen de gedragsbeoordelingen van de PI-leerkrachten en de verwijzende leerkrachten bleek er wat betreft de rangorde binnen groepen een redelijke overeenstemming te bestaan. De gevonden correlatie van 0.52 voor Totale Problemen ligt in dezelfde orde van grootte als correlaties die in andere onderzoeken met leerkrachten werden gevonden (zie Achenbach et al., 1987; Verhulst et al., 1997), zij het dat het in het hier gepresenteerde onderzoek om leerkrachten ging die het kind in wisselende situaties zien. Over externaliserende gedragingen bestaat meer overeenstemming dan over internaliserende gedragingen. Dit is ook conform de in de literatuur gerapporteerde bevindingen. Blijkbaar is externaliserend gedrag gemakkelijker observeerbaar en beter benoembaar dan internaliserend gedrag. Wat betreft het niveau van de scores bleek dat verwijzende leerkrachten in het externaliserende

gebied meer problemen aangaven dan PI-leerkrachten. Dit geldt zowel voor de hoofddimensie Externaliseren als voor de gedragschalen Delinquent Gedrag en Agressief Gedrag. Ook gaven zij meer problemen aan op de schalen Sociale Problemen en Aandachtsproblemen. Wat betreft internaliserende problemen werden geen verschillen gevonden. Dit komt overeen met resultaten uit eerder onderzoek naar de leerkrachtbeoordelingen van de externe PI-school in Duivendrecht (Moesker, 1993). Omdat het de beoordeling betrof rondom de opname lijkt het niet aannemelijk dat het gedrag van de kinderen in de (korte) tijd tussen de afnames van gemiddeld vier maanden wezenlijk was veranderd. Voor een verklaring omtrent de verschillen in de beoordeling wordt derhalve gekeken naar situationele aspecten en informantkenmerken (Achenbach et al., 1987).

Wat betreft de situationele aspecten is bekend dat zowel groepsleiders als leerkrachten bij een klinische populatie de eerste zes maanden *na* opname minder problemen aangaven dan ouders *voor* opname (Ten Brink, 1998; Jansen & Oud, 1990; Kroes & Van der Doef, 1994; Meijer & Veerman, 1989). Hierin zien we een parallel met de gedragsbeoordelingen van verwijzende leerkrachten (vlak *voor* opname) en die van PI-leerkrachten (vlak *na* opname). Het is mogelijk dat kinderen vlak na de opname, nog onder de indruk van de opname en de nieuwe omgeving, een (schijn)aanpassing vertonen en dat ze pas na enige tijd hun eigen ('ware') gezicht laten zien. De TRF-scores van PI-leerkrachten zouden dan deze schijnaanpassing weergeven. De lagere scores van PI-leerkrachten kunnen echter ook een uiting zijn van het feit dat de PI-school beter is toegerust voor de begeleiding van kinderen met emotionele en gedragsproblemen. De klassen zijn kleiner en het onderwijs is meer individueel gericht. Daardoor kent de onderwijsomgeving meer structuur, hetgeen de kans op het ontstaan van externaliserende gedragingen kleiner maakt. Tevens werken leerkrachten van de PI-school samen met groepsleiders en hulpverleners uit diverse disciplines die de kinderen in het kader van residentiële of dagbehandeling begeleiden. Het onderwijs is hiermee ingebed in een bredere behandelingscontext, waardoor in geval van moeilijk in de hand te houden gedrag de expertise binnen het PI snel kan worden geraadpleegd.

Kijkend naar informantkenmerken lijkt het aannemelijk dat verwijzende leerkrachten vanwege hun ervaring met 'gewone' kinderen goed kunnen inschatten wanneer een kind afwijkt van normaal. Het beeld zou echter ook enigszins vertekend kunnen zijn daar zij het gedrag, met name op het gebied van externaliseren, als lastig ervaren. Een hoge score zou er in dat geval op kunnen wijzen dat de leerkracht graag zou willen dat het kind verwezen wordt (Pameijer, Zijlmans & Vorstermans, 1994). Bij PI-leerkrachten is het beeld wellicht vertekend omdat zij hoofdzakelijk kinderen in de klas hebben met ernstige emotionele en gedragsproblemen en bovendien speciaal zijn opgeleid om onderwijs te geven aan kinderen met emotionele, gedrags- en leerproblemen. Het is daardoor goed mogelijk dat voor hen de norm voor afwijkend gedrag hoger is geworden.

Verder onderzoek zou zich kunnen richten op de vraag in hoeverre er bij andere PI-scholen en overige scholen van het cluster-4 eveneens sprake is van emotionele en gedragsproblemen, en tevens of deze problemen sterker zijn dan de emotionele en gedragsproblemen op scholen binnen de andere clusters. Ondanks mogelijke verschillen binnen de scholen van het cluster-4 onderwijs zou dit kunnen leiden tot het opstellen van een voor dit cluster typerend gedragsprofiel, waarmee het zich onderscheidt van andere clusters. Indien dit zou gelukken kunnen TRF-gegevens een rol spelen bij de zogeheten 'toelaatbaarheidsbepaling' zoals die momenteel ten aanzien van de verwijzing naar scholen voor speciaal onderwijs bediscussieerd wordt (Hover & Harperink, 1998). Hierbij gaat het om de vraag of een aangemeld kind al dan niet in aanmerking komt voor een vorm van speciaal onderwijs. Onderzoek naar emotionele en gedragsproblemen van kinderen binnen de verschillende onderwijsclusters kan leiden tot het opstellen van normatieve kaders, die het werk in de praktijk kunnen ondersteunen. Albrecht & Veerman (1998) lieten bijvoorbeeld zien dat ook bij kinderen van een aantal ZMLK-scholen emotionele en gedragsproblemen te constateren waren, maar dat deze minder ernstig waren dan de problemen van de kinderen van de PI-school. Indien op basis van verder onderzoek, gericht op het aantonen van binnen het cluster-4 aanwezige verschillen in emotionele en gedrags-

problemen, nog betrouwbare differentiaties tussen de scholen van cluster-4 te maken zijn, kan dit weer handvatten geven voor een meer verfijnde indicatiestelling naar de betreffende schooltypes. Dit alles in de wetenschap dat er meerdere leerling- en gezinsgebonden factoren zijn die de uiteindelijke indicatie zullen bepalen en ook in de wetenschap dat gegevens van gestandaardiseerde instrumenten als de TRF niet het laatste woord zullen hebben. Indicaties zullen ook op deze basis nooit met 100% zekerheid gesteld kunnen worden, wel mag men hopen met dit soort gegevens afnemende onzekerheid in oordeelsvorming te bewerkstelligen. Ook is het mogelijk dat bij de totstandkoming van normatieve kaders voor de verschillende clusters nieuwe inzichten gaan ontstaan van soorten onderwijs die mogelijk eenzelfde doelgroep bedienen. Of dat, andersom, het licht valt op scholen die geacht werden qua doelgroep bij elkaar te horen maar heel uiteenlopende groepen leerlingen blijken te hebben. Het werken met gestandaardiseerde instrumenten kan vanuit deze visie zowel behandelingsondersteunend als beleidsondersteunend zijn.

Noten

- 1 Andere clusters omvatten scholen voor visueel gehandicapten (cluster-1), scholen voor auditief en communicatief gehandicapten (cluster-2), scholen voor lichamelijk, verstandelijk en meervoudig gehandicapten (cluster-3). Om het beeld ten aanzien van de nieuwe indeling te completeren kan hier nog aan toegevoegd worden dat de huidige voorzieningen voor het MLK-, LOM- en (delen van) het IOBK-onderwijs samen met het reguliere onderwijs ondergebracht worden in de nieuwe Wet op het Primair Onderwijs (voor kinderen van 4 tot 12 à 13 jaar) of de Wet op het Voortgezet Onderwijs (voor kinderen vanaf 12 à 13 jaar).

Literatuur

- Achenbach, T.M. (1991). *Manual for the teacher's report form and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.
- Achenbach, T.M., McConaughy, S.H., & Howell, C.T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232.
- Albrecht, G., & Veerman, J.W. (1998). *Kinderen in het ZMLK-onderwijs in Amsterdam. Screeningsonderzoek naar emotionele en gedragsproblemen*. Duivendrecht: Paedologisch Instituut.
- Albrecht, G., Veerman, J.W., & Brink, L.T. ten, 1997. In H.M. Pijnenburg, C.M. van Rijswijk, A.J.J.M. Ruijsenaars, & J.W. Veerman (red.), *Pedologisch Jaarboek 1997* (pp. 181-191). Delft: Eburon.
- Bergh, C. van den, (1997). *Het gebruik van gedragsvragenlijsten voor leerkrachten. Onderzoek met de TRF op het Paedologisch Instituut Nijmegen*. Doctoraalscriptie Vrije Universiteit Amsterdam.
- Brink, L.T. ten (1998). *De ontwikkeling van kinderen tijdens een periode van klinische jeugdzorg. Status, beloop en prognose*. Dissertatie Vrije Universiteit Amsterdam.
- Brink, L.T. ten, & Veerman, J.W. (1997). *Experimentele klinische normen voor de CBCL*. Duivendrecht: Paedologisch Instituut.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Hover, C., & Harperink, M. (1998). *Verantwoording Uitwerking Onafhankelijke toelaatbaarheidsbepaling 2/3 onderwijs*. Den Haag: Smets+ Hover+.
- Huizen, J.H. van (1998). *TRF onderzoek kliniek PI 1989-1997. De beoordeling van emotionele en gedragsproblemen door leerkrachten van de kinderen in de kliniek van het Paedologisch Instituut Duivendrecht/Amsterdam rondom de opname in de periode 1989-1997*. Doctoraal werkstuk klinische kinder- en jeugdpsychologie, Universiteit van Amsterdam.
- Jansen, M.G., & Oud, J.H.L. (1990). *Residentiële hulpverlening geëvalueerd. Interim rapport 1*. Nijmegen: Instituut voor Orthopedagogiek, Katholieke Universiteit Nijmegen.
- Kroes, G., & van der Doef, P.L.M. (1994). Beoordeling van gedragsproblemen met de CBCL/TRF. In H.M. Pijnenburg, C.M. van Rijswijk, & en J.W. Veerman (red.), *Pedologisch Jaarboek 1994* (pp. 374-379). Delft: Eburon.
- Meijer, R.W.J., & Veerman, J.W. (1989). Gedragsbeoordeling door verschillende informanten. Literatuurbespreking en eigen onderzoeksgegevens. *Kind en Adolescent*, 10, 183-194.
- Moesker, M. (1993). Samenhang tussen cognitieve problematiek, sociaal-emotionele problemen en leerproblemen. Onderzoek naar de populatie

- van een PI-school. Doctoraalscriptie Vrije Universiteit Amsterdam.
- Offord, D.R., Boyle, M.H., Racine, Y., Szatmari, P., Fleming, J.E., Santford, M., & Lipman, E.L. (1996). Integrating assessment data from multiple informants. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35, 1078-1085.
- Pameijer, N.K., & Zijlman, S.W.M. (1994). De CBCL in een aantal pedologische instituten. In H.M. Pijnenburg, C.M. van Rijswijk, & J.W. Veerman, *Pedologisch Jaarboek 1994* (pp. 53-87). Delft: Eburon.
- Pameijer, N.K., Zijlman, S.W.M., & Vorstermans, M.G.C. (1994). Het gebruik van CBCL-lijsten in de praktijk: mogelijkheden en moeilijkheden. *Tijdschrift voor Orthopedagogiek*, 33, 87-103.
- Rijswijk, C.M. van (1996). De rugzak van de PI-leerling. In H.M. Pijnenburg, C.M. van Rijswijk, A.J.J.M. Ruijsenaars, & J.W. Veerman (red.), *Pedologisch Jaarboek 1996* (pp. 9-24). Delft: Eburon.
- Rugzak, De. *Beleidsplan voor het onderwijs aan kinderen met een handicap*. (1996). Zoetermeer: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers, and children on internalizing and externalizing behavior problems. *Journal of Clinical Child Psychology*, 22, 107-115.
- Veerman, J.W. (1990). *De ontwikkeling van kinderen na een periode van klinische jeugdhulpverlening. Een evaluatief-epidemiologisch onderzoek*. Amersfoort: Acco.
- Veerman, J.W. (1993). Evaluatief-epidemiologisch onderzoek in de klinische praktijk. Opzet en kader van follow-up onderzoek in een pedologisch instituut. In H.M. Pijnenburg & L.M.G. Curfs (red.), *Jaarboek onderzoek pedologische instituten 1992* (pp. 49-66). Cadier en Keer: SNZB.
- Verhulst, F.C., van der Ende, J., & Koot H.M. (1997). *Handleiding voor de Teacher's Report Form (TRF)*. Rotterdam: Sophia Kinderziekenhuis.

Auteurs

Drs. J.H. van Huizen is psycholoog en musicus. In het kader van zijn studie klinische kinder- en jeugdpsychologie aan de Universiteit van Amsterdam heeft hij het hier beschreven onderzoek uitgevoerd op het Paedologisch Instituut Duivendrecht/Amsterdam. Momenteel is hij in het kader van de opleiding

tot gezondheidszorgpsycholoog werkzaam bij de afdeling Jeugd en Jongeren van de RIAGG Zuid Nieuw-West te Amsterdam. Tevens is hij als hoofdvakdocent slagwerk Improviserend Musicus verbonden aan de Amsterdamse Hogeschool voor de Kunsten te Amsterdam.

Prof. dr. J.W. Veerman is psycholoog en hoofd van de afdeling Epidemiologisch Evaluatief Onderzoek (EPI) van PI Research te Duivendrecht. Tevens is hij vanwege de stichting De Waarden te Nijmegen als bijzonder hoogleraar verbonden aan de Katholieke Universiteit Nijmegen.

Correspondentie: PI Research, Afdeling EPI, Postbus 366, 1115 ZH Duivendrecht.

E-mail: jw.veerman@psy.vu.nl

Abstract

Teacher ratings of children's emotional and behavioral problems in a Dutch special education sample

J.H. van Huizen, J.W. Veerman. *Pedagogische Studiën*, 1999, 76, 320-331.

From the 332 children admitted to the Pedological Institute in the period 1989-1997 the Teacher's Report Form (TRF) was filled out by the referring teachers (N=223) and the internal teachers in the Pedological Institute (N=215). The Pedological Institute is one of the Dutch type of schools for special education. The TRF-scores indicated serious emotional/behavioral problems as compared to norm groups. Problems appeared both of the internalizing and externalizing type. Although the effect was small more externalizing problems were reported for girls. No effect was found for age and kind of treatment. Referring teachers reported more severe problems than the internal teachers, in particular with respect to externalizing problems. Agreement between both groups of teachers varied, but was high for externalizing behavior. In the discussion some explanations regarding the results are offered and the use of standardized instruments as well as the need to describe populations in special education are discussed.

Achtergronden, kenmerken en opbrengsten van islamitisch basisonderwijs*

G.W.J.M. Driessen en J.J. Bezemer

Samenvatting

Nederland telt momenteel 28 islamitische basisscholen; deze worden voornamelijk bezocht door kinderen van Marokkaanse en Turkse herkomst. Het onderwijs dat op deze scholen wordt gegeven heeft twee doelen, namelijk een hogere onderwijskwaliteit, i.c. leerprestaties leveren dan andere scholen en de kinderen opvoeden in een islamitische traditie. Dit artikel gaat allereerst in op de achtergronden van het islamitische onderwijs. Vervolgens wordt via een empirische studie nagegaan wat de kenmerken zijn van de leerlingen en hun ouders, en wat de prestaties zijn voor taal en rekenen. Daarbij wordt een vergelijking gemaakt tussen 16 islamitische scholen, 16 scholen met een overeenkomstige sociaal-etnische leerlingenpopulatie, en 432 scholen van een landelijk representatieve steekproef (in totaal meer dan 40000 leerlingen). De resultaten laten zien dat leerlingen op islamitische scholen zich met name onderscheiden doordat hun ouders meer belang hechten aan hun geloof en wat minder op de Nederlandse samenleving zijn gericht. Wat de taal- en rekenprestaties van de leerlingen betreft, zijn er nagenoeg geen verschillen tussen islamitische scholen en scholen met een soortgelijke leerlingenpopulatie; wel verkeren de leerlingen op deze scholen in een forse achterstandpositie ten opzichte van de leerlingen op scholen uit de representatieve steekproef.

1 Inleiding

1.1 De stichting van islamitische scholen

De Nederlandse samenleving kent in toeneemende mate een multicultureel karakter. In het onderwijs uit zich dit doordat op veel scholen, en dan met name in de grote steden, kinderen van Turkse, Marokkaanse en Surinaamse origine ver in de meerderheid zijn (Tesser, Van Praag, Van Dugteren, Herweijer & Van der Wouden, 1995). Allochtone groeperingen onderscheiden zich van de autochtone bevolking in hun culturele, linguïstische en godsdienstige voorkeuren, gedragingen en houdingen (Vermeulen, 1996). Evident is dat de Nederlandse samenleving daar nog niet in alle opzichten op is ingesteld. Voor een aantal allochtone ouders sluiten bijvoorbeeld vormgeving, inhoud en opbrengsten van het onderwijs aan hun kinderen onvoldoende aan bij hun wensen: de pedagogisch-didactische aanpak is volgens hen te permissief, er wordt onvoldoende rekening gehouden met en aandacht besteed aan de islam en koran, en het lukt de scholen niet de kinderen tot een acceptabel prestatieniveau te brengen. Gebruikmakend van de grondwettelijke vrijheid van onderwijs hebben zij – of, althans hun zaakwaarnemers (vgl. Alkan, 1996) – daarom eigen islamitische basisscholen gesticht. Dat is niet bepaald steeds van een leien dakje gelopen (cf. Rath, Penninx, Groenendijk & Meyer, 1997). Vanwege het ontbreken van medewerking van lokale en centrale overheden wordt daarom wel de term 'ontmoedigingsbeleid' gehanteerd (cf. Teunissen, 1990). Alhoewel al in 1980 initiatieven daartoe zijn genomen, heeft het tot 1988 geduurd voordat de eerste islamitische basisschool werd gesticht (Landman, 1992). Tot de start van het schooljaar 1997/98 waren er 29 basisscholen op islamitische grondslag; momenteel zijn er nog 28 scholen met in totaal ongeveer 7000 leerlingen, voornamelijk van Marokkaanse en Turkse herkomst.

* Het onderzoek waarover hier wordt gerapporteerd werd gesteund door de Stichting voor Gedragswetenschappen die wordt gesubsidieerd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Het artikel is gebaseerd op het rapport van G. Driessen & J. Bezemer (1999). *Islamitisch basisonderwijs. Schipperen tussen identiteit en kwaliteit?* Nijmegen: ITS.

Over de wenselijkheid van dit onderwijs bestaat veel discussie, zowel onder niet-moslims als onder moslims zelf (Rath, Penninx, Groenendijk & Meyer, 1996). Voorstanders van islamitisch onderwijs streven in hoofdzaak twee doelen na, namelijk versterking van de eigen identiteit, i.c. cultureel-religieuze persoonlijkheidsvorming in de islamitische geest, en verbetering van de onderwijskwaliteit, i.c. de prestaties van de kinderen (Shadid & Van Koningsveld, 1992a, 1992b). Volgens de betreffende ouders schiet het bestaande onderwijs hierin namelijk ernstig tekort. Hun kinderen zaten voorheen doorgaans op zogenaamde zwarte scholen en ondanks alle extra faciliteiten lukte het de meeste van deze scholen niet hun leerlingen in redelijke getale naar hogere types van voortgezet onderwijs te laten doorstromen. De ouders verwachten dat islamitische scholen dat wel zal lukken. Bovendien wordt aangevoerd dat 'eigen' scholen zullen leiden tot vergroting van de ouderbetrokkenheid, wat op zich weer een positief effect op de prestaties zal hebben. Voorstanders zien eigen scholen daarmee als een belangrijk middel voor emancipatie en integratie. Tegenstanders daarentegen vrezen dat islamitische scholen juist leiden tot isolement en segregatie, dat de Nederlandse normen en waarden onvoldoende tot hun recht komen, dat het puur gaat om nationaliteitsscholen, en dat het tot leegloop leidt bij bestaande scholen (Aarsen & Jansma, 1992; Kabdan, 1992; Karagül, 1994; Meyer, 1993).

Met de door Kabdan (1992, 1993) geuite kritiek kunnen de belangrijkste op landelijk en lokaal niveau levende bezwaren tegen aparte islamitische scholen tot twee hoofdpunten worden samengevat. Deze staan in feite diametraal tegenover de twee voornaamste doelstellingen van het islamitische onderwijs. Wat het eerste doel betreft, het verbeteren van de onderwijsprestaties van de kinderen, vreest Kabdan dat als gevolg van het ontbreken van een Nederlandstalig referentiekader de taalachterstand alleen maar zal toenemen. Aangezien er geen autochtone leeftijdgenoten zijn, zullen de kinderen in hun thuistaal communiceren, en niet in het Nederlands. Daar komt bij dat de kinderen ten behoeve van het koran-onderwijs Klassiek-Arabisch moeten leren. Dit is voor alle leerlingen een vreemde taal en betekent dus een ver-

zwaring van het lesprogramma. Ook met het tweede doel van het islamitische onderwijs, het versterken van de eigen identiteit, heeft Kabdan grote problemen. Hij verwacht dat seculier-wetenschappelijke kennis en onderwerpen die tegen de religieuze opvattingen ingaan, niet onderwezen worden. Dit beperkt niet alleen de doorstroommogelijkheden in het voortgezet onderwijs, maar ook de integratie in de Nederlandse samenleving. Wat dit laatste betreft vreest hij ook dat de in gang gezette emancipatie van moslimvrouwen weer zal worden teruggedraaid (cf. Lavrijsen, 1996). Ten slotte verdenkt hij de scholen, i.c. schoolbesturen er van dat onder de dekmantel van cultureel-religieuze persoonlijkheidsvorming nationalistisch-fundamentalistische ideeën worden overgedragen.

Tegen Kabdans vrees met betrekking tot de onderwijsprestaties kan worden ingebracht dat deze situatie niet uniek is voor islamitische scholen, maar in principe geldt voor alle zwarte scholen. Want ook daar is nauwelijks sprake van een Nederlands(talig) referentiekader en zullen de kinderen onderling neigen tot communiceren in de thuistaal. Eveneens zullen ze – na schooltijd – koranles volgen en daar Arabisch moeten leren (vgl. Driessen & Haanstra, 1996). Wat dat laatste betreft kan ook verweten worden naar de situatie in Friesland waar de kinderen verplicht zijn onder schooltijd Fries te leren (en waar de taal- en rekenprestaties overigens relatief laag zijn; vgl. Driessen & Withagen, 1998). Wat de religieuze vorming betreft kan een parallel worden getrokken met de situatie op streng-gereformeerde scholen. Ook op scholen van die richting worden volgens critici zeer traditionele opvattingen overgedragen die op gespannen voet staan met het functioneren in de hedendaagse samenleving.

Het voorafgaande overziend, kan men zich afvragen wat nu precies het unieke is van de bezwaren tegen het islamitische onderwijs en waarom die in het geval van deze nieuwe richting wel worden geuit en in het geval van andere, bestaande richtingen niet. Een en ander in een wat algemener kader plaatsend, lijkt het er – ondanks alle gelijkenissen – niet op dat de vraag naar pro of contra islamitisch onderwijs zonder meer vertaald kan worden naar de vraag naar pro of contra bijzonder onderwijs.

In Figuur 1 vatten we de in de literatuur

PRO

- de grondwettelijke vrijheid van stichting en het recht op bekostiging van bijzondere scholen
- versterking van het zelfbewustzijn, de eigen identiteit en de cultureel-religieuze persoonlijkheid door een islamitische opvoeding
- hogere betrokkenheid van de ouders door betere herkenning of identificatie en minder communicatieproblemen
- verbetering van de schoolresultaten door een betere afstemming op het eigene van allochtone leerlingen
- toename zelfvertrouwen en welbevinden van de leerlingen, en tevredenheid van de ouders over de school door betere aansluiting tussen thuis- en schoolcultuur
- versterking van de integratie en emancipatie van etnische minderheden door betere onderwijsresultaten en een sterkere eigen identiteit van de leerlingen en de instelling van cursussen in de avonden voor de ouders

CONTRA

Nadelige gevolgen van en bezwaren tegen ontbreken van interculturele contacten

- remming van integratie, bevordering van segregatie ('gettovorming', 'apartheid')
- vergroting van de taalachterstand
- strijdig met de doelstellingen van het beleid ten aanzien van intercultureel onderwijs
- ontbreken van de dialoog over culturele gewoontes
- slechte basis voor de brede, algemene ontwikkeling

Nadelige gevolgen van en bezwaren tegen islamitische grondslag

- kennisachterstand door het ontbreken van onderwijs in seculier-wetenschappelijke vraagstukken
- versterking van traditionele sekserolverhoudingen
- verbreiding van afwijkende normen en waarden (remming acculturatie)
- mogelijke banden met (militante) fundamentalistische organisaties
- politiek-ideologische motieven

Overige nadelige gevolgen en bezwaren

- leegloop van bestaande scholen
- aanwakking van racisme
- ongunstige effecten op de doorstroming naar het voortgezet onderwijs
- categorale voorziening
- matige onderwijskwaliteit door onprofessioneel bestuur

Figuur 1. Argumenten pro en contra islamitische basisscholen

genoemde argumenten pro en contra islamitische basisscholen samen (vgl. Driessen & Bezemer, 1999).

1.2 Opbrengsten van islamitisch onderwijs

De in Figuur 1 genoemde aspecten hebben betrekking op argumenten, nauwelijks op feiten. Over de opbrengsten van de islamitische scholen of de manier waarop zij zich onderscheiden van andere (zwarte) scholen was tot voor kort ook nagenoeg geen informatie voorhanden. Hoewel enig onderzoek is verricht, is dit vaak fragmentarisch van aard, heeft het meestal betrekking op case-studies (Aarsen & Jansma, 1992; Meyer, 1993) of op specifieke aspecten als ideologie en bestuurlijke vormgeving (Lammers, 1993; Shadid & Van Koningsveld, 1992a, 1992b). Shadid en Van Konings-

veld (1992a) wijzen er op dat er verschillen zijn qua leerlingenpopulatie, bestuurssamenstelling, en statuten. Het bijzondere karakter van de scholen komt tot uitdrukking in het geven van koran-onderwijs. De vormgeving en invulling van het onderwijs lijkt sterk afhankelijk van de plaats die de school inneemt op het continuüm liberaal - orthodox (Alkan, 1996; Canatan, 1993; Lammers, 1993). Zo wordt op een aantal scholen het lesmateriaal getoetst op de verenigbaarheid met islamitische normen, zitten jongens gescheiden van meisjes, gelden er specifieke gedrags- en kledingvoorschriften, en is er ruimte voor gebedsoefening (Teunissen, 1990). De indruk bestaat dat de meer liberale scholen nauwelijks of niet afwijken van niet-islamitische scholen (Aarsen & Jansma, 1992). De Onderwijsinspectie, die in 1989 en 1990

onderzoek verrichtte naar het functioneren en (vermeende) fundamentalistische karakter van islamitische scholen, verwacht een hogere doorstroom naar het voortgezet onderwijs.

Kwantitatief onderzoek naar de prestaties, houdingen, gedragingen en de gezinssituatie van de betreffende leerlingen is tot op heden enkel door Driessen verricht. In een tweetal studies (1996a, 1996b; zie ook 1997) analyseert hij de relatie tussen meer dan 80 verschillende leerling- en ouderkenmerken enerzijds en het al dan niet bezoeken van een islamitische basisschool anderzijds. Beide studies zijn gebaseerd op gegevens die in het schooljaar 1994/95 in het cohort-onderzoek Primair Onderwijs zijn verzameld bij schooldirecteuren, leerkrachten, leerlingen en hun ouders. Uit de analyses blijkt dat leerlingen op islamitische scholen zich nauwelijks onderscheiden van leerlingen op scholen met een soortgelijke leerlingenpopulatie ('vergelijkingsscholen'). De belangrijkste onderscheidende kenmerken ten aanzien van de houdingen, gedragingen en de gezinssituatie van de leerlingen omvatten de deelname aan koran-onderwijs, de frequentie waarmee huiswerk wordt opgegeven, het belang dat gehecht wordt aan geloof in de opvoeding, het tijdstip van opstaan tijdens werkdagen, en de tevredenheid over de school. Met uitzondering van het tijdstip van opstaan, vallen op alle kenmerken de scores hoger uit bij de leerlingen op islamitische scholen. Wat hun taalprestaties betreft, blijken leerlingen op islamitische basisscholen in de groepen 2 en 4 significant lagere scores te behalen dan leerlingen op vergelijkingsscholen. Vanaf groep 6 scoren beide groepen gelijk. Gedurende de schoolperiode lijken de leerlingen op islamitische scholen dan ook een sterkere ontwikkeling door te maken. Qua rekenprestaties scoren de leerlingen op islamitische scholen in jaargroep 2 nog significant slechter, maar vanaf groep 6 juist beter dan leerlingen op de vergelijkingsscholen. Hier is het verschil in ontwikkeling bijgevolg nog groter. Beide groepen leerlingen blijven overigens ver achter in zowel de taal- als rekenprestaties op leerlingen van de referentiescholen.

Bij deze onderzoeksresultaten dient er op te worden gewezen dat de ontwikkelingen in de taal- en rekenprestaties berusten op cross-sectionele vergelijkingen. Slechts op basis van

longitudinale gegevens kan vast komen te staan in hoeverre deze ontwikkelingen zich bij individuele leerlingen voordoen. Bovendien is onduidelijk of de prestaties van de leerlingen op de islamitische scholen volledig toegeschreven kunnen worden aan de kwaliteit van deze scholen. Doordat de onderzochte scholen op het meetmoment gemiddeld pas vijf jaar functioneerden, hebben met name de leerlingen in de hogere jaargroepen vaak één of meerdere jaren onderwijs op een andere basisschool genoten.

1.3 Een internationaal perspectief

De geschiedenis van islamitische basisscholen in Nederland staat niet op zichzelf. Ook in ons omringende landen als België, Duitsland, Frankrijk, Engeland en Denemarken is het *de jure* mogelijk om scholen op islamitische grondslag te stichten. En hoewel ook in deze landen grote aantallen moslims wonen, functioneren er nog nauwelijks door de overheid erkende en gesubsidieerde islamitische scholen. In Engeland bestond tot voor kort geen enkele door de staat gefinancierde islamitische basisschool, maar waren er wel circa 60 privé-scholen op islamitische grondslag opgericht. Per januari 1998 is daar verandering in gekomen; vanaf dat moment worden er namelijk twee scholen door de overheid bekostigd. In Denemarken zijn 14 islamitische scholen; daarnaast is enkel in Berlijn en Brussel sinds 1989 een door de overheid gesubsidieerde islamitische school te vinden (M. Parker-Jenkins, persoonlijke mededeling, 12 oktober 1998; Pedersen, 1996; Sunier & Meyer, 1997).

Voor deze opmerkelijke verschillen bestaan ten minste twee redenen. In de eerste plaats wordt een pas opgerichte school als gevolg van verschillende wetgevingen niet in alle landen in gelijke mate door de overheid erkend en/of gefinancierd, ook al wordt aan de stichtingsvoorwaarden voldaan. De meeste deelstaten in Duitsland en de landelijke overheid van Frankrijk gaan pas over op (gedeeltelijke) subsidiëring indien enkele jaren na de oprichting de school levensvatbaar blijkt te zijn. Tot die tijd dienen alle kosten door private bronnen te worden gedekt (Sunier & Meyer, 1997). In tegenstelling tot Engeland, kan men in Nederland bij de oprichting van islamitische scholen een beroep doen op principes van gelijkheid en

vrijheid van onderwijs zoals vastgelegd in de Grondwet en vindt volledige bekostiging plaats nadat is voldaan aan de stichtingsvoorwaarden. In België kan men sinds 1974 ook wijzen op de formele erkenning van de islam; hier wordt de huisvesting van scholen op religieuze grondslag echter niet bekostigd door de overheid (Dwyer & Meyer, 1996). In Denemarken worden de islamitische scholen voor 85% gefinancierd (Pedersen, 1996).

In de tweede plaats lijkt de behoefte aan eigen scholen onder de plaatselijke islamitische geloofsgemeenschappen niet altijd even sterk aanwezig te zijn. Dit hangt samen met de organisatie van het onderwijs (vgl. Wagten-donk, 1991). Zo wordt in België op veel grotere schaal dan in Nederland islamitisch godsdienstonderwijs aangeboden (Van Esch & Roovers, 1987; in Nederland gebeurt dat momenteel slechts in vier gemeentes), en zijn net als in Frankrijk voornamelijk openbare scholen gevestigd. Deze omstandigheden lijken de behoefte aan eigen scholen aanzienlijk te drukken. Bovendien heeft dit ook effect op het politieke klimaat. Men vindt dat er via het godsdienstonderwijs voldoende tegemoet gekomen wordt aan de godsdienstige behoeften van de moslims en dat aparte scholen verder niet nodig zijn. In Engeland zijn net als in Nederland de meeste basisscholen juist door de staat gefinancierde scholen op religieuze grondslag, waar veelal ook aan islamitische leerlingen christelijk godsdienstonderwijs wordt gegeven. Het aantal in deze landen gestichte islamitische scholen is dan ook relatief hoog.

Een opvallende overeenkomst tussen de landen is de weerstand die initiatiefnemers van de eerste islamitische scholen hebben ondervonden van grote delen van de ontvangende samenleving (waaronder overigens ook moslims zelf) en de argumenten die daarbij gebruikt worden. Evenals in Nederland, leidde ook in België de komst van de eerste islamitische school tot felle discussies, die voornamelijk via de media werden gevoerd (Dwyer, 1993; Dwyer & Meyer, 1995, 1996; Leman & Renaerts, 1996; Sunier & Meyer, 1997). De Coördinator en Hoge Commissaris Immigrantenbeleid spraken zich daarbij openlijk uit tegen de komst van de school in Brussel. In Engeland lijkt de weerstand bij de overheid het

grootst. Het Britse Ministerie van Onderwijs weigerde zelfs nadat zij door het Hoger Gerechtshof in het ongelijk gesteld was, een 'voluntary-aided' status te verlenen aan een islamitische basisschool in Brent. Met deze status zou de school het recht op overheidssubsidie hebben verworven. De angst voor raciale en religieuze segregatie en de ideologische rol van de gevestigde kerk in Engeland spelen hierbij een belangrijke rol (Cumper, 1990; Parker-Jenkins, 1991; Dwyer & Meyer, 1996). Met de inwerkingtreding van de Education Reform Act uit 1988 is in Engeland overigens een nieuwe financieringsmogelijkheid ontstaan waarvan inmiddels ook islamitische basisscholen gebruik hebben gemaakt (Cumper, 1990; M. Parker-Jenkins, persoonlijke mededeling, 12 oktober 1998). Ook de motieven die voor de oprichting van islamitische basisscholen in Nederland zijn genoemd, komen grotendeels overeen met de motieven van de Britse moslims. Het belangrijkste ongerief is ook bij de Britse ouders het tekort aan mogelijkheden om hun geloof te belijden volgens bepaalde islamitische principes als het vasten en de gebedsoefening (Parker-Jenkins, 1991). Daarnaast vormen gemengd zwemmen en gymmen, seksuele voorlichting en kledingvoorschriften aanleiding om naar alternatieve scholen te zoeken (net als in Denemarken, vgl. Pedersen, 1996).

Uit het voorgaande kan worden geconcludeerd dat Nederland de meeste, door de overheid gesubsidieerde islamitische basisscholen telt van geheel Noord-West Europa. Dit is in de eerste plaats het gevolg van een relatief gunstige en nauw omschreven wet- en regelgeving. In de tweede plaats is de behoefte aan islamitische scholen groot, omdat er nauwelijks islamitisch godsdienstonderwijs in het reguliere basisonderwijs wordt aangeboden en islamitische leerlingen daar veelal christelijk godsdienstonderwijs dienen te volgen.

2 Probleemstelling

Hierboven hebben we een overzicht gegeven van argumenten die door voorstanders worden aangevoerd met betrekking tot de stichting van islamitische scholen. Uit de literatuur (vgl. Driessen & Bezemer, 1999) kan worden afge-

leid dat er, daarbij aansluitend, bepaalde verwachtingen leven ten aanzien van dit onderwijs. Hierbij zet men islamitische scholen af tegen zoveel mogelijk vergelijkbare scholen, in de praktijk andere zwarte scholen. Men verwacht dat het onderwijs op islamitische scholen onderwijskundig beter is afgestemd op de specifieke etnisch-culturele leerlingenpopulatie en bovendien dat er ook op religieus gebied meer aansluiting is tussen het school- en thuismilieu. Daardoor zullen de leerlingen zich beter thuis voelen op school en ook meer zelfvertrouwen krijgen. Tevens is de verwachting dat, omdat de ouders een bewustere keuze hebben gemaakt voor deze scholen, zij ook meer betrokken zullen zijn bij het onderwijs aan hun kinderen. Ten slotte verwacht men dat vanwege de adequatere afstemming tussen onderwijs en specifieke leerlingenpopulatie en de grotere betrokkenheid van de ouders het prestatieniveau van leerlingen op islamitische scholen hoger zal liggen en ook dat de ouders tevredener zullen zijn met het onderwijs.

In hoeverre deze verwachtingen in de praktijk ook waar worden gemaakt is onduidelijk. Niet alleen zijn de meningen over het islamitische basisonderwijs sterk verdeeld, ook zijn ze nauwelijks gebaseerd op empirische gegevens. Waar het vooralsnog in voldoende mate aan ontbreekt zijn kwantitatieve en kwalitatieve gegevens over het onderwijs zelf: de achtergronden van de leerlingen en hun ouders, de pedagogisch-didactische vormgeving van het onderwijs, en het rendement in termen van onderwijsprestaties.

In het onderzoek waarvan we hierna verslag doen richten we ons op gegevens die direct te maken hebben met de leerlingen: hun achtergronden en prestaties.¹ We willen de volgende *onderzoeksvragen* beantwoorden:

- Wat zijn de achtergronden en prestaties van kinderen op islamitische scholen?
- Waarin verschillen zij van hun jaargenoten op overigens vergelijkbare, niet-islamitische scholen?

Meer specifiek willen we nagaan welke verschillen er bestaan in het thuismilieu van de leerlingen, welke verschillen er zijn qua intelligentie en taal- en rekenprestaties, en welke verschillen er zijn wat betreft het welbevinden en zelfvertrouwen van de leerlingen.

In het onderzoek maken we een vergelijking

tussen de leerlingen op islamitische scholen en leerlingen op scholen met een qua socio-etnische achtergrond soortgelijke leerlingenpopulatie. Hiervoor maken we gebruik van gegevens die in het schooljaar 1996/97 zijn verzameld in het kader van het cohortonderzoek Primair Onderwijs. Omdat we ook beschikken over eerdere gegevens uit dat cohort uit 1994/95, stelt dit ons in staat om een rechtstreekse longitudinale vergelijking te maken en de eerdere resultaten te toetsen op hun stabiliteit.

3 Methode

3.1 Steekproeven

De basis voor het onderzoek wordt gevormd door gegevens uit het cohort-Primair Onderwijs ('PRIMA'). Dit cohort is in 1994/95 van start gegaan met een eerste dataverzamelingsronde bij 55000 leerlingen uit de groepen 2, 4, 6 en 8 van 692 basisscholen. Er is informatie verzameld bij directies, leerkrachten en ouders en er zijn verschillende toetsen afgenomen bij de leerlingen (voor een uitgebreide verantwoording van PRIMA-1, zie Van Langen, Vierke & Robijns, 1996). In 1996/97 heeft een tweede dataverzamelingsronde plaatsgevonden bij 622 scholen met circa 57000 leerlingen. Het merendeel van de scholen uit 1994/95 werkte ook in 1996/97 mee. Kenmerkend voor het cohortonderzoek is dat leerlingen bij wie in 1994/95 toetsen zijn afgenomen, in 1996/97 opnieuw zijn getoetst. Uiteraard geldt dit alleen voor die leerlingen die in 1994/95 in groep 2, 4 en 6 zaten en die in 1996/97 bij een normale doorstroom – dus zonder doubleren – in groep 4, 6 en 8 terecht zijn gekomen. De groep 8 leerlingen uit 1994/95 waren in 1996/97 in principe doorgestroomd naar het voortgezet onderwijs, en de groep 2 leerlingen uit 1996/97 zijn per definitie nieuwe leerlingen. In totaal is 43.6% van de leerlingen uit de eerste meting ook bij de tweede meting betrokken geweest. De opzet van PRIMA-2 staat uitvoerig beschreven in Driessen, Van Langen, Portengen en Vierke (1998).

De steekproef van scholen uit PRIMA-2 staat in dit artikel centraal. Van die scholen vormen er 432 een landelijk representatieve steekproef. Deze zogenaamde referentiesteek-

proef telt 39450 leerlingen. Aan PRIMA nemen 16 islamitische scholen deel; dit is ruim de helft van het totaal. Opgeteld zitten er 1670 leerlingen in de groepen 2, 4, 6 en 8. In de analyses worden deze leerlingen vergeleken met leerlingen op eveneens 16 scholen met een overeenkomstige leerlingenpopulatie, de zogenaamde vergelijkingsscholen. Op deze scholen zitten 1834 leerlingen. Als criterium bij de vorming van de groep van vergelijkingsscholen is gebruik gemaakt van de zogenaamde schoolscore van het Ministerie van OCW. Deze score, die is gebaseerd op een combinatie van de opleiding, het beroep en de etnische herkomst van de ouders van de leerlingen, geeft een indicatie van de sociaal-etnische achterstand van een school. We hebben voor dit criterium gekozen, omdat juist deze achterstandspositie en de lagere prestaties die daar doorgaans het gevolg van zijn, een van de redenen was om aparte islamitische scholen te stichten. Analyses die we uitvoerden wezen uit dat er geen verschillen zijn tussen de schoolscores van de 16 aan PRIMA deelnemende islamitische scholen, de overige 13 islamitische scholen in Nederland en de 16 vergelijkingsscholen. Vervolgens hebben we ook nog gekeken naar de specifieke etnische samenstelling van de scholen. Uit deze analyses bleek dat er geen verschillen waren tussen de PRIMA islamitische scholen en de overige islamitische scholen, maar dat er wel meer Marokkaanse en minder Surinaamse leerlingen op de islamitische scholen zaten dan op de vergelijkingsscholen.

3.2 Onderzoeksinstrumenten

In PRIMA is een grote verscheidenheid aan schriftelijke onderzoeksinstrumenten gehanteerd. Voor het onderhavige onderzoek zijn de volgende van belang:

- Het *groepsformulier* bevat administratieve gegevens zoals naam van de leerling, geboortedatum, geslacht, en geboorteland en opleidingsniveau ouders.
- De groepsleerkrachten hebben voor elk van hun leerlingen een zogenaamd *leerlingprofiel* ingevuld. Hierin zijn vragen opgenomen over het gedrag en de houding van de leerling, het thuis klimaat en enkele onderwijsbijzonderheden.
- Bij de ouders van de leerlingen uit de groep 2 is een uitgebreide *oudervragenlijst* afge-

nomen. Hierin is geïnformeerd naar een aantal achtergrondkenmerken van de ouders, het kind en de gezinssituatie.

- Bij de leerlingen van groep 4, 6 en 8 zijn twee *non-verbale intelligentietests* afgenomen, namelijk Figuren samenstellen en Exclusie. Er is met name voor non-verbale tests gekozen om zo eventuele test- en itembias jegens allochtone leerlingen te voorkomen.
- Wat de cognitieve capaciteiten betreft is in elke jaargroep een *taal- en rekentoets* afgenomen. Voor groep 2 waren dat Begrippen en Ordenen uit het Cito-Leerlingvolgsysteem; deze geven een indicatie van het niveau voorbereidend lezen en rekenen. Begrippen omvat 60 en Ordenen 42 meerkeuze-items. Voor de groepen 4, 6 en 8 zijn speciaal voor dit onderzoek door het Cito ontwikkelde taal- en rekentoetsen afgenomen. Voor taal betreft het voor de groepen 4, 6 en 8 respectievelijk 60, 65 en 64 meerkeuze-opgaven; voor rekenen gaat het voor alle drie de groepen om 40 meerkeuze-opgaven. De scores voor de toetsen zijn via een kalibratieprocedure zodanig geschaald, dat ze binnen de vakgebieden taal, respectievelijk rekenen over de vier jaargroepen heen vergelijkbaar zijn (vgl. Vierke, 1995).
- Voor de leerlingen uit groep 8 is door de groepsleerkracht of directie van de school tevens een *uitstroomformulier* ingevuld. Hierin wordt onder meer gevraagd naar het advies voor voortgezet onderwijs dat de betrokken leerling van de directie heeft ontvangen en de scores op de Cito-Eindtoets basisonderwijs.
- De leerlingen uit groep 6 en 8 hebben zelf een *schoolwelbevinden-vragenlijst* ingevuld. Deze bestaat uit 10 opvattingen over de verhouding school-leerling; deze zijn samengevoegd tot de dimensies 'welbevinden' en 'zelfvertrouwen'.

Voor een gedetailleerde verantwoording van deze instrumenten en kenmerken verwijzen we naar Driessen, Van Langen, Portengen en Vierke (1998).

4 Resultaten

4.1 Een eerste vergelijking

Hierna volgen de resultaten van de variantie-analyses die we uitvoerden. Centraal daarbij staat steeds de vergelijking van de gemiddelden van de leerlingen van de islamitische scholen met die van de vergelijkingsscholen. Om te toetsen of er een relevant verschil is richten we ons op de Eta^2 -coëfficiënt. We spreken van relevantie bij $Eta^2 > .02$, ofwel wanneer meer dan 2% variantie wordt gebonden.² Om een idee te geven van de relatieve positie van de leerlingen op de twee te vergelijken categorieën, presenteren we tevens de gemiddelden van de referentiecategorie (de landelijk representatieve steekproef).

Het groepsformulier

Het groepsformulier bevat een serie gegevens uit de schooladministraties. Een eerste kenmerk is de gezinssamenstelling; hierbij gaat het om eenoudergezin versus volledig gezin. Wat het geboorteland van de vader, respectievelijk moeder betreft onderscheiden we twee categorieën: de ouders zijn geboren in Marokko, Turkije of Suriname versus de ouders zijn elders geboren. Dit onderscheid hebben we gemaakt in verband met de belangrijkste doelgroepen van islamitisch onderwijs. Van zowel de vader als de moeder is ook het opleidingsniveau bekend; hierbij zijn als categorieën onderscheiden: (1) maximaal lager onderwijs, (2) maximaal lbo, (3) maximaal mbo, (4) hbo/wo. Bij het geslacht gaat het om het percentage jongens. De leeftijd is uitgedrukt in jaren. Met betrekking tot de verblijfsduur van het kind is een onderscheid gemaakt tussen leerlingen die altijd in Nederland hebben gewoond en leerlingen die niet altijd in Nederland hebben gewoond.³

De analyses laten allereerst zien dat er slechts minimale verschillen zijn tussen de jaargroepen. Vervolgens blijkt dat de verschillen tussen islamitische en vergelijkingsscholen over het algemeen ook zeer gering zijn. In het oog springt vooral het gegeven dat leerlingen op islamitische scholen aanzienlijk vaker in volledige gezinnen opgroeien dan leerlingen op vergelijkingsscholen. Grofweg genomen gaat het bij de islamitische scholen om circa 95% en bij de vergelijkingsscholen om 80%

van de gezinnen ($Eta^2 \approx .05$). Verder geldt voor groep 8 van de islamitische scholen dat er iets meer vaders in Turkije of Marokko geboren zijn ($Eta^2 = .03$). Nadere analyse wijst uit dat deze twee gegevens gerelateerd zijn: in de vergelijkingscategorie zitten met name meer Suri-naamse gezinnen, en die groepering kent relatief veel eenoudergezinnen (alleen moeders). Verder zijn er geen relevante verschillen.

Het leerlingprofiel

Via door de groepsleerkrachten ingevulde leerlingprofielen is informatie verkregen over: deelname aan de lessen Nederlands als Tweede Taal (NT2), het volgen van Onderwijs in Allochtone Levende Talen (OALT), het opkrijgen van huiswerk, en of de leerling het vorige jaar op dezelfde school heeft gezeten. Dit laatste punt kan van belang zijn, omdat een aantal islamitische scholen pas sinds enkele jaren bestaat, en leerlingen dus later vanuit een andere school in een van de hogere jaargroepen zijn ingestroomd. Verder bevat het profiel gegevens over de inschatting van de leerkrachten over het vermoedelijke type voortgezet onderwijs waarvoor de leerling geschikt is ('prognose v.o.'). De categorieën hierbij variëren van (1) ivbo/vbo tot (5) vwo. Ten slotte zijn op basis van een serie items uit het profiel een aantal cognitieve en niet-cognitieve kenmerken geconstrueerd.⁴ De cognitieve factoren omvatten: cognitieve capaciteiten (hoe goed kan hij/zij leren?) en onderpresteren (de leerling presteert onder zijn/haar kunnen; in feite een negatief attribuut). De niet-cognitieve factoren, voornamelijk gedrags- en houdingskenmerken, zijn: zelfvertrouwen; de werkhouding ten opzichte van schoolse zaken; sociaal gedrag (hoe is de leerling in zijn/haar omgang met anderen?); welbevinden (voelt de leerling zich op zijn/haar gemak op school?); onderwijs-ondersteunend thuisklimaat. De scores op deze kenmerken hebben een range van 1 tot 5, en kunnen inhoudelijk worden geïnterpreteerd in termen van negatief tot positief.

Evenals bij het groepsformulier bleken zich ook bij de leerlingprofielen nauwelijks verschillen voor te doen tussen de vier jaargroepen. In groep 2 en 4 wordt meer NT2 gevolgd en in groep 6 en 8 wordt meer huiswerk gegeven. Wanneer we dan naar de verschillen tussen islamitische en vergelijkingsscholen kij-

ken, valt op dat op islamitische scholen in de groepen 2 en 4 ongeveer dubbel zo veel leerlingen NT2 volgen ($Eta^2 \approx .05$); voor OALT geldt dat in alle jaargroepen ($Eta^2 \approx .25$). OALT wordt door bijna alle leerlingen op islamitische scholen gevolgd. In verhouding tot de leerlingen uit de referentiecategorie wordt er op islamitische en vergelijkingsscholen aanzienlijk meer NT2 en OALT gevolgd, maar dat ligt gezien de bijna geheel allochtone leerlingenpopulatie van beide schoolcategorieën ook voor de hand. Duidelijk wordt ook dat op islamitische scholen wat meer leerlingen uit groep 4 en 6 in het voorafgaande jaar nog op een andere school hebben gezeten ($Eta^2 \approx .03$). Wat de overige cognitieve en non-cognitieve kenmerken betreft zijn er geen verschillen.

De oudervragenlijst

De ouders van de leerlingen uit groep 2 is gevraagd een uitgebreide vragenlijst in te vullen over de gezinssituatie van de leerling. De aldus verkregen gegevens zijn onderverdeeld in een viertal inhoudelijke categorieën:

- *gezinsstructurele kenmerken*: gezinssamenstelling, aantal kinderen, en voor beide ouders: opleiding, geboorteland, nationaliteit, en inkomen;
- *gezinsculturele kenmerken*: zelftoerekening van de taal-, cultuur- en geloofsgemeenschap van vader, moeder en kind, en voor beide ouders: beheersing Nederlands, informele spreektaal en leesgedrag;
- *gezinspedagogische kenmerken*: voor beide ouders het belang dat ze hechten aan het geloof in de opvoeding, het belang dat ze

hechten aan het leren van de eigen taal, ouder-kindactiviteiten, bezoek ouderbijeenkomsten, en tevredenheid met school;

- *leerlingenkenmerken*: bezoek kinderdagverblijf en peuterspeelzaal, deelname aan ouder-kindprogramma's, aspecten met betrekking tot 'schoolrijpheid', slaaptijden, tv-kijken, attitude tegenover school, relatie met leerkracht, en spreektaal.

Vanwege het grote aantal variabelen beperken we ons bij de presentatie tot die kenmerken waarop er relevante verschillen bestaan tussen de islamitische en vergelijkingsscholen. Tabel 1 geeft daarvan een overzicht per schoolcategorie. Enkele van de in de tabel vermelde variabelen verdienen een korte toelichting. De hoogst voltooide opleiding van de ouders kan variëren van (1) maximaal l.o. tot (7) w.o. Bij de 'spreektaal ouders Nederlands' gaat het om het percentage ouders dat onderling meestal Nederlands spreekt (en niet een buitenlandse taal). Bij de kenmerken 'moeder/vader islam' geven de percentages het aandeel moeders en vaders weer dat zichzelf tot de islamitische geloofsgemeenschap rekent. Bij 'Nederlandse cultuurgemeenschap moeder/vader/kind' duiden de percentages op het aandeel moeders en vaders dat zichzelf en hun kind tot de Nederlandse cultuurgemeenschap rekent. Het belang van het geloof in de opvoeding is aangegeven op een schaal die varieert van (1) totaal niet belangrijk tot (5) zeer belangrijk. Bij de 'spreektaal kind Nederlands' wordt het percentage domeinen (met vader, moeder, broers/zussen, vrienden/vriendinnen) vermeld waarin het kind Nederlands spreekt.

Tabel 1

De gegevens van de oudervragenlijst naar islamitische, vergelijkings- en referentiescholen (alleen relevante verschillen; gemiddelden)

	islam.	verg.	Eta^2	ref.
aantal kinderen	3.1	2.7	.04	2.4
opleiding vader	3.1	2.4	.03	3.8
geboorteland gezin (% beiden NL)	.5	7.5	.03	85.8
% spreektaal ouders Nederlands	9.5	24.3	.04	72.4
% moeder islamitisch	99.5	66.7	.19	5.8
% vader islamitisch	99.0	70.2	.16	5.8
% NL cultuurgemeenschap moeder	5.4	18.9	.04	91.6
% NL cultuurgemeenschap vader	3.2	16.4	.05	92.1
% NL cultuurgemeenschap kind	5.9	26.7	.08	93.1
belang geloof bij opvoeding, moeder	4.6	3.9	.11	2.9
belang geloof bij opvoeding, vader	4.7	4.0	.11	2.8
uren tv-kijken, wekdagen	2.2	2.8	.05	1.5
% spreektaal kind Nederlands	42.7	62.6	.08	96.3

Bij de *gezinsstructurele* kenmerken bestaan er op een drietal punten relevante verschillen tussen islamitische en vergelijkingsscholen. Ten eerste hebben de ouders van leerlingen op islamitische scholen gemiddeld meer kinderen (3.1 vs. 2.7). Ten tweede zijn de vaders van deze leerlingen naar eigen zeggen wat hoger opgeleid (mavo vs. lbo/mavo). Ten derde ligt het percentage ouders dat allebei in Nederland is geboren bij de vergelijkingsscholen 7% hoger dan bij de islamitische scholen. Van de ouders uit de laatstgenoemde schoolcategorie is slechts .5% allebei in Nederland geboren. De grootste verschillen doen zich echter voor bij de *gezinsculturele* en *gezinspedagogische* kenmerken. Uit de tabel blijkt in de eerste plaats dat bijna alle ouders van leerlingen van islamitische scholen (99-100%) zichzelf tot de islamitische geloofsgemeenschap rekenen, terwijl dit percentage bij de vergelijkingscategorie 67-70% bedraagt. Bij de moeders bindt dit verschil 19%, en bij de vaders 16% van de variantie. Ten tweede zijn er bij de islamitische scholen nauwelijks ouders die zichzelf of hun kind tot de Nederlandse cultuurgemeenschap rekenen (3-6%), terwijl 16-27% van de ouders uit de vergelijkingscategorie dit wel doet. De gebonden variantie bij dit verschil bedraagt 4-8%. Ten derde spreekt 10% van de ouders van islamitische scholen onderling meestal Nederlands, tegenover 24% van de ouders uit de vergelijkingscategorie. Ten slotte blijkt dat de ouders uit de islamitische scholencategorie er veel sterker aan hechten dat hun kind volgens het eigen geloof wordt opgevoed ($Eta^2 = .11$). Wat de *leerlingkenmerken* betreft komt naar

voren dat de kinderen op islamitische scholen wat minder vaak tv-kijken dan kinderen op vergelijkingsscholen. Wel is het zo dat ze aanzienlijk minder frequent Nederlands spreken.

Samenvattend kan worden vastgesteld dat de 'eigen' (d.w.z. uit het herkomstland) religie, taal en cultuur binnen gezinnen uit de islamitische scholencategorie een sterkere positie innemen dan binnen gezinnen uit de vergelijkingscategorie, terwijl gezinnen uit de vergelijkingscategorie sterker georiënteerd lijken op de Nederlandse samenleving.

Intelligentie

Wat de scores op beide non-verbale intelligentietests betreft kunnen we kort zijn. Voor geen van de drie jaargroepen waarin deze tests zijn afgenomen zijn er verschillen tussen de leerlingen op islamitische scholen en de leerlingen op vergelijkingsscholen. Wel zijn er verschillen met de leerlingen uit de referentiesteekproef. Voor de twee subtests gaat het om .5 tot 2.2 punten (bij standaarddeviaties van 3.1, resp. 3.3).

Taal- en rekenprestaties

Tabel 2 geeft de taal- en rekenprestaties weer. Omdat de toetsscores zijn gekalibreerd kunnen ze per vak over de jaargroepen heen worden vergeleken. In de tabel staan de gemiddelde scores en standaarddeviaties van de leerlingen op de islamitische en vergelijkingsscholen samen. De gegevens in de tabel betreffen cross-sectionele data: ze zijn op één moment, in het schooljaar 1996/97, gemeten. In de tabel voeren we ook een quasi-longitudinale vergelij-

Tabel 2

De taal- en rekenprestaties op islamitische, vergelijking- en referentiescholen: cross-sectioneel (gemiddelden)

	islam.		verg.		SD	Eta ²	ref.	
TAAL								
groep 2	940		941		26	.00	972	
groep 4	997	(+57)	1001	(+60)	29	.01	1041	(+69)
groep 6	1048	(+51)	1047	(+46)	27	.00	1080	(+39)
groep 8	1087	(+39)	1088	(+40)	33	.00	1122	(+42)
		(147)		(147)				(150)
REKENEN								
groep 2	861		859		54	.00	900	
groep 4	1035	(+174)	1035	(+175)	61	.00	1075	(+175)
groep 6	1124	(+89)	1122	(+88)	39	.00	1143	(+68)
groep 8	1204	(+80)	1184	(+62)	50	.04	1211	(+68)
		(343)		(325)				(311)

Tabel 3

De taal- en rekenprestaties op islamitische, vergelijkings- en referentiescholen: longitudinaal (gemiddelden)

	TAAL				REKENEN			
	islam.	verg.	Eta ²	ref.	islam.	verg.	Eta ²	ref.
groep 2 ('94)	937	944		975	849	875		912
groep 4 ('96)	996	1005		1043	1039	1048		1078
	+59	+61	.00	+68	+191	+172	.02	+166
groep 4 ('94)	989	993		1032	1008	1006		1048
groep 6 ('96)	1048	1048		1081	1127	1122		1144
	+59	+55	.00	+49	+119	+117	.00	+96
groep 6 ('94)	1037	1039		1072	1118	1110		1135
groep 8 ('96)	1087	1089		1122	1211	1186		1212
	+50	+49	.00	+50	+94	+76	.05	+77

king uit: we kijken hoeveel de gemiddelden per jaargroep zijn toegenomen. Deze gegevens staan tussen haakjes in de tabel.

De taalscores maken duidelijk dat er voor geen van de jaargroepen relevante verschillen zijn tussen de islamitische en vergelijkingsscholen. Wel wordt opnieuw duidelijk dat er erg grote verschillen zijn tussen de prestaties van de leerlingen op de islamitische en vergelijkingsscholen enerzijds en de referentiescholen anderzijds. Als we vervolgens kijken in hoeverre de prestaties voor taal per jaargroep zijn gestegen, dan blijkt ook hier dat de verschillen minimaal zijn. In totaal zijn de leerlingen van zowel de islamitische, als van de vergelijkingsscholen over de jaargroepen heen 147 punten gestegen. Dit komt praktisch overeen met het aantal punten dat de leerlingen in de referentiegroep tussen groep 2 en groep 8 zijn gestegen (nl. 150). Dit laat overigens onverlet dat de leerlingen op islamitische en vergelijkingsscholen op een lager niveau zijn gestart, en ook op een lager niveau zijn geëindigd in het basisonderwijs.

De gegevens ten aanzien van de rekenprestaties laten een iets ander beeld zien dan bij de taalprestaties. Bij rekenen zijn de verschillen binnen de groepen 2, 4 en 6 eveneens minimaal. In groep 8 echter is er een verschil van 20 punten, ofwel een effectgrootte van .40. Ook de Eta²-waarde laat met 4% verklaarde variantie zien dat hier van een relevant verschil gesproken kan worden, en wel in het voordeel van de islamitische scholen. We zien ook dat de gemiddelde 'voortgang' van de islamitische scholen iets gunstiger is (343 vs. 325 punten), maar dit komt dus geheel voor rekening van het ver-

schil in groep 8. Uit de tabel kan ook worden afgelezen dat de referentiescholen het aanzienlijk beter doen, maar dat de totale voortgang met 311 punten onder die van de vergelijkingsscholen en meer nog onder die van de islamitische scholen blijft. Voor rekenen kan bovendien worden opgemerkt dat de verschillen tussen de islamitische en referentiescholen in groep 8 nog maar gering zijn, namelijk 7 punten.

Behalve over cross-sectionele gegevens beschikken we ook over longitudinale gegevens. Immers, een deel van de leerlingen die in het schooljaar 1996/97 aan PRIMA-2 hebben deelgenomen waren eerder, in 1994/95, ook al bij PRIMA-1 betrokken. Van deze leerlingen kunnen we dus nagaan in hoeverre hun prestaties zich in de loop van de 2 jaar ontwikkeld hebben. In Tabel 3 staan de betreffende gegevens; behalve de gemiddelde prestaties in de verschillende jaargroepen bevat de tabel ook de 'ontwikkelingsscores' (in feite de gemiddelde verschillen die op individueel niveau zijn berekend; ook wel aangeduid als 'leerwinst').

In Tabel 3 zijn vooral de ontwikkelingsscores van belang. Voor taal laat de tabel zien dat er praktisch geen verschillen zijn in ontwikkeling tussen islamitische en vergelijkingsscholen. Vergeleken met de referentiescholen zijn die er echter weer wel. Wat betreft de overgang van groep 2 naar 4 zijn de referentiescholen enigszins in het voordeel; voor de overgang van groep 4 naar 6 echter zijn ze in het nadeel. Voor de overgang van groep 6 naar 8 is er geen verschil. Grofweg kunnen we waarschijnlijk voor de hele periode concluderen dat de taalontwikkeling voor alle drie de categorieën

ongeveer parallel loopt. Ten overvloede zij opgemerkt dat een en ander niet wegneemt dat de leerlingen uit de referentiescholen in alle jaargroepen flink hogere prestaties leveren.

Voor *rekenen* ligt de situatie iets anders dan voor taal. Het blijkt dat de islamitische scholen in de overgang van groep 2 naar 4, maar vooral in de overgang van groep 6 naar 8 meer vooruit zijn gegaan dan de vergelijkingsscholen (resp. 2 en 5% verklaarde variantie). Wat groep 2 betreft laat de tabel zien dat er in het schooljaar 1994/95 nog 26 punten verschil waren tussen de beide categorieën van scholen, in het nadeel van de islamitische scholen. Twee jaar later, in 1996/97, is dit verschil teruggebracht tot 9 punten. In de overgang van groep 4 naar groep 6 is de mate van ontwikkeling ongeveer gelijk, terwijl de leerlingen op islamitische scholen in de laatste schooljaren een grotere vooruitgang boeken dan hun jaargroepgenoten op de vergelijkingsscholen. Vergelijken we dit met de situatie op de referentiescholen, dan valt op dat de leerlingen op de islamitische scholen veel meer vooruit zijn gegaan dan de leerlingen op de referentiescholen. Per (twee) jaar bedraagt het verschil respectievelijk 25, 23 en 17 punten.

Uitstroomgegevens groep 8

Bij de overgang van het basis- naar het voortgezet onderwijs zijn enkele vragen aan de groepsleerkracht van groep 8 voorgelegd. Het betreft de volgende onderwerpen. Spijbelen: Heeft deze leerling dit schooljaar wel eens gespijbeld? (1) nooit – (5) zeer vaak. Drop-out: Is deze leerling volgens uw inschatting een potentiële schoolverlater? (1) zeker niet – (5) zeker wel. Eindtoets basisonderwijs: Welke score heeft deze leerling op deze toets behaald? Advies: Welk advies voor voortgezet onderwijs heeft de leerling ontvangen? (1) vso - (11) vwo.

De analyses laten zien dat er enige relevante verschillen optreden tussen islamitische en vergelijkingsscholen. Op islamitische scholen wordt wat minder gespijbeld ($Eta^2 = .03$), en volgens de groepsleerkrachten is de kans dat de betreffende leerlingen drop-out worden ook geringer ($Eta^2 = .05$). Opvallend is dat de islamitische scholen er in beide opzichten ook wat gunstiger afsteken ten opzichte van de referentiescholen. Wat de score op de Eindtoets betreft gaat het om een verschil van 4 punten in het

voordeel van de islamitische scholen; dit is ongeveer de helft van een standaarddeviatie (deze is nl. 9). De totaalscore op de referentiescholen ligt vervolgens weer 5 punten hoger dan die op de islamitische scholen. Qua advies voortgezet onderwijs is er geen relevant verschil tussen islamitische en vergelijkingsscholen; gemiddeld ligt dit advies op mavo-niveau. Binnen de referentiecategorie ligt het iets hoger, namelijk tussen mavo en mavo-havo in.

Schoolwelbevinden

De leerlingen uit groep 6 en 8 hebben een vragenlijstje ingevuld op basis waarvan twee dimensies zijn onderscheiden: schoolwelbevinden en zelfvertrouwen. De scores op de oorspronkelijke items zijn zodanig omgezet dat ze kunnen worden geïnterpreteerd in termen van (1) negatief tot (5) positief. Uit vergelijking van de gemiddelden blijkt dat er geen verschillen zijn tussen islamitische en vergelijkingsscholen. Vermeldenswaardig is dat in groep 8 de leerlingen op de referentiescholen een iets lagere score halen voor zowel welbevinden als voor zelfvertrouwen.

4.2 Nadere analyses

In het voorafgaande hebben we een gedetailleerde vergelijking gemaakt van islamitische met vergelijkingsscholen. Op enkele punten riepen deze analyses nieuwe vragen op. In deze paragraaf proberen we die te beantwoorden.

De subgroep Marokkaanse en Turkse leerlingen

Bij de vergelijking van de islamitische en vergelijkingsscholen hebben we één criterium gehanteerd, namelijk de schoolscore. Deze geeft een indicatie van de sociaal-etnische samenstelling van de leerlingenpopulatie, en daarmee van de onderwijsachterstanden op schoolniveau. Uit de gegevens is echter naar voren gekomen dat er ondanks een nagenoegelijke schoolscore toch sprake is van enig verschil in etnische herkomst tussen de twee schoolcategorieën. Hoewel het merendeel van de leerlingen van Marokkaanse en Turkse origine is, zitten er op vergelijkingsscholen met name minder Marokkaanse en meer Surinaamse kinderen dan op islamitische scholen. Om de vergelijking zuiverder te krijgen hebben we daarom alle hierboven gepresenteerde analyses herhaald, maar dit keer alleen voor de subgroep Marokkaanse en Turkse leerlingen.

Op basis van deze analyses kan wat de prestaties en uitstroomgegevens betreft worden geconcludeerd dat voor zover al sprake was van verschillen, deze wat groter zijn geworden in het voordeel van de islamitische scholen. Echter, het eindresultaat blijft onveranderd, namelijk dat er zich met uitzondering van de rekenprestaties in groep 8 geen relevante verschillen voordoen tussen islamitische en vergelijkings-scholen. Met betrekking tot de andere kenmerken is het zo dat door de beperking van de steekproef tot de Turkse en Marokkaanse leerlingen de verschillen in het merendeel van de gevallen kleiner zijn geworden.

Leertijd

Als we naar de kenmerken kijken waarop islamitische scholen verschillen van vergelijkings-scholen, dan zien we enkele aspecten die verwant zijn aan concepten als 'time on task' en 'opportunity to learn' (cf. Scheerens & Bosker, 1997). Concreet bedoelen we de volgende kenmerken: huiswerk, NT2, OALT, en vorig jaar zelfde school.

Bij huiswerk wordt de leertijd uitgebreid doordat de leerling na schooltijd thuis nog expliciet een tijd met het leerproces bezig is. Bij NT2 krijgt de leerling ook extra instructie, al dan niet in de vorm van een zogenaamd pull-out model, alhoewel doorgaans gewoon binnen de reguliere schooltijden. Bij OALT zijn er in principe verschillende varianten. De leerling kan bijvoorbeeld met een hele klas onderwijs in de 'eigen' taal volgen, maar ook kan er sprake zijn van een pull-out model, waarbij een leerling of een deel van de leerlingen voor een of meerdere uren uit de klas wordt gehaald en apart les krijgt in de eigen taal. In beide gevallen kan het zijn dat de leerling onderdelen van het reguliere onderwijs mist. Daarnaast is er de optie dat het OALT na schooltijd wordt gevolgd, waardoor de leertijd wordt uitgebreid. Bij het laatste kenmerk uit het bovengenoemd rijtje gaat het om de vraag of de leerling in het jaar dat de PRIMA-toetsen zijn afgenomen voor het eerst op de betreffende school zat, of dat hij/zij zich het voorafgaande jaar nog op een andere school bevond. In het laatste geval is – in ieder geval – onduidelijk of eventuele effecten van de school te maken hebben met de huidige of de vroegere school. Daarnaast kan het zijn dat de leerling die het toetsjaar voor het

eerst op deze school zat rechtstreeks uit het buitenland is gekomen, een zij-instromer dus.

Wat de effecten van al deze aspecten zijn, is – zo blijkt uit de literatuur – vooralsnog onduidelijk (cf. Borman & D'Agostino, 1996; Driesen, 1996c, 1998; Rossi & Montgomery, 1994). Daarom zijn we nagegaan of ze tevens samenhangen met de taal- en rekenprestaties van de leerlingen. Via covariantie-analyse is bekeken of er na controle voor het betreffende kenmerk nog prestatieverschillen bestaan tussen islamitische en vergelijkings-scholen.

De analysesresultaten maken duidelijk dat er weliswaar 'effecten' zijn van met name NT2 en OALT op taal en van NT2 op rekenen, maar dat na controle voor deze kenmerken de toename in prestatieverschillen tussen islamitische en vergelijkings-scholen maar in één analyse (OALT en rekenen groep 8) enige relevantie heeft. Anders gezegd: als er al prestatieverschillen bestaan tussen beide schoolcategorieën, dan veranderen die maar in een enkel geval als rekening wordt gehouden met de achtergrondkenmerken.

5 Samenvatting en conclusies

In dit artikel hebben we een beschrijving gegeven van een aantal achtergrondkenmerken en cognitieve en niet-cognitieve kenmerken van leerlingen op islamitische scholen. Deze hebben we vergeleken met die van leerlingen op scholen met een vergelijkbare sociaal-etnische leerlingenpopulatie en leerlingen op scholen uit een landelijk representatieve steekproef. Wat de *gezinskenmerken* betreft blijkt dat islamitische scholen zich vooral van vergelijkings-scholen onderscheiden doordat de ouders van de leerlingen zich enerzijds vaker tot het islamitische geloof rekenen en meer belang hechten aan de plaats van het geloof in de opvoeding. Anderzijds rekenen de betreffende ouders zichzelf en hun kinderen minder vaak tot de Nederlandse cultuurgemeenschap en zijn ze minder op het spreken van de Nederlandse taal gericht. Bovendien gaat het bij deze ouders vaker om volledige gezinnen, hebben ze meer kinderen en zijn de vaders wat hoger opgeleid. Ook kijken de kinderen minder naar de tv. Er zijn geen verschillen geconstateerd wat betreft de tevredenheid van de ouders met de school.

Wat de *onderwijskenmerken* betreft, blijkt dat leerlingen van islamitische scholen vaker NT2- en OALT-lessen volgen, maar ook vaker in het voorafgaande jaar op een andere school hebben gezeten. Ten aanzien van die kenmerken die we als 'cognitieve en niet-cognitieve *effectvariabelen*' zouden kunnen aanduiden, valt het volgende op. De leerlingen van beide schoolcategorieën verschillen niet in non-verbale intelligentie en evenmin in schoolwelbevinden en zelfvertrouwen. Verschillen wat betreft hun cognitieve capaciteiten, onderwijsondersteunend thuisklimaat, werkhouding en sociaal gedrag zijn er volgens de groepsleerkrachten evenmin. Wel liggen op de islamitische scholen de rekenprestaties in groep 8 iets hoger dan op vergelijkingsscholen, en zijn ze qua rekenen ook wat meer vooruitgegaan in de periode van groep 2 naar groep 4 en met name van groep 6 naar groep 8. In de andere leerjaren zijn er geen verschillen in rekenprestaties. Qua taalprestaties zijn er evenmin verschillen. De scores op de Eindtoets basisonderwijs laten zien dat de islamitische scholen ook op die toets wat beter presteren. Al moet daarbij het voorbehoud worden gemaakt dat een aanzienlijk groter deel van de islamitische scholen dan van de vergelijkingsscholen aan de Eindtoets heeft deelgenomen. Op islamitische scholen wordt ook minder gespijeld en is volgens de inschatting van de groepsleerkrachten de kans dat de leerlingen drop-outs worden ook geringer. Bij dit alles moet overigens niet uit het oog worden verloren dat zowel islamitische als vergelijkingsscholen er wat betreft taal- en rekenprestaties, intelligentiescores, Eindtoetsresultaten en adviezen voortgezet onderwijs aanzienlijk ongunstiger voorstaan dan referentiescholen. Deze verschillen variëren met betrekking tot het merendeel van de kenmerken van een halve tot driekwart standaarddeviatie.

Als we de huidige resultaten van PRIMA-2 leggen naast de eerder gerapporteerde bevindingen van PRIMA-1, dan blijkt dat ze groten-deels in één lijn liggen. Met name de taal- en rekenachterstand die de leerlingen op islamitische scholen in de kleutergroepen nog hadden ten opzichte van de leerlingen op vergelijkingsscholen lijkt nu verdwenen. Het betreft hier de nieuw ingestroomde leerlingen. De groep 6 leerlingen van islamitische scholen die in PRIMA-1 zo hoog scoorden op rekenen zitten

nu in groep 8, en scoren daar wederom hoog. Ook wat betreft de overige kenmerken die in beide studies zijn onderzocht lijken er weinig verschillen te bestaan.

In de inleiding van dit artikel hebben we enkele verwachtingen van voorstanders van islamitisch onderwijs geëxpliciteerd. Zij verwachten dat het onderwijs beter is afgestemd op de specifieke leerlingenpopulatie en dat de ouders meer betrokken zijn bij het onderwijs aan hun kinderen. Tevens verwachten ze dat door deze afstemming en betrokkenheid het zelfvertrouwen, welbevinden en prestatieniveau van de kinderen op islamitische scholen hoger ligt, en bovendien dat de ouders tevredener zijn. Nu we de analyseresultaten op een rij hebben staan, moeten we concluderen dat – voorzover we kunnen nagaan op basis van de hier gepresenteerde gegevens – slechts weinig van de verwachtingen waar zijn gemaakt. Weliswaar volgen kinderen van islamitische scholen meer NT2 en OALT, voor meer ouder-school contacten en een grotere ouderbetrokkenheid hebben we geen bevestiging gevonden. Evenmin is gebleken dat het zelfvertrouwen en welbevinden op de islamitische scholen hoger liggen dan op vergelijkbare scholen. Wat de prestaties betreft geldt alleen voor de rekenprestaties dat er in groep 8 en in de periode van groep 6 naar groep 8 relevante verschillen zijn in het voordeel van de islamitische scholen. En ofschoon uit PRIMA-1 bleek dat de ouders van islamitische scholen wat tevredener waren, in PRIMA-2 is deze bevinding niet bestendigd.

Bij het voorafgaande kunnen overigens enkele kanttekeningen worden geplaatst. Uit de schooleffectiviteitsliteratuur is bekend dat rekenen een typisch schoolse vaardigheid is. Taal daarentegen, is iets wat in principe overal geleerd wordt, niet alleen op school, maar ook thuis en op straat. Dat op islamitische scholen juist de rekenprestaties hoger liggen kan daarom wellicht op het conto van de school worden geschreven. De oudergegevens hebben laten zien dat er relatief weinig verschillen zijn. Eén aspect echter is mogelijk direct relevant voor de interpretatie van de taalprestaties, namelijk het feit dat in de thuissituatie van leerlingen op islamitische scholen aanzienlijk minder vaak Nederlands wordt gesproken. Dit zou als een extra nadelige randconditie kunnen worden

opgevat. Dit doortrekkend zou kunnen worden geopperd dat het misschien eerder voor de hand had gelegen dat de islamitische scholen lagere taalprestaties zouden leveren. Dat dit niet het geval is spreekt dan in het voordeel van deze scholen.

De doelen die de islamitische scholen zich stellen zijn tweeledig, namelijk het verbeteren van de onderwijsprestaties van hun leerlingen en hen opvoeden in de geest van de islam. Op basis van onze bevindingen kunnen we – net als we eerder ook al deden – de conclusie trekken, dat leerlingen op islamitische scholen het niet slechter doen dan leerlingen op andere scholen met een vergelijkbare sociaal-etnische achterstand. Dit betekent dat de vrees die door sommigen wordt geuit dat de prestaties in de kernvakken zouden lijden onder het specifieke karakter van de islamitische scholen geen bevestiging vindt. Tegelijkertijd blijkt echter dat islamitische scholen het doorgaans ook niet beter doen dan vergelijkbare scholen. Vergeleken met referentiescholen doen ze het zelfs aanzienlijk slechter. Dit betekent dat de islamitische scholen het doel van verbetering van de prestaties vooralsnog niet gerealiseerd hebben.

Bij deze conclusies kan worden aangetekend dat de islamitische 'zuil' nog maar een relatief korte bestaansgeschiedenis kent. Op het moment van de dataverzameling van PRIMA-2 hadden pas enkele van deze scholen een complete lichte leerlingen bij het voortgezet onderwijs afgeleverd. Het merendeel van de scholen bestond nog pas rond de vijf jaar. Ook kan worden opgemerkt dat de ondersteuning van deze scholen vanuit de koepel nog in haar kinderschoenen staat, zeker vergeleken met die van de gevestigde koepels. Daarbij komt nog dat ook de schoolbesturen van de islamitische scholen relatief onervaren zijn. In het licht van deze contextfactoren is het daarom niet meer dan opportuun af te wachten hoe de scholen bij een volgende PRIMA-meting zullen presteren. Gelet op de bevinding dat de leerlingen op islamitische scholen in groep 8 betere rekenprestaties leveren en ook meer vooruitgang hebben geboekt bij de overgang van groep 6 naar groep 8 dan die op vergelijkingsscholen vormt overigens wel een aanwijzing dat er in principe een groeipotentieel ligt. Analyse van de gegevens die in het schooljaar 1998/99 bij de derde meting van PRIMA wor-

den verzameld kunnen uitsluitsel geven of deze ontwikkeling zich doorzet.

In het onderhavige onderzoek is voornamelijk gekeken naar cognitieve en niet-cognitieve effecten. De kwantitatieve gegevens die wij hebben geanalyseerd betreffen schriftelijk verkregen informatie (toetsen, vragenlijsten). Om het beeld van het islamitisch onderwijs te completeren zou echter ook nog meer-kwalitatief getint onderzoek uitgevoerd dienen te worden, bijvoorbeeld in de vorm van gesprekken met de leerlingen over hun houding, gedrag en opvattingen over hun (toekomstig) functioneren binnen de Nederlandse samenleving. Onderzoek naar de motieven van ouders zou meer licht kunnen werpen op de achtergronden van de schoolkeuzes (vgl. Dijkstra, Dronkers & Hofman, 1997). Daarnaast zou het longitudinale karakter moeten worden uitgebouwd door de leerlingen te blijven volgen in het voortgezet onderwijs en op de arbeidsmarkt om na te gaan hoe ze uiteindelijk functioneren in de Nederlandse maatschappij. Dergelijk aanvullende onderzoek is met name van belang om andere veronderstelde effecten van het islamitische onderwijs, zoals segregatie en vervreemding van de Nederlandse samenleving dan wel integratie en emancipatie, te toetsen.

Noten

- 1 Elders gaan we in op gegevens die te maken hebben met het niveau van de klas en school; vergelijk Driessen en Bezemer (1999).
- 2 In statistische handboeken (bv. Cohen & Cohen, 1983) wordt wat betreft de ondergrens voor een zwakke samenhang vaak een correlatie van .10 genoemd. Dit komt overeen met 1% 'verklaarde' variantie. Omdat deze correlatie ook nog ruis in de vorm van meet- en steekproeffouten bevat, nemen we als criterium voor relevantie een (nominaal-metrische samenhang) van .15, ofwel circa 2% verklaarde variantie.
- 3 Het groepsformulier bevatte ook een vraag met betrekking tot het doubleren in het betreffende leerjaar. Helaas bleek dit gegeven in de door ons gebruikte bestanden niet compleet en bovendien waren er verschillen in respons tussen de islamitische en vergelijkingsscholen. Vandaar dat we af hebben moeten zien van gebruik ervan. Inzicht in het aandeel zittenblijvers op islamiti-

sche en vergelijkingsscholen is van belang om een juiste interpretatie te geven van de toetsprestaties. Immers, als er in de ene categorie meer gedoubleerd wordt dan in de andere kan dat tot een overschatting van het prestatieniveau leiden.

- 4 De cognitieve en niet-cognitieve kenmerken uit het leerlingprofiel zijn gebaseerd op inschatten van groepsleerkrachten met betrekking tot een 30-tal aspecten. Hoewel de alfa-betrouwbaarheden van de geconstrueerde schalen met minimaal .75 als voldoende kunnen worden gekwalificeerd (vgl. Driessen, Van Langen, Portengen & Vierke, 1998), kan er twijfel worden geuit wat betreft hun validiteit. De veronderstelling is dat bij die inschattingen niet alleen de leerling, maar ook de leerkracht een variantiebron is. Sommige leerkrachten oordelen immers systematisch positiever dan anderen. Daarom zouden alleen leerlingen binnen een klas met elkaar vergeleken kunnen worden. Strikt genomen is dit correct. Dat we toch besloten hebben de betreffende variabelen in de analyses te betrekken heeft ten minste twee redenen. Op de eerste plaats zullen er weinig beoordelingen en inschattingen zijn die niet worden beïnvloed door het referentiekader van de beoordelaar en de context van de beoordeling. Veel opinie- en attitude-onderzoek bijvoorbeeld, zou daarmee bij voorbaat zinloos zijn. Op de tweede plaats zou verondersteld kunnen worden dat islamitische scholen er juist vanwege hun door sommigen bekritiseerde positie bij gebaat zijn zich positiever te etaleren dan andere scholen. Als dit zo zou zijn, dan zouden er dus systematische verschillen op moeten treden tussen de scholen categorieën. Dit is echter alleen het geval als de prognose voortgezet onderwijs en het onderwijsondersteunend thuisklimaat van de referentiecategorie worden vergeleken met die van de twee andere categorieën. En dat ligt precies in de lijn der verwachtingen en geeft daarmee ondersteuning van de validiteit van de gegevens.

Literatuur

Aarsen, L., & Jansma, P. (1992). *Een islamitische basisschool. Onderzoek naar verwachtingen van moslimouders en leerkrachten aangaande de islamitische basisschool*. Nijmegen: KUN.

Alkan, M. (1996). Islamitische initiatieven in het onderwijs. In R. Lavrijsen (Ed.), *Islam in een ontzuilde samenleving. Discussies over vrouwenemancipatie, kunst en onderwijs* (pp. 113-146), Amsterdam: Koninklijk Instituut voor de Tropen/Soeterij.

Borman, G., & D'Agostino, J. (1996). Title I and student achievement: A meta-analysis of federal evaluation results. *Educational Evaluation and Policy Analysis*, 18, 309-326.

Canatan, K. (1993). De islamitische scholen zijn een alternatief voor de mislukte Nederlandse scholen. *Vernieuwing*, 52 (6), 12.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum Associates.

Cumper, P. (1990). Muslim schools: the implications of the Education Reform Act 1988. *New Community*, 16, 379-389.

Dijkstra, A.B., Dronkers, J., & Hofman, R. (Eds.) (1997). *Verzuiling in het onderwijs. Actuele verklaringen en analyse*. Groningen: Wolters-Noordhoff.

Driessen, G. (1996a). De gezinssituatie van leerlingen op islamitische basisscholen. Een verkenning van hun achtergronden. *Nederlands Tijdschrift voor Opvoeding, Onderwijs en Vorming*, 12, 247-265.

Driessen, G. (1996b). Prestaties, gedrag en houding van leerlingen op islamitische basisscholen. Een vergelijking van islamitische scholen en scholen met een overeenkomstige sociaal-etnische leerlingenpopulatie. *Migrantenstudies*, 12, 136-151.

Driessen, G. (1996c). Minority Language and Culture Teaching in the Netherlands: Policies, arguments, evaluation and prospects. *Compare*, 26, 315-332.

Driessen, G. (1997). Islamic primary schools in the Netherlands: The pupils' achievement levels, behaviour and attitudes and their parents' cultural backgrounds. *The Netherlands' Journal of Social Sciences*, 33, 42-66.

Driessen, G. (1998). *The limits of educational policy and practice? The case of ethnic minority pupils in the Netherlands*. Nijmegen: ITS.

Driessen, G., & Bezemer, J. (1999). *Islamitisch basisonderwijs. Schipperen tussen identiteit en kwaliteit?* Nijmegen: ITS.

Driessen, G., & Haanstra, F. (1996). *Achtergrondkenmerken van leerlingen in het primair onderwijs. Beschrijvende rapportage op basis van het PRIMA-cohortonderzoek 1994/95*. Ubbergen: Tandem Felix.

- Driessen, G., Langen, A. van, Portengen, R. & Vierke, H. (1998). *Basisonderwijs: Veldwerkverslag, leerlinggegevens en ouder vragenlijsten. Basisrapportage PRIMA-cohortonderzoek. Tweede meting 1996-1997*. Nijmegen: ITS.
- Driessen, G., & Withagen, V. (1998). Taalvariatie en onderwijsprestaties van autochtone basisschool-leerlingen. *Taal en Tongval*, 50, 2-24.
- Dwyer, C. (1993). Constructions of Muslim identity and the contesting of power: the debate over muslim schools in the United Kingdom. In P. Jackson & J. Penrose (Eds.), *Constructions of race, place and nation* (pp. 143-159). London: UCL.
- Dwyer, C., & Meyer, A. (1995). The institutionalisation of Islam in The Netherlands and in the UK: the case of Islamic schools. *New Community*, 21, 37-54.
- Dwyer, C., & Meyer, A. (1996). The establishment of Islamic schools. A controversial phenomenon in three European countries. In W. Shadid & P. van Koningsveld (Eds.), *Muslims in the margin. Political responses to the presence of Islam in Western Europe* (pp. 218-242). Kampen: Kok Pharos.
- Esch, W. van, & Roovers, M. (1987). *Islamitisch godsdienstonderwijs in Nederland, België, Engeland en West-Duitsland*. Nijmegen: ITS.
- Kabdan, R. (1992). Op weg naar maatschappelijke spanningen op religieuze gronden. Islamitische scholen in Nederland zijn onwenselijk. *Vernieuwing*, 51(9), 4-8.
- Kabdan, R. (1993). Islamitische scholen zijn verre van alternatief. *Vernieuwing*, 52(6), 13.
- Karagül, A. (1994). *Islamitisch godsdienstonderwijs op de basisschool in Nederland. Theorie en praktijk in vergelijking met enkele Europese en Moslimse landen*. Amsterdam: UvA.
- Lammers, M. (1993). *Ideologie en praktijk. Een onderzoek naar het bijzondere karakter van islamitische basisscholen in Nederland*. Amsterdam: UvA.
- Landman, N. (1992). *Van mat tot minaret. De institutionalisering van de islam in Nederland*. Amsterdam: VU Uitgeverij.
- Langen, A. van, Vierke, H., & Robijns, M. (1996). *Veldwerkverslag basisonderwijs en speciaal onderwijs. Technische rapportage PRIMA-cohortonderzoek 1994/95*. Nijmegen/Amsterdam: ITS/SCO.
- Lavrijsen, R. (Ed.) (1996). *Islam in een ontzuilde samenleving. Discussies over vrouwenemancipatie, kunst en onderwijs*. Amsterdam: Koninklijk Instituut voor de Tropen/Soeterijn.
- Leman, J., & Renaerts, M. (1996). Dialogues at different institutional levels among authorities and Muslims in Belgium. In W. Shadid & P. van Koningsveld (Eds.), *Muslims in the margin. Political responses to the presence of Islam in Western Europe* (pp. 164-181). Kampen: Kok Pharos.
- Meyer, A. (1993). *Islam en onderwijs in Utrecht. De oprichting van islamitische scholen en islamitisch godsdienstonderwijs op openbare scholen*. Nijmegen: KUN.
- Parker-Jenkins, M. (1991). Muslim matters: the educational needs of the muslim child. *New Community*, 17, 569-582.
- Peach, C., & Glebe, G. (1995). Muslim minorities in western Europe. *Ethnic and Racial Studies*, 18, 26-45.
- Pedersen, L. (1996). Islam in the discourse of public authorities and institutions in Denmark. In W. Shadid & P. van Koningsveld (Eds.), *Muslims in the margin. Political responses to the presence of Islam in Western Europe* (pp. 202-217). Kampen: Kok Pharos.
- Rath, J., Penninx, R., Groenendijk, K., & Meyer, A. (1996). *Nederland en zijn islam. Een ontzuilende samenleving reageert op het ontstaan van een geloofsgemeenschap*. Amsterdam: Het Spinhuis.
- Rath, J., Penninx, R., Groenendijk, K., & Meyer, A. (1997). De Nederlandse samenleving en haar islam. Reacties op de opbouw van islamitische instituties. *Migrantenstudies*, 13, 69-79.
- Rossi, R., & Montgomery, A. (1994). *Educational reforms and students at risk. A review of the current state of the art*. Washington DC: US Department of Education.
- Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Shadid, W., & Koningsveld, P. van (1992a). Islamitische scholen. De verschillende scholen en hun achtergronden. *Samenwijs*, 12, 227-233.
- Shadid, W., & Koningsveld, P. van (1992b). Islamic primary schools. In W. Shadid & P. van Koningsveld (Eds.), *Islam in Dutch society: Current developments and future prospects* (pp. 107-122). Kampen: Kok Pharos.
- Sunier, T., & Meyer, A. (1997). Religie. In H. Vermeulen (Ed.), *Immigrantenbeleid voor de multiculturele samenleving* (pp. 95-125). Amsterdam: Het Spinhuis.
- Tesser, P., Praag, C. van, Dugteren, F. van, Herweijer, L., & Wouden, H. van der (1995). *Rapportage minderheden 1995. Concentratie en segregatie*. Rijswijk: SCP.

Teunissen, J. (1990). Basisscholen op islamitische en hindoeïstische grondslag. *Migrantenstudies*, 6, 45-57.

Vermeulen, H. (Ed.) (1996). *Immigrantenbeleid voor de multiculturele samenleving*. Amsterdam: Het Spinhuis.

Vierke, H. (1995). *De PRIMA-toetsen gecalculeerd. De ontwikkeling van vaardigheidsscores over de leerjaren heen op basis van de jaargroepoetsen in het cohort primair onderwijs (PRIMA)*. Nijmegen: ITS.

Wagtendonk, K. (1991). Islamic schools and Islamic religious education. A comparison between Holland and other West European countries. In W. Shadid & P. van Koningsveld (Eds.), *The integration of Islam and Hinduism in Western Europe* (pp. 154-173). Kampen: Kok Pharos.

Auteurs

G. Driessen is als onderzoeker op het terrein van het onderwijs werkzaam op het Instituut voor Toegepaste Sociale Wetenschappen (ITS) van de Katholieke Universiteit Nijmegen.

J. Bezemer liep ten tijde van het hier gerapporteerde onderzoek stage op het ITS vanuit het Werkverband Taal en Minderheden van de Katholieke Universiteit Brabant.

Correspondentie-adres: ITS, Postbus 9048, 6800 KJ Nijmegen.

Abstract

Backgrounds, features and yields of Islamic primary schools

G.W.J.M. Driessen & J.J. Bezemer. *Pedagogische Studiën*, 1999, 76, 332-349.

Since the 1988/89 school year, 29 Islamic schools have been founded in the Netherlands. These schools are attended by about 7,000 pupils, largely of Moroccan and Turkish descent. Their main aims are to improve the pupils' achievements, and to strengthen the pupils' Islamic identity. This article first of all explores the backgrounds of Islamic education. Then, the achievement levels, behaviour, attitudes and family backgrounds of pupils at 16 Islamic schools are compared with pupils at 16 schools with a similar socio-ethnic population and pupils at 432 schools from a nationally representative sample (in total more than 40,000 pupils). With regard to their achievements, behaviour and attitudes the analyses show that there are only very small if any differences between the Islamic and similar schools. However, they perform far below the pupils at the schools from the representative sample. With regard to the family background characteristics, there are some minor differences between the pupils at Islamic schools and pupils at similar schools. The parents are more oriented towards the Muslim community and they attach more importance to the role of religion in the upbringing of their children. At the same time they are less oriented towards the Dutch society and they speak less Dutch. For the time being, it is possible to conclude that pupils at Islamic schools do not do any worse, nor any better than pupils at schools with a similar socio-ethnic disadvantage. Time will tell whether they will succeed in raising their educational performance up to the level of non-disadvantaged pupils, and how the pupils will eventually function in Dutch society.

De invloed van achtergrondtelevisie op de uitvoering van gemakkelijke en moeilijke huiswerktaken

M. M. Pool, T. H. A. van der Voort, J. W. J. Beentjes en C. M. Koolstra*

Samenvatting

Langs experimentele weg werd nagegaan wat de invloed is van het gebruik van twee soorten achtergrondtelevisieprogramma's op de uitvoering van zowel gemakkelijke als moeilijke huiswerktaken. Leerlingen uit de tweede klas van de MAVO en HAVO ($N = 90$) werden gematched op leesvaardigheid. Vervolgens werden zij at random toegewezen aan één van de drie experimentele condities waarin huiswerk werd gemaakt met op de achtergrond ofwel: (a) Engelstalige videoclips, of (b) een Nederlandstalige soap-serie, of (c) stilte (controlegroep). Uit het experiment kwam naar voren dat de uitvoering van het huiswerk alleen werd belemmerd als er een soap-programma aanstond. Dit distractie-effect kwam alleen tot uitdrukking in een verminderde prestatie bij moeilijke huiswerkopgaven. Achtergrondtelevisie leidde niet tot een verlenging van de tijd benodigd voor het maken van de huiswerktaken.

1 Inleiding.

Onderwijskundig mediaonderzoek is over het algemeen gericht op de positieve rol die media in het leerproces kunnen spelen. Media worden gezien als een hulpmiddel voor de docent, en het onderzoek is erop gericht vast te stellen hoe media de leerprestaties binnen de schoolcontext kunnen verhogen. In buitenschoolse situaties gebruiken scholieren de media echter niet altijd om hun leerprestaties te verbeteren, en worden media ook als achtergrond bij het maken van huiswerk gebruikt. Vroeger ge-

bruikten scholieren uitsluitend geluidsmedia als achtergrond, met name de radio, grammofoonplaten of cassettebandjes. Tegenwoordig maken scholieren hun huiswerk ook bij de televisie. Volgens een Amerikaanse studie die 15 jaar geleden werd uitgevoerd (Patton, Stinard, & Routh, 1983) maakt ongeveer 30% van de scholieren van 10 tot 15 jaar het huiswerk wel eens bij de televisie. Volgens een meer recente Nederlandse studie (Beentjes, Koolstra, & Van der Voort, 1996) combineert ruim 50% van de leerlingen uit klas 2 en 4 van het voortgezet onderwijs het maken van huiswerk wel eens met televisie kijken. Het hogere percentage gevonden in de meest recente studie is wellicht toe te schrijven aan het feit dat onder scholieren het bezit van televisieapparatuur op de eigen kamer steeds meer gemeengoed begint te worden (Van der Voort et al., 1998).

Anders dan media gebruikt in de schoolcontext, zou het gebruik van media als secundaire activiteit bij het huiswerk de leerprestaties wel eens eerder kunnen hinderen dan bevorderen. Wanneer het maken van huiswerk wordt gecombineerd met televisie kijken, wordt het cognitieve verwerkingsapparaat immers opgezadeld met een extra taak die de aandacht van het huiswerk kan afleiden. Diverse experimentele studies suggereren dat het gebruik van televisie op de achtergrond inderdaad de prestaties op cognitieve taken kan verlagen (Armstrong, 1993, 1997; Armstrong, Boiarsky, & Mares, 1991; Armstrong & Greenberg, 1990; Armstrong & Sopory, 1997; Furnham, Gunter, & Peterson, 1994). Het is echter de vraag of de in deze studies gevonden distractie-effecten ook gelden voor de gebruikelijke huiswerksituatie thuis, omdat de taken die in deze experimenten gebruikt werden in drie opzichten verschilden van gewoon huiswerk: (a) er werd gebruik gemaakt van cognitieve tests (bv. een test voor ruimtelijk inzicht of creativiteit) die qua inhoud vaak ver afstonden van gebruikelijke huiswerktaken; (b) er werden doorgaans kortduren-

* Dit onderzoek vond plaats met financiële steun van SVO. De auteurs danken Lissa Moerman, Jacqueline Tournier-Hoogewij, Esther van Renesse, en Miranda Vroom voor hun bijdrage aan de verzameling van de gegevens.

de taken gebruikt die slechts enkele minuten vergden, terwijl huiswerktaken doorgaans meer tijd en een langduriger concentratie vereisen; en (c) er werd meestal een tijdslimiet gesteld aan de uitvoering van de taak, terwijl het in de huiswerksituatie zelden of nooit voorkomt dat het aantal minuten dat aan een taak mag worden besteed begrensd is. Een belangrijk doel van het hier te presenteren onderzoek was dan ook na te gaan of de distractie-effecten van achtergrondtelevisie vastgesteld bij cognitieve tests zich ook voordoen bij taken die een grotere gelijkenis vertonen met reguliere huiswerktaken.

1.1 Theoretische verklaringen van distractie-effecten

Voor het optreden van distractie-effecten van achtergrondtelevisie op de huiswerkprestaties zijn vier theoretische verklaringen in omloop (Armstrong & Greenberg, 1990) die elkaar geenszins uitsluiten. De meest gangbare verklaring is gebaseerd op een '*beperkte-aandachtscapaciteits-theorie*' (bv. Kahneman, 1973) waarin gesteld wordt dat de mens bij uitvoering van huiswerk of andere cognitieve taken beschikt over een beperkte capaciteit voor het verwerken van informatie. Bij het gelijktijdig verwerken van de informatie vervat in de huiswerктаak (de primaire taak) en het televisieprogramma (de secundaire taak) is er sprake van twee taken, die elk voor zich een beroep doen op de beperkte aandachtscapaciteit. Met name bij moeilijke huiswerktaken die een sterk beroep doen op de aandachtscapaciteit dreigt het gevaar dat de totale aandachtscapaciteit door de aanwezigheid van achtergrondtelevisie wordt overschreden, met als gevolg dat de aandachtscapaciteit die beschikbaar is voor het maken van huiswerk wordt aangetast en de prestaties worden verminderd. Dit interferentie-effect wordt door Kahneman *capaciteitsinterferentie* genoemd.

Behalve via capaciteitsinterferentie kunnen huiswerkprestaties ook gehinderd worden door *structurele interferentie*, hetgeen zich voor kan doen wanneer het huiswerk en het achtergrondtelevisieprogramma een beroep doen op hetzelfde subsysteem van informatieverwerking (Kahneman, 1973). In zo'n geval kunnen de twee taken zelfs met elkaar interfereren wanneer de totale aandachtscapaciteit niet wordt

overschreden. Structurele interferentie kan zich bijvoorbeeld voordoen wanneer het maken van verbale huiswerkopgaven wordt gecombineerd met het kijken naar een praatprogramma, omdat bij beide activiteiten sprake is van de verwerking van linguïstische informatie.

Ten derde kan interferentie met huiswerk optreden doordat saillante, aandachttrekkende elementen uit het televisieprogramma *oriënterende reacties* oproepen die de aandacht op het beeld richten. Hierdoor wordt de aandacht van het huiswerk afgeleid, waarmee het huiswerk tijdelijk van een primaire in een secundaire taak verandert. Volgens Armstrong en Greenberg (1990) is de interferentie die voortvloeit uit het oproepen van oriënterende reacties vooral schadelijk bij huiswerktaken die continu aandacht vergen, zoals bijvoorbeeld het zich eigen maken van basisideeën uit een moeilijke studietekst.

Tenslotte kan achtergrondtelevisie de huiswerkprestaties beïnvloeden door invloed uit te oefenen op het niveau van *fysiologische activatie* van de leerling. Zoals bekend, zijn leerprestaties doorgaans kromlijng gerelateerd aan het activatieniveau (in de vorm van een omgekeerde U-kromme), waarbij het optimale activatieniveau hoger ligt voor gemakkelijke dan voor moeilijke taken (Yerkes & Dodson, 1908). Daarom kan de activerende werking die van achtergrondtelevisie uitgaat de uitvoering van simpele taken bevorderen, en de uitvoering van moeilijke huiswerktaken juist belemmeren (Anderson & Revelle, 1982).

1.2 Eerder onderzoek naar distractie-effecten

Experimentele studies naar de distractie-effecten van achtergrondtelevisie op de prestaties op cognitieve tests lieten zien dat achtergrondtelevisie vooral interfereert bij moeilijke cognitieve taken die een langdurige aandacht vereisen. Meer in het bijzonder zijn er aanwijzingen dat drie typen moeilijke taken door achtergrondtelevisie kunnen worden geschaad: (a) begrijpend-lezen-taken (Armstrong, 1997; Armstrong, et al., 1991; Armstrong & Greenberg, 1990; Furnham, et al., 1994); (b) de *Toren-van-Hanoi-puzzel*, een non-verbale test van het vermogen om ruimtelijke problemen op te lossen die veronderstelt dat men in staat is om middeldoel-analyses uit te voeren en planmatig te handelen (Armstrong & Greenberg, 1990); en

(c) de *Uses test*, een creativiteitsmaat die verlangt dat de proefpersoon zoveel mogelijk toepassingen bedenkt voor een dagelijks gebruiksvoorwerp (Armstrong & Greenberg, 1990). In de studie van Armstrong en Greenberg (1990) werd geen nadelige invloed vastgesteld bij simpele taken die bestonden uit verzamelingen van eenvoudige, op zichzelf staande problemen die geen langdurige aandacht behoeften. Maar andere studies wijzen uit dat ook bij simpele taken interferentie kan optreden, zij het in mindere mate dan bij moeilijke taken (bv. Patton, Routh, & Offenbach, 1981).

De bevinding dat vooral bij moeilijke taken distractie-effecten optreden, is door de betrokken onderzoekers geïnterpreteerd als evidentie die pleit voor de capaciteitsinterferentie-verklaring, die poneert dat interferentie vooral optreedt bij moeilijke taken die in combinatie met achtergrondtelevisie de aandachtscapaciteit overschrijden. De bevinding dat moeilijke taken het meest te lijden hebben van achtergrondtelevisie lijkt echter evenzeer te pleiten voor de oriënterende reacties-verklaring en de fysiologische activatie-verklaring, die eveneens voorspellen dat interferentie zich primair voordoet bij moeilijke taken. Daarnaast zijn er enkele aanwijzingen dat distractie-effecten zijn toe te schrijven aan structurele interferentie (Armstrong, 1993; Armstrong & Sopory, 1997).

Het distractie-effect van achtergrondtelevisie kan behalve in een afname van de prestaties tot uitdrukking komen in een verlenging van de tijd die de leerling nodig heeft om de taak af te ronden, een criterium dat alleen bruikbaar is wanneer geen tijdslimiet aan de afronding van een taak wordt gesteld. In de enkele gevallen waarin de tijd die beschikbaar was voor de taak werd vrijgelaten (Armstrong & Greenberg, 1990; Patton, et al., 1981), leidde achtergrondtelevisie vooral bij moeilijke taken tot een verlenging van de tijd benodigd om de taak af te ronden.

Distractie-effecten van achtergrondtelevisie zijn aangetroffen bij uiteenlopende soorten televisieprogramma's, waaronder actieprogramma's (Armstrong & Greenberg, 1990), comedy-programma's (Armstrong & Sopory, 1997), soap-series (Armstrong, 1997; Armstrong et al., 1991; Armstrong & Greenberg, 1990; Furnham et al., 1994) en reclamespotjes

(Armstrong, 1993; Armstrong et al., 1991). Slechts twee studies waren erop gericht verschillen op te sporen in de mate waarin verschillende soorten televisieprogramma's afleidend werken. Er werden geen verschillen in interferentie-effecten gevonden tussen programma's met een sterk verbale inhoud, zoals de soap-serie *Dynasty*, en programma's met een hoog actiegehalte, zoals *The A-Team* (Armstrong & Greenberg, 1990), en evenmin tussen soap-series en reclameuitzendingen (Armstrong et al., 1991).

In slechts één studie is expliciet geprobeerd de invloed van achtergrondtelevisie op de uitvoering van *huiswerktaken* vast te stellen (Cool, Yarbrough, Patton, Runde, & Keith, 1994). In twee kleinschalige experimenten, beide uitgevoerd bij kinderen uit groep 8, werd nagegaan of achtergrondtelevisie en -radio interfereerden met de uitvoering van wiskundeopgaven en leestaken. In beide experimenten werd het best gepresteerd zonder achtergrondmedium en het slechtst in de televisieconditie. De gevonden verschillen in prestaties tussen de studiecondities waren echter niet statistisch significant, hetgeen volgens de auteurs een uitvloeisel kan zijn geweest van de gebruikte kleine steekproeven en de daarmee gepaard gaande beperking van de statistische *power* voor het blootleggen van significante verschillen. De aanwezigheid van televisie op de achtergrond leidde echter wel tot een grote (en statistisch significante) vermindering van het aantal wiskundeopgaven dat de deelnemers maakten.

1.3 De huidige studie

Zoals eerder besproken, was het hoofddoel van de hier te presenteren studie om – net als in het onderzoek van Cool et al. (1994) – na te gaan in hoeverre de distractie-effecten van achtergrondtelevisie op cognitieve tests aangetroffen in eerder onderzoek generaliseerbaar zijn naar de natuurlijke huiswerk situatie. Het tweede doel was na te gaan in hoeverre de distractie-effecten van achtergrondtelevisie afhankelijk zijn van de moeilijkheidsgraad van de huiswerkopgaven. Het derde doel was na te gaan of de distractie-effecten van achtergrondtelevisie afhankelijk zijn van het type televisieprogramma waarmee het maken van huiswerk wordt gecombineerd.

Gepoogd werd de studie zodanig op te zet-

ten dat de onderzoeksresultaten zo goed mogelijk generaliseerbaar waren naar de natuurlijke condities waaronder leerlingen thuis huiswerk maken. De volgende keuzen waren erop gericht de ecologische validiteit van de studie te bevorderen:

(a) De studie werd uitgevoerd bij die leerlingen uit het algemeen vormend voortgezet onderwijs die blijkens eerder onderzoek (Beentjes et al., 1996) het meest frequent met de televisie aan huiswerk maken, namelijk leerlingen van de MAVO en HAVO.

(b) Gebruik werd gemaakt van televisieprogramma's die blijkens eerder onderzoek (Beentjes, et al., 1996) het vaakst door middelbare scholieren met het maken van huiswerk worden gecombineerd, namelijk Engelstalige videoclips en een Nederlandstalige soap-serie (*Goede Tijden, Slechte Tijden*).

(c) De geselecteerde huiswerktaken bestonden alle uit 'maakwerk', dat wil zeggen taken waarbij leerlingen antwoorden op vragen of opgaven moesten opschrijven of invullen. Dit type taken wordt volgens een studie van Beentjes et al. (1996) door middelbare scholieren aanzienlijk vaker met televisie kijken gecombineerd dan 'leertaken', waarbij leerlingen taken uit het hoofd moeten leren of zich de hoofdlijnen uit een studietekst eigen moeten maken.

(d) In plaats van korte cognitieve tests werd gebruik gemaakt van wat langere huiswerktaken op het terrein van de schoolvakken wiskunde en Nederlands.

(e) Net als in de thuissituatie, mochten de deelnemers zelf bepalen hoeveel tijd zij aan een taak besteedden.

1.4 Hypothesen

In het onderzoek werden drie hypothesen getoetst. De eerste hypothese was dat elk van de twee typen achtergrondtelevisieprogramma's die in het onderzoek waren betrokken – soap-serie en videoclips – zouden interfereren met het huiswerk. Omdat de deelnemers zelf mochten bepalen hoeveel tijd zij aan de huiswerktaken besteedden, verwachtten we dat het distractie-effect niet alleen tot uitdrukking zou komen in verminderde prestaties maar ook in een verlenging van de tijd gebruikt om de huiswerktaken af te ronden. Dat soap-programma's distractie-effecten kunnen oproepen werd reeds in diverse studies vastgesteld (Arm-

strong, 1997; Armstrong et al., 1991; Armstrong & Greenberg, 1990; Furnham et al., 1994). Bij ons weten, is echter nog niet eerder nagegaan of ook videoclips interfereren met huiswerkprestaties of huiswerktijd. Wel is onderzocht of alleen auditief gepresenteerde muziek (zonder televisiebeelden) de prestaties op cognitieve taken in de weg zitten. Hoewel simpele instrumentale muziek de prestaties op cognitieve taken zelfs bleek te kunnen verhogen (Kiger, 1989; Matthews & Brunson, 1979; Strube, Turner, Patrick, & Perrillo, 1983), bleek moeilijke instrumentale muziek (Fogelson, 1973) en vooral vocale muziek de prestaties juist te hinderen (Daoussis & McKelvie, 1986; Etaugh & Ptasnik, 1982; Furnham & Bradley, 1997; Martin, Wogalter, & Forlano, 1988; Salamé & Baddeley, 1989). Omdat videoclips naast vocale muziek ook nog beelden bevatten die af kunnen leiden, verwachtten we dat videoclips minstens even sterk als vocale muziek de prestaties zouden belemmeren.

Ten tweede verwachtten we te vinden dat het distractie-effect bij beide programmatypen – soap-programma en videoclips – het sterkst zou uitvallen bij de moeilijke huiswerktaken. Deze hypothese is in overeenstemming met de bevindingen uit eerder onderzoek naar de distractie-effecten van zowel achtergrondtelevisie (Armstrong & Greenberg, 1990; Armstrong et al., 1991; Patton, Routh, & Offenbach, 1981) als andersoortige taak-irrelevante stimuli (Graydon & Eysenck, 1989; Sanders & Baron, 1975). Zoals eerder besproken, is de aanname dat distractie-effecten sterker zullen uitvallen bij moeilijke dan bij gemakkelijke taken in overeenstemming met meerdere verklaringsmodellen, namelijk de capaciteitsinterferentie-hypothese, de oriënterende reacties-verklaring en de fysiologische activatie-verklaring.

Tenslotte verwachtten we te vinden dat de huiswerkprestaties meer te lijden zouden hebben van het gebruik van een soap-programma dan van videoclips op de achtergrond, omdat er tenminste vier redenen zijn om aan te nemen dat soap-series een sterker beslag leggen op de aandachtscapaciteit dan videoclips. Ten eerste leidt het gesproken woord, dat in soap-series centraal staat, meer af dan de muziek (Salamé & Baddeley, 1989) die in videoclips een centrale plaats inneemt. Ten tweede bieden soap-programma's samenhangende informatie die

kennis vergt van het voorafgaande, terwijl muziekprogramma's ongerelateerde videoclips bevatten die ook zonder kennis van het voorafgaande te volgen zijn. Ten derde bevatten de afleveringen van soap-series iedere dag weer nieuwe informatie, terwijl de meeste videoclips die worden uitgezonden reeds veel vaker zijn uitgezonden en dus overbekend zijn. Ten vierde, bevat de audiotrack van soap-programma's veel cues zoals geluidseffecten, kreten en gelach die een leerling kunnen aansporen de aandacht op het beeld te richten (Anderson & Lorch, 1983). Videoclips bevatten daarentegen zelden auditieve cues die oriënterende reacties oproepen. Tenslotte was er in de hier te presenteren studie nog een extra reden om aan te nemen dat het gebruikte soap-programma (*Goede Tijden, Slechte Tijden*) sterker zou afleiden dan videoclips, aangezien in deze soap-serie de moedertaal van de leerlingen wordt gesproken, terwijl de videoclips Engels-talig zijn. Er zijn aanwijzingen dat gesprekken op de achtergrond sterker afleiden wanneer de moedertaal in plaats van een vreemde taal te horen is (Martin, et al., 1988).

2 Methode

2.1 Deelnemers

De studie werd uitgevoerd bij een steekproef van 45 meisjes en 45 jongens uit klas 2 ($M = 13$ jaar en 10 maanden) van drie scholen voor voortgezet onderwijs in de regio Leiden. De deelnemers zaten op de MAVO of de HAVO. Voorafgaand aan het experiment werd bij de leerlingen uit 8 tweede klassen ($N = 187$) klassikaal een gestandaardiseerde toets voor begrip-pend lezen (Van Kan & Aarnoutse, 1993) afgenomen. Bij die gelegenheid werd tevens gevraagd of zij, tegen een vergoeding van 20 gulden, op een later moment in hun vrije tijd op de universiteit aan het experiment wilden deelnemen. De 119 leerlingen (64%) die bereid waren deel te nemen, werden ingedeeld in subgroepen met een identieke score op de begrip-pend leestoets. Binnen elke subgroep werden één of meer groepen van 3 leerlingen geselecteerd, totdat het totaal aantal van 90 deelnemers was bereikt.

2.2 Proefopzet

De leerlingen opgenomen in elk van de 30 subgroepen van 3 leerlingen met identieke leesvaardigheidsscores werden *at random* toegewezen aan één van de drie experimentele condities waarin huiswerk werd gemaakt met op de achtergrond: (a) Engelstalige videoclips, (b) een Nederlandstalige soap-serie, of (c) geen televisie. Alle deelnemers maakten zowel gemakkelijke als moeilijke huiswerktaken. De volgorde van aanbieding werd systematisch gevarieerd. Er was dus sprake van een 3 (achtergrondconditie: soap vs. videoclips vs. stilte) $\times$ 2 (gemakkelijke vs. moeilijke taken) factorieel design, met achtergrondconditie als een tussen-proefpersonen-factor en moeilijkheidsgraad als binnen-proefpersonen-factor. Echter, omdat de leerlingen geplaatst waren in blokjes van drie deelnemers met een vergelijkbaar leesvaardigheidsniveau, werd de factor achtergrondconditie in de analyses gebruikt als een 'binnen-proefpersonen'-factor of liever een 'binnen-blokken'-factor, een aanpak die de foutenvariantie kan verkleinen (Kirk, 1968).

2.3 Achtergrondtelevisieprogramma's

In de conditie waarin het maken van huiswerk gecombineerd werd met het kijken naar videoclips, werd gebruik gemaakt van een videoband met 32 van de televisie opgenomen Engelstalige videoclips met een totale duur van 110 minuten. De videoclips waren achter elkaar gemonteerd onder weglating van interviews, aankondigingen door de presentator en reclamespotjes. Ter bevordering van de ecologische validiteit werd gebruik gemaakt van recente videoclips die twee maanden voorafgaand aan het experiment door de televisie waren uitgezonden en dus nog betrekkelijk vers in het geheugen lagen. Ook in de thuissituatie ziet men doorgaans videoclips die meerdere keren op de televisie zijn geweest en daarom vers in het geheugen liggen.

In de conditie waarin het maken van huiswerk werd gecombineerd met het kijken naar een soap-programma, werd gebruik gemaakt van een videoband bestaand uit vijf afleveringen van de bij de Nederlandse jeugd zeer populaire soap-serie *Goede Tijden, Slechte Tijden*. De totale duur van de band was 107 minuten. De afleveringen waren achter elkaar gemonteerd onder weglating van herhalingen

van fragmenten uit eerdere afleveringen, reclame, en (behalve bij de eerste aflevering) de standaardbeginbeelden. Anders dan bij video-clips, is de inhoud van afleveringen van soap-series die in de thuissituatie onder het huiswerk worden bekeken voor de leerlingen nieuw, omdat de afleveringen voor het eerst worden uitgezonden. Uit een oogpunt van ecologische validiteit was het dus het best geweest wanneer in het experiment gebruik gemaakt zou zijn van nog niet eerder uitgezonden afleveringen van *Goede Tijden, Slechte Tijden*. Omdat geen nieuwe afleveringen beschikbaar waren, werd gebruik gemaakt van juist relatief oude afleveringen, die vier jaar voorafgaand aan het experiment waren uitgezonden, dit om de kans te verkleinen dat de deelnemers de afleveringen reeds kenden.

2.4 Huiswerktaken

De deelnemers maakten allen drie *gemakkelijke taken*: (a) een hoofdreken-taak bestaand uit een serie van 75 opgaven (De Vos, 1992) waarin kinderen moesten demonstreren dat zij in staat waren om, zonder het gebruik van kladpapier, op getallen van één of twee cijfers de rekenkundige bewerkingen optellen, aftrekken, vermenigvuldigen en delen uit te voeren; (b) een synoniementaak bestaand uit 30 vragen (Aarnoutse, 1992), waarin kinderen gevraagd werd voor een gegeven woord het synoniem aan te wijzen uit een viertal alternatieven; en (c) een tegenstellingentaak bestaand uit 39 items (Aarnoutse, 1992), waarin kinderen gevraagd werd voor een gegeven woord het woord met een tegenovergestelde betekenis aan te wijzen uit een viertal alternatieven.

Naast de drie gemakkelijke taken maakte elke deelnemer twee *moeilijke taken*: (a) een reken-taak bestaand uit 15 redactiesommen (De Vos, 1995), en (b) een begrijpend-lezen-taak bestaand uit drie teksten met in totaal 17 meerkeuzevragen. De opgaven uit de leestaak waren ontleend aan verschillende leesvaardigheids-toetsen ontwikkeld door het CITO. De vragen vereisten dat leerlingen het centrale thema van een tekst aanwezen, een verbinding legden tussen verschillende delen van een tekst of gevolgtrekkingen maakten uit de informatie vervat in een tekst.

2.5 Afhankelijke variabelen

De distractie-effecten van achtergrondtelevisie werden vastgesteld aan de hand van vier afhankelijke variabelen: (a) de proportie juist beantwoorde vragen bij de drie gemakkelijke huiswerkopgaven; (b) de proportie juist beantwoorde vragen bij de twee moeilijke huiswerkopgaven; (c) de hoeveelheid tijd die gemiddeld per vraag aan de beantwoording werd besteed bij het gemakkelijke huiswerk; (d) de hoeveelheid tijd die gemiddeld per vraag aan de beantwoording werd besteed bij het moeilijke huiswerk.

2.6 Procedure

De deelnemers kwamen op eigen gelegenheid naar het universiteitsgebouw waar het experiment werd gehouden, daarbij geholpen door een van tevoren thuis gestuurde routebeschrijving. Bij het gebouw aangekomen, meldde de leerling zich bij de receptie, en werd hij of zij opgehaald door één van de twee getrainde proefleiders die met de afname van het experiment waren belast. De proefleider bracht de deelnemer naar de onderzoeksruijme. De afname geschiedde individueel. Door het lot was bepaald welke proefleider aan welke deelnemers was toebedeeld.

In de experimenteerruimte stonden een tafel, twee stoelen en een televisietoestel met videorecorder. De beeldbuis was 58 cm. De deelnemer was gezeten aan een tafel die in het midden van de kamer stond opgesteld. Het televisietoestel stond op ooghoogte van de deelnemer, op een afstand van ongeveer drie meter. Het toestel stond schuin tegenover de deelnemer in een hoek van de kamer. Bij rechtshandige deelnemers stond het televisietoestel 45 graden naar rechts en bij linkshandige deelnemers 45 graden naar links. De proefleider zat gedurende het hele experiment schuin tegenover de deelnemer aan tafel in de hoek waar geen televisietoestel was opgesteld.

Nadat de deelnemer had plaatsgenomen, legde de proefleider de bedoeling van het onderzoek uit: 'Je doet mee aan een onderzoek over huiswerk. We hebben een vijftal taal- en rekentaken. Deze taken zouden je huiswerk kunnen zijn. Bij elke taak geef ik in het begin aan wat de bedoeling is. Ik blijf hier, als je klaar bent, dan zeg je het maar.' Voor de perso-

nen in de controleconditie volgde direct een uitleg van de eerste huiswerktaak.

De personen in de televisiecondities werd gezegd: 'Er zijn veel kinderen die hun huiswerk bij de televisie maken. We willen graag weten hoe goed kinderen in staat zijn hun huiswerk bij de televisie te maken. Daarom zetten we nu de televisie aan.' Vervolgens werd de televisie op huiskamersterkte aangezet (gemiddeld 60 – 70 dB) en werd de eerste huiswerktaak uitgelegd. De televisie bleef vanaf dat moment continu aanstaan, ook tijdens de momenten waarop de huiswerktaken werden uitgelegd. Na iedere huiswerktaak noteerde de proefleider de tijd die de deelnemer over de taak had gedaan en werd de bedoeling van de volgende taak uitgelegd. Wanneer de deelnemer alle taken had gemaakt, werd het experiment afgesloten door de beloofde beloning van 20 gulden te overhandigen.

3 Resultaten

Datacontroles wezen uit dat de verdelingen van de afhankelijke variabelen (de proportie juiste antwoorden en de hoeveelheid tijd besteed aan het beantwoorden van de huiswerktaken) voldeden aan de aannames die ten grondslag liggen aan variantie-analyse. Een uitbijter-analyse met gebruikmaking van afstandsmaten van Cook (1977) leverde geen datapunten op met een afstand groter dan 1 (Cook & Weisberg, 1982). Daarom werden de scores van alle 90 deelnemers toegelaten tot de analyse.

3.1 Huiswerkprestaties

Een 3 (soap vs. videoclips vs. stilte) X 2 (gemakkelijke vs. moeilijke taken) variantie-analyse werd uitgevoerd op de prestatiescores (proportie juist beantwoorde vragen). In deze analyse werden de leerlingen uit elk blok van drie deelnemers (die gematched waren op leesvaardigheid) behandeld alsof zij één persoon waren die onderworpen was aan herhaalde metingen (Kirk, 1968; Tabachnick & Fidell, 1989). De factor achtergrondconditie werd dus – evenals de moeilijkheidsgraad – behandeld als een binnen-proefpersonen-factor. Voor elk van de drie onderzoeksgroepen, geeft Tabel 1 de proportie juist beantwoorde vragen, voor zowel gemakkelijke als moeilijke huiswerktaken.

Tabel 1

Proportie goede antwoorden bij gemakkelijke en moeilijke huiswerktaken, gemaakt met of zonder televisie op de achtergrond

	Huiswerktaken			
	Gemakkelijk		Moeilijk	
	M	SD	M	SD
Achtergrondtelevisie				
Soap opera	.88	.05	.56	.12
Videoclips	.87	.06	.60	.12
Controlegroep (stilte)	.88	.06	.63	.13

Er was sprake van een hoofdeffect van moeilijkheidsgraad, $F(1, 29) = 670.16$, $MSE = .01$, $p = .000$, $\eta^2 = .96$, uitgaande van $\eta^2 = SS_{\text{effect}} / (SS_{\text{effect}} + SS_{\text{error}})$ (Tabachnick & Fidell, 1989). De waarde van η^2 van .96 geeft aan dat er sprake was van een groot hoofdeffect (Cohen, 1988). Het hoofdeffect hield in dat, conform onze bedoeling, de taken die geacht werden relatief gemakkelijk te zijn inderdaad beter waren gemaakt dan de als moeilijk bestempelde taken. Er was geen hoofdeffect van achtergrondconditie, $F(2, 58) = 1.55$, $MSE = .01$, $p > .20$, hetgeen wil zeggen dat de leerlingen die het huiswerk bij de televisie maakten de taken ongeveer even goed maakten als de controlegroep. Er was echter wel een significante interactie tussen moeilijkheidsgraad en achtergrondconditie, $F(2, 58) = 3.88$, $MSE = .00$, $p < .026$; $\eta^2 = .13$. De gevonden waarde van η^2 duidt erop dat er sprake was van een middelgroot effect (Cohen, 1988). Post-hoc vergelijkingen met gepaarde t -tests wezen uit dat leerlingen die waren blootgesteld aan de soap-programma's significant minder goed presteerden dan de leerlingen uit de controlegroep, een effect dat beperkt bleef tot de moeilijke huiswerktaken, $t(29) = 2.66$, $p = .012$. Het laatste effect behield zijn significantie wanneer een Bonferroni-correctie werd toegepast, leidend tot een stringenter significantieniveau ($p < .013$). Geen enkele van de overige gepaarde t -tests leverde significante verschillen op, alle p 's $> .23$.

In een aanvullende analyse werd nagegaan of het bij moeilijke taken gevonden distractie-effect van soap-programma's verschillend uitviel voor de moeilijke leestaak (begrijpend-lezen-taak) en de moeilijke rekentaak (redactiesommen). Blijkens een 3 (soap vs. video-

clips vs. stilte) X 2 (begrijpend lezen vs. redactiesommen) binnen-proefpersonen-variantie-analyse was er geen significante interactie tussen achtergrondconditie en type taak ($F(2,58) = .19$, $MSE = .02$, $p = .831$), hetgeen wil zeggen dat het distractie-effect van het soapprogramma bij beide typen taken ongeveer even sterk was.

3.2 Huiswerktijd

Een 3 (soap vs. videoclip vs. stilte) X 2 (gemakkelijke vs. moeilijke taken) variantie-analyse werd uitgevoerd op de huiswerktijd (aantal seconden dat gemiddeld per vraag aan de beantwoording werd besteed), met achtergrondconditie en moeilijkheidsgraad als binnen-proefpersonen-factoren. Voor elk van de drie achtergrondcondities, vermeldt Tabel 2 het gemiddeld aantal seconden besteed per vraag bij gemakkelijke en moeilijke huiswerktaken.

Tabel 2
Gemiddeld aantal seconden besteed per vraag bij gemakkelijke en moeilijke huiswerktaken, gemaakt met of zonder televisie op de achtergrond

	Huiswerktaken			
	Gemakkelijk		Moeilijk	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Achtergrondtelevisie				
Soap opera	6.35	1.53	43.69	11.40
Videoclips	6.51	1.80	44.75	10.18
Controlegroep (stilte)	6.06	1.61	42.50	10.93

Er was een hoofdeffect voor moeilijkheidsgraad, $F(1, 29) = 945.73$, $MSE = 66.35$, $p = .000$, $\eta^2 = .97$, hetgeen inhield dat de leerlingen gemiddeld meer tijd besteedden aan moeilijke vragen dan aan gemakkelijke vragen. Het effect van achtergrondconditie was niet significant, $F(2, 58) = .47$, $MSE = 59.02$, $p = .629$. De interactie tussen achtergrondconditie en moeilijkheidsgraad was evenmin significant, $F(2, 58) = .32$, $MSE = 37.91$, $p = .729$. De aanwezigheid van achtergrondtelevisie had dus geen invloed op de hoeveelheid tijd die de leerlingen aan de beantwoording van de huiswerkvragen besteedden.

4 Discussie

De resultaten van ons experiment waren slechts ten dele in overeenstemming met onze hypothesen. Conform de verwachting was de interferentie van achtergrondtelevisie met de huiswerkprestaties het sterkst bij de moeilijke huiswerktaken en bij het gebruik van een soapprogramma op de achtergrond. Maar tegen de verwachting in, werden bij de gemakkelijke huiswerktaken en bij het gebruik van videoclips op de achtergrond in het geheel geen distractie-effecten gevonden.

De opzet van het huidige experiment laat geen dwingende conclusies toe omtrent het soort psychologisch mechanisme dat verantwoordelijk was voor de bij moeilijke huiswerktaken en het soapprogramma gesignaleerde distractie-effecten. De bevinding dat wel bij moeilijke en niet bij gemakkelijke huiswerktaken een distractie-effect werd aangetroffen is in beginsel verenigbaar met drie typen theoretische verklaringen, namelijk capaciteitsinterferentie, oriënterende reacties en fysiologische activatie. Het is zeer wel denkbaar dat de combinatie van moeilijke taken en soapprogramma's tot een overschrijding van de verwerkingscapaciteit van leerlingen, en dus tot capaciteitsinterferentie, heeft geleid. Hoewel we niet systematisch geobserveerd hebben hoe vaak de leerlingen door auditieve cues naar de televisie gingen kijken, lijkt het eveneens plausibel dat oriënterende reacties een verklaring voor het gevonden distractie-effect bieden. Zoals eerder opgemerkt bevat de audiotrack van soapprogramma's veel auditieve cues die een leerling kunnen aansporen de aandacht op het beeld te richten.

Het lijkt echter onwaarschijnlijk dat het gevonden distractie-effect voor soapprogramma's kan worden toegeschreven aan een verhoogde fysiologische activatie. Ten eerste omdat dit programmatype de kenmerken ontbeert die blijkens onderzoek activerend werken, namelijk spanning, hilarische humor en agressie (Zillmann, 1982). Ten tweede omdat in geval fysiologische activatie het distractie-effect zou verklaren, we ook hadden moeten vinden dat videoclips afleidend zouden werken, omdat dit snel gesneden programmatype geacht mag worden minstens evenveel arousal op te roepen als soapprogramma's.

Het lijkt eveneens onwaarschijnlijk dat het gevonden distractie-effect werd veroorzaakt door structurele interferentie. Indien het gevonden distractie-effect veroorzaakt zou zijn door structurele interferentie, zou men immers verwachten dat de sterk talige soapprogramma's vooral afleidend zouden werken bij de meest talige moeilijke huiswerктаak, namelijk de begrijpend leestest. Dit bleek echter niet het geval te zijn: het distractie-effect was ongeveer even groot bij de moeilijke leestaak als bij de moeilijke rekentaak.

De bevinding dat zich bij de gemakkelijke huiswerktaken in het geheel geen distractie-effecten voordeden, is niet zo verrassend, omdat in eerder onderzoek bij gemakkelijke taken soms wel (bv. Patton et al., 1981) en soms geen distractie-effecten (bv. Armstrong & Greenberg, 1990) werden aangetroffen.

Verrassender is dat we bij videoclips geen distractie-effecten aantreffen. Immers, eerder onderzoek liet zien dat reeds louter auditief gepresenteerde vocale muziek de prestaties op cognitieve taken kan verlagen (bv. Martin et al., 1988; Salamé & Baddeley, 1989), hetgeen de verwachting wettigde dat als aan de muziek ook nog beelden zijn toegevoegd er minstens even grote of sterkere distractie-effecten zouden optreden. In de betreffende studies ging het echter om cognitieve taken die een sterk beroep deden op de aandachtscapaciteit, omdat zij onder tijdsdruk moesten worden uitgevoerd (Kahneman, 1973). Daarnaast werden de liedjes die in deze experimenten werden gebruikt in de moedertaal van de deelnemers gezongen, een factor die eveneens de kans op interferentie-effecten doet toenemen (Martin, et al., 1988). Mogelijk werden in onze experimenten bij videoclips geen distractie-effecten aangetroffen, omdat de clips in een vreemde taal werden gezongen en de tijd benodigd voor het huiswerk werd vrijgelaten.

Het verbaast ons eveneens dat achtergrondtelevisie geen enkel effect had op de hoeveelheid tijd die de leerlingen aan de huiswerktaken besteedden, omdat eerder onderzoek suggereert dat als de tijd die beschikbaar is voor de taak wordt vrijgelaten, achtergrondtelevisie vooral bij moeilijke taken de afronding kan vertragen.

De gevonden distractie-effecten van achtergrondtelevisie deden zich niet alleen in een

beperkt aantal situaties voor (alleen bij moeilijke taken en bij soap-series), maar voorzover er negatieve effecten werden aangetroffen waren die ook betrekkelijk klein (de proportie correct beantwoorde vragen bij de moeilijke huiswerктаak werd door soap-programma's op de achtergrond met zeven procentpunten verlaagd).

Een mogelijke verklaring voor de betrekkelijk geringe omvang van de gevonden distractie-effecten is dat de gegeven instructie de leerlingen kan hebben gemotiveerd om de achtergrondtelevisie zoveel mogelijk te negeren. De letterlijke instructie die stelde dat we 'graag [willen] weten hoe goed kinderen in staat zijn om hun huiswerk bij de televisie te maken' kan de leerlingen hebben uitgedaagd om te bewijzen dat zij inderdaad heel wel in staat waren om hun huiswerk bij de televisie te maken, met als mogelijk gevolg dat de distractie-effecten werden ingeperkt. Het onderzoek van Armstrong en Greenberg (1990) heeft aangetoond, dat wanneer leerlingen worden aangemoedigd om de televisie zoveel mogelijk te negeren, de distractie-effecten inderdaad verkleind kunnen worden.

Een tweede mogelijke verklaring voor de beperkte omvang van de gevonden distractie-effecten is dat de lijfelijke aanwezigheid van de proefleider tijdens het maken van de taken de leerlingen gemotiveerd kan hebben om de aandacht zoveel mogelijk op de huiswerktaken te richten. De deelnemers kunnen daarom extra hun best hebben gedaan om de televisie zoveel mogelijk te negeren, met het gevolg dat de distractie-effecten werden verkleind. De aanwezigheid van een proefleider die met een stopwatch in de hand de deelnemer op de vingers keek, kan mede verklaren waarom geen distractie-effecten in termen van tijd gevonden werden. Omdat de deelnemers zich bekeken voelden en de tijd die zij per taak nodig hadden ter plaatse geklokt werd, voelden zij zich wellicht niet vrij om tussentijds hun pen neer te leggen en een tijdje naar de televisie te kijken.

In de thuissituatie komt het zelden of nooit voor dat leerlingen tijdens het huiswerk op hun vingers worden gekeken of dat wordt bijgehouden hoe lang zij erover doen. Evenmin is het in de thuissituatie gebruikelijk dat kinderen worden aangemoedigd om aan te tonen hoe goed zij in staat zijn om hun huiswerk bij de televisie te maken. Op deze twee punten liet de ecologi-

sche validiteit van het uitgevoerde experiment dus te wensen over. In een vervollexperiment dat thans lopende is, zullen we daarom nagaan hoe de distractie-effecten uitvallen onder omstandigheden die meer vergelijkbaar zijn met de thuissituatie, namelijk als de leerling zich tijdens het huiswerk niet bekeken voelt en als geen instructie wordt gegeven die de deelnemer aanmoedigt om de televisie te negeren.

Literatuur

- Aarnoutse, C. A. J. (1992). *Synoniementest, Tegenstellingentest, Voegwoordentest, Hoofdgedachtentest*. Nijmegen: Berkhout.
- Anderson, D. R., & Lorch, E. P. (1983). Looking at television: Action or reaction? In J. Bryant & D. R. Anderson (Eds.), *Children's understanding of television: Research on attention and comprehension* (pp. 1-33). New York: Academic Press.
- Anderson, K. J., & Revelle, W. (1982). Impulsivity, caffeine, and proofreading: A test of the Easterbrook hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 614-624.
- Armstrong, G. B. (1993). Cognitive interference from background television: Structural effects on verbal and spatial processing. *Communication Studies*, 44, 56-70.
- Armstrong, G. B. (1997, May). *Context and effects of background television on reading: A test of television effects on encoding, recall, and recognition*. Paper presented at the Annual Conference of the International Communication Association, Montreal, Canada.
- Armstrong, G. B., Boiarsky, G. A., & Mares, M. L. (1991). Background television and reading performance. *Communication Monographs*, 58, 235-253.
- Armstrong, G. B., & Greenberg, B. S. (1990). Background television as an inhibitor of cognitive processing. *Human Communication Research*, 16, 355-386.
- Armstrong, G. B., & Sopory, P. (1997). Effects of background television on phonological and visuo-spatial working memory. *Communication Research*, 24, 459-506.
- Beentjes, J. W. J., Koolstra, C. M., & Van der Voort, T. H. A. (1996). Combining background media with doing homework: Incidence of background media use and perceived effects. *Communication Education*, 45, 59-72.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cook, R. D. (1977). Detection of influential observations in linear regression. *Technometrics*, 19, 15-18.
- Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman & Hall.
- Cool, V. A., & Yarbrough, D. B., Patton, J. E., Runde, R., & Keith, T. Z. (1994). Experimental effects of radio and television distractors on children's performance on mathematics and reading assignments. *Journal of Experimental Education*, 62, 181-194.
- Daoussis, L., & McKelvie, S. J. (1986). Musical preferences and effects of music on a reading comprehension test for extraverts and introverts. *Perceptual and Motor Skills*, 62, 283-289.
- Etaugh, C., & Ptasnik, P. (1982). Effects of studying to music and post-study relaxation on reading comprehension. *Perceptual and Motor Skills*, 55, 141-142.
- Fogelson, S. (1973). Music as a distractor on reading-test performance of eighth grade students. *Perceptual and Motor Skills*, 36, 1265-1266.
- Furnham, A., & Bradley, A. (1997). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extraverts. *Applied Cognitive Psychology*, 11, 445-455.
- Furnham, A., Gunter, B., & Peterson, E. (1994). Television distraction and the performance of introverts and extraverts. *Applied Cognitive Psychology*, 8, 705-711.
- Graydon, J., & Eysenck, M. W. (1989). Distraction and cognitive performance. *European Journal of Cognitive Psychology*, 1, 161-179.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Kan, N. van, & Aarnoutse, C. (1993). *Begrijpend leestest bestemd voor het eerste leerjaar van het MAVO, HAVO en VWO*. Nijmegen: Berkhout.
- Kiger, D. M. (1989). Effects of music information load on a reading comprehension task. *Perceptual and Motor Skills*, 69, 531-534.
- Kirk, R. E. (1968). *Experimental design: Procedures for the behavioral sciences*. Belmont, CA: Brooks/Cole.
- Martin, R. C., Wogalter, M. S., & Forlano, J. G. (1988). Reading comprehension in the presence of unat-

tended speech and music. *Journal of Memory and Language*, 27, 382-398.

- Matthews, K. A., & Brunson, B. I. (1979). Allocation of attention and the Type A coronary-prone behavior pattern. *Journal of Personality and Social Psychology*, 37(11), 2081-2090.
- Patton, J. E., Routh, D. K., & Offenbach, S. I. (1981). Televised classroom events as distractors for reading-disabled children. *Journal of Abnormal Child Psychology*, 9, 355-370.
- Patton, J. E., Stinard, T. A., & Routh, D. K. (1983). Where do children study? *Journal of Educational Research*, 76, 280-286.
- Salamé, P., & Baddeley, A. D. (1989). Effects of background music on phonological short-term memory. *Quarterly Journal of Experimental Psychology Human Experimental Psychology*, 41, 107-122.
- Sanders, G. S., & Baron, R. S. (1975). The motivating effects of distraction on task performance. *Journal of Personality and Social Psychology*, 32, 956-963.
- Strube, M. J., Turner, C. W., Patrick, S., & Perrillo, R. (1983). Type A and Type B attentional responses to aesthetic stimuli: Effects on mood and performance. *Journal of Personality and Social Psychology*, 45, 1369-1379.
- Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics*. New York: Harper & Row.
- Voort, T. H. A. van der, Beentjes, J. W. J., Bovill, M., Gaskell, G., Koolstra, C. M., Livingstone, S., & Marseille, N. (1998). Young people's ownership and uses of new and old forms of media in Britain and the Netherlands. *European Journal of Communication*, 13, 457-477.
- Vos, T. de (1992). *Tempo Test Rekenen*. Nijmegen: Berkhout.
- Vos, T. de (1995). *Niveau Test Rekenen*. Nijmegen: Berkhout.
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology of Psychology*, 18, 459-482.
- Zillmann, D. (1982). Television viewing and arousal. In D. Pearl, L. Bouthilet, & J. Lazar (Eds.), *Television and behavior: Ten years of scientific progress and implications for the eighties* (Vol. 2, 53-67). Washington, U.S. Government Printing Office.

Auteurs

M. M. Pool is ontwikkelingspsychologe en methodologe en als Assistent in Opleiding verbonden aan de vakgroep Algemene en Gezinspedagogiek, Universiteit Leiden. Zij doet promotie-onderzoek naar de invloed van achtergrondmedia op de huiswerkprestaties van middelbare scholieren.

T. H. A. van der Voort is als hoogleraar gezinspedagogiek verbonden aan de Universiteit Leiden en is wetenschappelijk directeur van ISED.

J. W. J. Beentjes is als universitair hoofddocent verbonden aan de vakgroep Communicatiewetenschap, Katholieke Universiteit Nijmegen.

C. M. Koolstra is als universitair docent verbonden aan de vakgroep Algemene en Gezinspedagogiek, Universiteit Leiden.

Correspondentie-adres: M. M. Pool, Afdeling Algemene en Gezinspedagogiek, Universiteit Leiden, Postbus 9555, 2300 RB Leiden, Telefoon: 071-5274080, Fax: 071-5273945, E-mail: Pool@fsw.leidenuniv.nl

Abstract

The impact of background television on the performance of easy and difficult homework assignments

M.M. Pool, T.H.A. van der Voort, J.W.J. Beentjes and C.M. Koolstra. *Pedagogische Studiën*, 76, 350-360.

An experiment was conducted to investigate the effects of two types of background television programs on students' performance on easy and difficult homework assignments. Students in Grade 8 ($N = 90$) were matched on reading proficiency. Subsequently, the participants were randomly assigned to one of the three experimental conditions in which they did their homework (a) while a Dutch soap opera was on, or (b) while English-language video clips were on, or (c) in quiet. Results indicated that the execution of difficult homework was hindered when a soap opera was on. The distraction effect of the soap opera was expressed in a performance decrement. Background television did not lead to an extension of performance time.

Sociaal-wetenschappelijke literatuur over onderwijs in Zuid-Afrika: van verzuiling tot eensgezindheid vanuit verscheidenheid*

C.C. Wolhuter

Samenvatting

Het doel van dit onderzoek is het in kaart brengen van de sociaal-wetenschappelijke literatuur over onderwijs en de evaluatie van onderwijs in Zuid-Afrika, met name in het licht van de grote maatschappelijke veranderingen na 1994. In de literatuur uit het tijdperk van vóór 1994 kunnen drie denkrichtingen onderscheiden worden: een conservatieve school die de status-quo wilde behouden, een liberale, en een neo-marxistische die uitging van het conflict. Volgens eerstgenoemde school was het toenmalige onderwijsstelsel onlosmakelijk verbonden met een onwrikbare en boven alle kritiek verheven maatschappelijke orde; volgens laatstgenoemde was hetzelfde systeem volstrekt verwerpelijk en onhervormbaar en was een totale maatschappelijke (op politiek, sociaal, economisch en onderwijsgebied) hervorming noodzakelijk. Deze situatie heeft een constructieve wisselwerking tussen wetenschap en onderwijs in de weg gestaan; de wetenschap kon onmogelijk een bijdrage leveren aan de verbetering van de praktijk. Sinds de onderwijshervormingen en de maatschappelijke veranderingen van na 1994 bestaan er nog steeds drie paradigma's in de sociaal-wetenschappelijke literatuur, maar nu is er wel brede eensgezindheid ten aanzien van de grondslagen en de doelstellingen van het nieu-

we onderwijssysteem. Zinnvolle interactie tussen theorie en praktijk behoort nu wel tot de mogelijkheden. De meest veelbelovende manier om een theorie tot stand te brengen die als richtlijn kan gaan fungeren voor de hervormingen in het onderwijs in Zuid-Afrika is het opzetten van comparatieve studies met andere landen die ervaring hebben met de doelstellingen die Zuid-Afrika voor haar onderwijs heeft geadopteerd.

De probleemstelling

Het onderwijs in Zuid-Afrika ondergaat op dit moment radicale veranderingen als onderdeel van de brede maatschappelijke reconstructie van het land na 1994. Het ideaal is een dialectische wisselwerking tussen theorie (onderwijskunde) en praktijk (de praktijk van het onderwijs) die bestaat uit reflectie en evaluatie van de praktijk door de wetenschap, die resulteren in verbeteringen die op hun beurt weer de basis vormen voor een nieuwe reeks reflecties en evaluaties. Hierdoor rijst de vraag wat de huidige stand is van de wetenschappelijke theorievorming met betrekking tot het onderwijs in Zuid-Afrika. Vanuit het perspectief van het buitenland is de vraag des te prangender in het licht van de academische boycot die Zuid-Afrika gedurende enkele decennia getroffen heeft (vgl. Harricombe en Lancaster, 1995, om een inzicht te krijgen in de omvang en de gevolgen van deze boycot). Dit heeft ervoor gezorgd dat Zuid-Afrika voor het buitenland een blinde vlek op de internationale landkaart van het onderwijs is geworden. Daartegenover staat een hernieuwde belangstelling voor de onderwijsontwikkelingen in Zuid-Afrika na de opheffing van de boycot vanwege de huidige maatschappelijke veranderingen en de onderwijsaspecten daarvan. Deze hernieuwde belangstelling is onder meer zichtbaar in speciale themanummers die twee vaktijdschriften aan de onderwijshervormingen in Zuid-Afrika hebben gewijd (*Comparative Education*, jaargang 31(2), 1995; *International Review of Education*, jaargang 43(2), 1997). Ten slotte is zo'n studie noodzakelijk voor het verkennen van mogelijkheden voor vergelijkende studies – een gids voor toekomstige hervormingen in Zuid-Afrika, verrijkt met de ervaringen uit het buitenland. Het doel van dit onderzoek is dus

* Ik wil hierbij mijn dank betuigen aan Prof. M. Kleijwegt van de Universiteit van Zuid-Afrika voor de taalverzorging in verband met het omzetten van het manuscript in aanvaardbaar Nederlands.

verslag te doen van de evaluatie van het onderwijs in Zuid-Afrika, vooral in het licht van de maatschappelijke veranderingen na 1994.

Het onderzoek is beperkt tot de onderwijssociologische literatuur, omdat, zoals hierboven al werd aangeduid, de aanleiding tot deze studie was gelegen in de maatschappelijke veranderingen in Zuid-Afrika sinds 1994. De hervormingen in het onderwijs zijn bestudeerd als onderdeel van deze maatschappelijke veranderingen. De centrale vraag die hierbij werd gesteld is in hoeverre deze veranderingen een uitwerking hebben gehad op de sociaal-wetenschappelijke literatuur over het onderwijs.

Aangezien de sociaal-wetenschappelijke literatuur over het onderwijs in Zuid-Afrika alleen te begrijpen valt tegen de achtergrond van de geschiedenis van Zuid-Afrika in het algemeen en de geschiedenis van het onderwijs in het bijzonder wordt hier ten behoeve van de ongeïnjwilde lezer een overzicht geboden van het onderwijs voor en na 1994 alvorens word gegaan tot een bespreking van de sociaal-wetenschappelijke literatuur uit diezelfde periodes.

1 De groei van segregatie binnen het onderwijs in Zuid-Afrika: drie paradigma's

1.1 De geschiedenis van het onderwijs in Zuid-Afrika: Segregatie

Formeel onderwijs in Zuid-Afrika werd ingesteld met de vestiging van de V.O.C. in de Kaap in 1652. Onderwijs gescheiden naar ras was vanaf de beginjaren een kenmerk van de geschiedenis van Zuid-Afrika. Onderwijs aan blanken was de verantwoordelijkheid van het koloniale bestuur en later van de regering van Zuid-Afrika, terwijl onderwijs aan zwarten werd gegeven door zendelingenorganisaties uit Europa. Met de machtsovername door de Nationale Partij (wat hoofdzakelijk met steun van de Afrikaners plaatsvond) werd segregatie op basis van ras vastgelegd in een reeks gedetailleerde wetten (het systeem van Apartheid). De bedenkers van Apartheid geloofden dat de scheiding van rassen (en van de verschillende

cultuurgroepen binnen de zwarte bevolkingsgroep) elke cultuurgroep in staat zou stellen om zich afzonderlijk te ontwikkelen en welvarend te zijn. Voor dit doel werden tien autonome staten (zgn. 'thuislanden' of 'Bantustans') voor de verschillende groepen binnen de grenzen van Zuid-Afrika gesticht. Elke groep zou over een eigen regering, eigen scholen, universiteiten, enz. de beschikking hebben. Eén van de eerste activiteiten die de regering van de Nationale Partij op het gebied van onderwijs ondernam was het instellen, in 1949, van een speciale commissie, de commissie Eiselen, die tot doel had het onderwijs aan zwarten te onderzoeken. Dr. W.W.M. Eiselen was de voorzitter van deze commissie. Haar taak was het formuleren van voorschriften en doelstellingen met betrekking tot onderwijs aan zwarten als een afzonderlijk ras, gebaseerd op hun eigen cultuur en uitgaande van hun eigen behoeftes. Naar aanleiding van het rapport van de commissie Eiselen is de Bantoe Onderwijs wet (Wet nr. 47 uit 1953) van kracht geworden. Verantwoordelijkheid voor en controle op de onderwijsvoorziening aan zwarten is toen overgedragen van de zendelingenorganisaties aan de regering die voor het doel een Ministerie voor Bantoe Onderwijs in het leven riep. In de loop der tijd zijn afzonderlijke ministeries (en administratieve structuren) ingesteld voor alle andere bevolkingsgroepen: Indiërs en kleurlingen ('Moelatte') zowel als in alle thuislanden. De Bantoe Onderwijs wet heeft vanaf het begin veel verzet opgeroepen. Onder Engelssprekende blanken, die de economische sector beheersten, terwijl de Afrikaners het politieke toneel domineerden, bestond grote twijfel over de praktisch uitvoerbaarheid van een rigoreus segregatie-beleid. De volgende zinsnede uit een rapport uit 1952 opgesteld door het 'Institute of Race Relations' (gezeteld aan de Engelse Universiteit van de Witwatersrand) in antwoord op het wetsontwerp over het Bantoe onderwijs kan ter illustratie dienen: '... Africans are not culturally, economically and politically independent, but they are an integral part of South Africa ... education must be directed towards training the individual for the requirements of modern society' (geciteerd uit Mphahlele & Mmincle, 1997). Het meeste verzet kwam echter van de kant van de zwarte gemeenschap.

De meest vooraanstaande politieke organisatie onder zwarte Zuid-Afrikaners was het ANC. Deze organisatie was in 1912 opgericht als gevolg van bezwaren tegen de feit dat stemrecht was voorbehouden aan leden van de blanke bevolkingsgroep. Het belangrijkste uitgangspunt van het ANC was het delen van macht en het tot stand brengen van een gemeenschappelijk Zuid-Afrika waaraan alle bevolkingsgroepen – zwarten, blanken, Indiërs en kleurlingen – zouden deelhebben (vgl. Karis & Gerhart, 1977; Pampallis, 1991) – integratie als tegenhanger van het beleid van segregatie gevoerd door te regering. De zwarte gemeenschap en het ANC stonden zeer wantrouwend tegenover de Bantoe Onderwijs Wet. Ze waren bang dat Bantoe onderwijs van een lager niveau zou zijn en dat het met name ontworpen was om zwarten te veroordelen tot een positie van permanente minderwaardigheid binnen de Zuid-Afrikaanse samenleving (vgl. Karis & Gerhart, 1977). Op een nationale bijeenkomst in 1954 riep het ANC op om de Bantoe Onderwijs Wet af te wijzen (vgl. Karis & Gerhart, 1977). Tijdens een grote bijeenkomst van 'Mensen van Zuid-Afrika' die het ANC in 1955 organiseerde is het 'Vrijheidsmanifest' ('Freedom Charter') opgesteld. De openingszin van dit manifest luidt als volgt: 'We, the people of South Africa, declare ... that South Africa belongs to all who live in it, black and white (geciteerd uit van Jaarsveld, 1971). Het manifest spreekt zich uit tegen ongelijkheid en ten gunste van: '... equal rights and opportunities ... for all our people ... All our people shall have equal right to use their own languages and to develop their own culture and customs ... The doors of learning and culture shall be opened: the government shall discover, develop and encourage national talent for the enhancement of our cultural life. All the cultural treasures of mankind shall be opened to all. The aim of education should be to teach the youth to love their people and their culture, to honour human brotherhood, liberty and peace. Education shall be free, compulsory, universal and equal for all children' (geciteerd uit van Jaarsveld, 1971). Nadat dit standpunt in 1956 door het ANC congres is aanvaard is het 'Vrijheidsmanifest' het officiële beleidsdocument van het ANC geworden. Het 'Vrijheidsmanifest' werd een van de centrale richtlijnen van het ANC en is voortdu-

rend de basis geweest voor later politiek beleid.

Het verzet van zwarte onderwijsleiders, onderwijzers en leerlingen tegen de Bantoe Onderwijs Wet kwam tot uiting in schoolboycots en betogingen en is tot volle uitbarsting gekomen in de Soweto-onlusten van 1976-1977, die zich vervolgens over het hele land verspreid hebben (vgl. Christie, 1991). Het regeringsbeleid van naar ras gescheiden onderwijs is op die manier getart en vervolgens als ongelijk en onwettig veroordeeld. Gedurende de tweede helft van de jaren '80 verkeerde het zwarte onderwijs voortdurend in een crisis, gekenmerkt door boycots van leerlingen, nauw verweven met het brede verzet tegen de regering en het politieke bestel. Dat vond op zulke grote school plaats dat onderwijs in zwarte scholen in feite tot stilstand kwam (vgl. Christie, 1991; Lemmer, 1993). Als gevolg hiervan was het onderwijs tegen het einde van het tijdperk van het vorige politieke bestel in een legitimiteitscrisis terechtgekomen – het model van naar ras gescheiden scholen werd door een meerderheid van de bevolking (zwarten vormen 76,3% van de totale bevolking van Zuid-Afrika) afgewezen.

1.2 De sociaal-wetenschappelijke studie van het onderwijs in Zuid-Afrika vóór 1994

Schriewer en Keiner (1992) betogen dat sociale studies het kenmerk dragen van hun oorsprong. Sociaal onderzoek is voornamelijk ingebed in en beïnvloed door het institutionele kader wat bepaald wordt voor verschillende nationale of socio-culturele contexten. De laatste worden gevormd door specifieke sociale situaties. Aangezien het systeem van strikt op basis van ras geschieden onderwijsvormen de schepping was van een bepaald deel van de bevolking (de blanke Afrikaner op wie de Nationale Partij steunde), en niet de steun had van het andere deel van de blanken (de Engelsprekenden) en als onaanvaardbaar werd gezien door de zwarte meerderheid, is het te verwachten dat de sociaal-wetenschappelijke literatuur over onderwijs ook uiteenvalt in drie verschillende richtingen. Dit word geïllustreerd aan de hand van een uiteenzetting over de sociaal-wetenschappelijke literatuur, met betrekking tot het onderwijs uit het tijdperk vóór 1994.

1.2.1 Het conservatieve pro-status-quo paradigma

Dit paradigma is hoofdzakelijk door Afrikaanssprekende blanke onderwijskundigen die aan Afrikaanse universiteiten doceerden voorgestaan. Hun vereniging was de 'Opvoedkunde Vereniging van Suid-Afrika'. Van de literatuur afkomstig uit deze stroming zijn de belangrijkste werken: J.C. Coetzee (1963) *Onderwys in Suid-Afrika: 1952-1960* (Pretoria: J.L. van Schaik); P.A. Duminy (1963) *Afsonderlike Onderwysgeriewe vir verskillende rasse-groepe in Suid-Afrika* (Antwerpen: Standaard Boekhandel); R.M. Ruperti (1974) *Die Onderwysstelsel in Suider Afrika* (Pretoria: J.L. van Schaik) en (1977) *Praktiese Perspektiewe uit die Vergelykende Opvoedkunde* (Pretoria: J.L. van Schaik); P.C. Potgieter (1980) *Moral education in South Africa, Journal of Moral education*, 9, 130-133; en H.J.S. Stone (1989) *Onderwysvoorsiening vir nasionale verskeidenheid en gemeenskaplikheid* (Potchefstroom: Potchefstroomse Universiteit vir Christelike Hoër Onderwys). Binnen dezelfde denkrichting valt ook het eerder vermelde verslag van de commissie Eiselen. Melding moet ook gemaakt worden van het rapport over onderwijs van de 'Raad vir Geesteswetenskaplike Navorsing', het 'De Lange verslag', opgesteld in 1981. Dit omvattende onderzoek naar onderwijsvoorzieningen in Zuid-Afrika werd uitgevoerd in opdracht van de regering. Alhoewel het rapport een onderwijssysteem op basis van ras niet expliciet aanbeveelt, wordt er met klem de nadruk gelegd op het in acht nemen van de culturele diversiteit van het land bij het formuleren van onderwijsbeleid (vgl. Raad vir Geesteswetenskaplike Navorsing, 1981). De internationale parallellen van dit paradigma moeten vooral gezocht worden in de raakvlakken die het heeft met het structureel-functionalisme dat ervan uitgaat dat het karakter van een samenleving (wat betreft politiek, economie, culturele samenstelling, enz.) het onderwijssysteem bepaalt. De ondersteuners van deze theorie schreven vanuit de overtuiging dat elke culturele groep in Zuid-Afrika zijn eigen scholen en onderwijssysteem behoorde te hebben. Barnard (1984) vat zijn uiteenzetting over onderwijs in Zuid-Afrika als volgt samen: 'In de RSA (Republiek van Suid-Afrika), met zijn beleid van parallelle ontwikkeling van de verschillen-

de bevolkingsgroepen, voorziet het onderwijssysteem in volkseigen onderwijs met de bedoeling om de talenten van iedere leerling op een optimale manier en zonder druk tot zijn recht te laten komen. Deze differentiatie ... is noodzakelijk omdat er verschillende bevolkingsgroepen ... in het ... land wonen die radicaal van elkaar verschillen wat betreft herkomst, ras, cultuur, levens- en wereldbeschouwing, taal en wat dies meer zij'.

1.2.2 Het liberale paradigma

Dit paradigma is vooral aangehangen door docenten aan Engelissprekende universiteiten en hun vereniging was de 'Kenton Education Association'. De literatuur afkomstig uit deze stroming is: E.G. Malherbe (1925) *Education in South Africa, volume 1, 1652-1922* (Cape Town: Juta) en (1977) *Education in South Africa, volume 2, 1923-1975* (Cape Town: Juta); A.L. Behr (1988) *Education in South Africa: 1652-1988: Origins, Issues and Trends* (Cape Town: Academica); en K.B. Hartshorne (1992) *Crisis and challenge: Black education 1910-1990* (Cape Town: Oxford University Press). Wat betreft de internationale herkomst van deze ideeën, deze zijn vooral terug te vinden in theorie van het menselijk kapitaal ('human capital') (vgl. Behr, 1988) en in theorieën over modernisering (vgl. King, 1979). Volgens deze denkbelden wordt onderwijs niet slechts gezien als een instrument dat bijdraagt aan de economische ontwikkeling en modernisering van Zuid-Afrika, maar bestaat er een grenzeloos vertrouwen in het vermogen van onderwijs om de totale sociaal-economische en politieke problematiek op te lossen, iets wat treffend tot uitdrukking komt in de monografie van G. Bozzoli (1977) *Education is the key to change in South Africa* (Pietermaritzburg: The Natal Witness).

De geleerden die deze theorie aanhangen beschrijven en beoordelen het Zuid-Afrikaanse onderwijssysteem vanuit de liberale principes: Individuele vrijheden en rechten, gelijke kansen, de rechtstaat, een vrije economie en minimale inmenging van de staat in de burgerlijke samenleving worden hoog aangeslagen. Onderwijs gescheiden naar ras wordt verworpen, vooral als dit voor de regering wordt opgelegd, tezamen met de principes van ongelijkheid die hieraan ten grondslag liggen (vgl.

Mackenzie; Moodie, 1994; Bozzoli, 1977; en voor het laatste aspect: King, 1979; Horrell, 1968). De principes komen het sterkst tot uitdrukking in de publicatie van K.B. Hartshorne (1989) *Can Separate be Equal? A Commentary on the White Paper on Education* (Durban: University of Natal Centre for Applied Social Science). Vanuit deze overwegingen is een geleidelijke hervorming van het onderwijssysteem bepleit, echter zonder dat de omverwerping van de politieke en economisch orde als een voorvereiste voor een succesvolle onderwijsvernieuwing is gesteld (vgl. Hartshorne, 1992; Behr, 1988; Moodie, 1994).

1.2.3 Het neo-marxistische conflict paradigma

De aanhangers van deze richting moeten vooral gezocht worden onder zwarte academici werkzaam aan zwarte universiteiten en onder de jongere blanke generatie (van na 1976) verbonden aan Engelssprekende blanke universiteiten. Hun vereniging heet de 'Southern African Society of Education'. Aanzetten tot deze denkbeelden zijn te vinden in J.G. Tabata (1960) *Education for Barbarism* (London: Unity Movement of South Africa) en in het UNESCO-rapport (1972) *Apartheid: its effects on education, science, culture and information* (Paris: UNESCO). Over het algemeen, echter, wordt de publikatie van P. Kallaway (ed.) (1984) *Apartheid and Education: the education of black South Africans* (Johannesburg: Ravan) als de meest vooraanstaande verwoording van dit paradigma gezien. Ook buitenlandse geestverwanten hebben een bijdrage geleverd, bijvoorbeeld, S. Bowles (1993) *Education and the capitalist economy: the politics of cumulative reform, Education with production*, 10, 35-39. Ook het verslag van de 'National Education Policy Investigation' (het 'NEPI-document', 1993) mag hier niet onvermeld blijven.

Wat het internationale kader van deze ideeën betreft interpreteren de auteurs het onderwijssysteem in Zuid-Afrika in termen van de neo-marxistische theorie van economische reproductie, zoals geformuleerd door Bowles, Gintis en Althusser. Bovendien gebruikt Nasson (1988) de theorie van culturele reproductie, zoals ontwikkeld door Bourdieu en Passeron. Naidoo (1990) verklaart het studentenactivisme in Freieriaanse termen van

een transformatie van onderwijs en samenleving in Zuid-Afrika.

Dit paradigma wordt in het bijzonder gekenmerkt door een volslagen afwijzing van het politieke (Apartheid) en economische (kapitalisme) systeem (vgl. Christie, 1991; Cross, 1986). Vanuit dit perspectief wordt onderwijs gescheiden naar ras, en met name het Bantoe onderwijs, gezien als een blauwdruk voor het continueren van blanke overheersing (politiek zowel als economisch) en het veroordelen van zwarten tot een positie van permanente onderworpenheid (vgl. Christie, 1991; Nasson & Samuel, 1990). De titels van de publicaties van M. Nkomo (ed.) (1990) *Pedagogy of Domination: towards a democratic education in South Africa* (Trenton, New Jersey: Africa World Press) en van P. Christie & G. Collins (1982) *Bantu education: Apartheid ideology or labour production?*, *Comparative Education*, 18, 59-75, vatten de grondslagen van dit paradigma kernachtig samen.

1.3 Samenvatting: verzuijing zonder dialoog

In de sociaal-wetenschappelijke literatuur over onderwijs in Zuid-Afrika van vóór 1994 kan men drie denkrichtingen onderscheiden. Ten eerste bestond er een conservatief pro-status-quo paradigma waarin het bestaande onderwijssysteem deel uitmaakte van een onwrikbare en boven alle kritiek verheven maatschappelijke orde. Precies vanwege deze uitgangspunten hebben de aanhangers van deze school zich volledig afgezonderd en zich niet laten verleiden tot een dialoog met diegenen die met hen van mening verschilden. Lijnrecht hiertegenover stonden de zwarte meerderheid en hun academische woordvoerders die het bestaande onderwijssysteem volledig verwierpen. Zij beschouwden het systeem als onherformbaar en waren met niets anders tevreden dan met een totale heroriëntatie van maatschappij en onderwijs. Alle drie de denkrichtingen waren verschanst in hun eigen universiteiten, publicaties en verenigingen, zonder onderling contact. Deze situatie maakte een wisselwerking tussen wetenschap en praktijk, zoals uiteengezet aan het begin van dit artikel, onmogelijk. Aangezien de drie scholen hun bestaan en hun bestaansrecht te danken hadden aan het maatschappelijk bestel en het daaruit voortvloeiende onderwijssysteem wat vóór

1994 van kracht was, rijst natuurlijk de vraag in hoeverre de nieuwe maatschappelijke orde die na 1994 tot stand is gekomen en die een einde maakte aan het oude onderwijssysteem van invloed is geweest op de sociaal-wetenschappelijke studie van het onderwijs.

2 Het samensmelten van denkbeelden in de voetsporen van de onderwijshervormingen en -integratie na 1994

2.1 De onderwijshervormingen na 1994: de grondslagen van het nieuwe onderwijsbestel

In 1994 hebben een nieuwe grondwet en een nieuw politiek bestel hun intrede gedaan. Het ANC heeft de rol van regerende partij van de Nationale Partij overgenomen en een beleid uitgestippeld overeenkomstig haar politieke idealen, zoals uiteengezet onder 1.1, en gegrondvest op de volgende principes die in vele opzichten precies het tegenovergestelde zijn van de beginselen die het beleid van de Nationale Partij bepaalden: gelijkberechtiging, desegregatie, culturele verscheidenheid (multiculturalisme) en democratisering (vgl. Republic of South Africa, 1997; African National Congress 1994-2). In economisch opzicht is er weinig veranderd – het economische systeem dat vóór 1994 bestond (kapitalisme) is grotendeels onaangetast gebleven (vgl. Johnson, 1995; Fataar, 1997). Terwijl er vóór 1994 tussen ANC en haar bondgenoten een socialistische dialoog bestond, heeft de werkelijkheid van de dag geresulteerd in een meer pragmatisch beleid na 1994.

De verschillende op rassenscheiding gebaseerde onderwijssystemen en hun administratieve structuren zijn in 1994 samengevoegd in één Ministerie van Onderwijs. De volgende wetten zijn afgekondigd: de Zuid-Afrikaanse scholen wet (Wet nr. 84 van 1996); de Wet op het Hoger Onderwijs (Wet nr. 10 van 1997) en de Wet op de Zuid-Afrikaanse Kwalifikatie Overheid (Wet nr. 58 van 1995). Verder heeft het Ministerie van Onderwijs een ontwerpvoorstel ('groenskrif') op het universitair onderwijs het licht laten zien en heeft de 'Raad vir Geestewetenskaplike Navorsing' (Engels: Human Sciences Research Council, kortweg

HSRC) op verzoek van de regering een enquête uitgevoerd naar de fysieke infrastructuur van de scholen, wat uitmondde in de publikatie HSRC (1997), School Register of Needs Survey (Pretoria: HSRC). Uit deze documenten blijkt heel duidelijk dat de uitgangspunten van het nieuwe onderwijsbeleid eveneens gelijkberechtiging, desegregatie, culturele verscheidenheid en democratisering zullen zijn (vgl. Republic of South Africa, 1996; African National Congress, 1994-1). Het totale onderwijsbestel is nu afgestemd op het ontwikkelen van de talenten van de gehele bevolking met als uiteindelijk doel de economische ontwikkeling van het land (vgl. Department of Education, 1998; Department of Education, 1997).

Ten einde het onderwijs hiermee in overeenstemming te brengen zijn twee grootschalige projekten ingevoerd: Ten eerste, een onderwijsvorm die gebaseerd is op het aanleren van vaardigheden (OBE = 'outcomes based education') en niet op kennisoverdracht, en een hierop gebaseerde leergang ('curriculum 2005' genaamd) die het oude leersysteem van kennisoverdracht moet vervangen (vgl. Department of Education, 1997; Department of Education, 1998). Ten tweede is de Nationale Kwalifikatie Owerheid in het leven geroepen om het grote aantal onderwijsprogramma's dat in Zuid-Afrikaanse scholen wordt aangeboden te erkennen en doorstroming van leerlingen tussen de verschillende schooltype's te vergemakkelijken en een opleidingsnetwerk voor alle leeftijden te scheppen (vgl. Fataar, 1997; Republiek van Suid-Afrika, 1995).

2.2 Ontluikende veranderingen in het sociaal-wetenschappelijke onderzoek naar onderwijs

Bestudering van de sociaal-wetenschappelijke literatuur die na 1994 is verschenen wijst uit dat de drie stromingen nog steeds bestaan.

2.2.1 Het voormalige conservatieve pro-status-quo paradigma: heroriëntatie en constructieve betrokkenheid

Na 1994 is een aanhoudende stroom literatuur over het onderwijs in Zuid-Afrika vanuit deze hoek verschenen. De belangrijkste literatuur afkomstig uit deze stroming omvat de volgende publicaties: A.J.J. Labuschagne, J.C. Kok & C.P.H. Myburgh (1997) Vryheid, Verdraagsaamheid en die regte van minderhede as

opvoedingsaangeleentheid in 'n demokrasie, *Suid-Afrikaanse Tydskrif vir Opvoedkunde*, 17, 136-143; S.J. Berkhout (1996) Treading the traditional maze: the South African Schools Bill, 1996, in: J. de Groof & E. Bray (eds) *Education under the new constitution in South Africa* (pp. 106-118) (Leuven: Acco); J.C. Claassen (1995) The education system of South Africa: A Comparative Perspective, in E. Dekker & O.J. van Schalkwyk (eds) *Modern Education Systems* (second edition) (pp. 447-495) (Durban: Butterworths); en H.J. Steyn & R. Vanderstraeten (1998) Minority group education in South Africa: a comparative perspective, *South African Journal of Education*, 18, 89-96. De auteurs van deze werken staan volmondig achter de principes waarop het nieuwe onderwijssysteem is gebaseerd, d.w.z. desegregatie, multiculturalisme, democratisering en gelijkberechtiging (vgl. Labuschagne, Kok & Myburgh, 1997; Liebenberg, 1998; Steyn & Vanderstraeten, 1998), alhoewel deze principes lijnrecht tegenover hun uitgangspunten van vóór 1994 staan. Ze nemen ook op een constructieve manier deel aan het debat over het opzetten van een nieuw onderwijssysteem. Er bestaat een sterk gevoel dat gelijke kansen en democratie ook van toepassing moeten zijn op onderwijs in de moedertaal, het handhaven van godsdienst en levensbeschouwing en de daarmee samenhangende waarden en dat deze rechten bij de wet moeten worden beschermd (vgl. Steyn & Vanderstraeten, 1998; Claassen, 1995) overigens zonder dat één van de beginsels van het nieuwe onderwijsbestel, bijv. desegregatie verworpen wordt. De auteurs binnen deze stroming kunnen nog steeds als conservatief bestempeld worden. Ten eerste word hun werk gekenmerkt door een overwegend behoudend karakter wat betref de rol van opvoeding en onderwijs (het bewaren en overdragen van traditionele cultuur, taal, godsdienst, levensbeschouwingen en algemene waarden). Ten tweede dienen ze als conservatief bestempeld te worden aangezien hun opvattingen zich sterk onderscheiden van de twee andere stromingen die onderwijs vooral beschouwen als een middel tot maatschappelijke verandering en algemene vooruitgang.

2.2.2 Het liberale paradigma: constructieve betrokkenheid door trouw te blijven aan de oude principes

Vanuit de hoek van het liberale paradigma is na 1994 eveneens een constante stroom literatuur verschenen, waarvan hier enkele voorbeelden: D.L. Brook (1996) From exclusion to inclusion: Racial Politics and South Africa's education reform, *Anthropology and Education Quarterly*, 27, 204-231; M.J. Kahn (1996) Five Years gone: A case study of education policy implementation in the transition to democracy in South Africa, *International Journal of Educational Development*, 16, 281-289; C.G. Mackenzie (1994) Black students at 'white' universities: the character and provision of liberal higher education institutions in post-apartheid South Africa, *Compare*, 24, 67-78; D. Johnson (1995) The challenges of educational reconstruction and transformation in South Africa, *Comparative Education*, 31, 131-140, en K. Dovey (1996) Becoming a 'Learning nation': What are South Africa's chances?, *International Journal of Lifelong Education*, 15, 353-369.

De auteurs die werken vanuit het liberale paradigma bestuderen evalueren het nieuwe onderwijsbestel voortdurend tegen de achtergrond van de liberale waarden (zoals uiteengezet in 1.2.2). Ze aanvaarden de principes waarop het nieuwe bestel is gebouwd (democratisering, multiculturalisme, gelijkberechtiging, desegregatie) – die vormen immers deel van de liberale waarden. Bovengenoemde publicaties worden gekenmerkt door een sterke optimistische ondertoon. Het wordt nog steeds zeer belangrijk gevonden dat het onderwijs word ingeschakeld als een instrument om het land economisch te ontwikkelen (vgl. Kahn, 1995; Lewin, 1995). Er wordt vooral aandacht besteed aan struikelblokken op de weg naar gelijkberechtiging en *de facto* integratie (vgl. Brook, 1996; Johnson, 1996; Adler, 1995) en aan het bedenken van maatregelen om die hindernissen uit de weg te ruimen (vgl. Mackenzie, 1994; Kahn, 1996; Adler, 1995).

2.2.3 Het voormalige neo-marxistische conflict – paradigma: de scheiding tussen afwijzers en opbouwers

In de kring van de aanhangers van dit paradigma kan men na 1994 twee denkrichtingen

onderscheiden: die van de opbouwers en die van de afwijzers. Aan de ene kant staan diegenen die de huidige politieke- en onderwijsontwikkelingen aanvaarden. Ze zijn van mening dat de vooruitgang die is geboekt voldoende is om hun idealen (desegregatie, multiculturalisme, gelijkberechtiging en democratisering) te verwezenlijken en zodoende gelijke onderwijskansen voor iedereen te scheppen. Deze groep lijkt in de meerderheid te zijn; de belangrijkste publicaties vanuit deze stroming zijn de volgende: P. Kallaway *et al.* (eds) *Education after apartheid: South African education in transition* (Cape Town: University of Cape Town Press); C Hopfer (1997) Empowering adult education in Namibia and South Africa, *International Review of Education*, 43, 43-59; M. Kgobe (1997) The National Qualification Framework in South Africa and 'out of school youth': problems and possibilities, *International Review of Education*, 43, 317-330; en A Fataar (1997) Access to schooling in a post-apartheid South Africa: Linking concepts to contexts, *International Review of Education*, 43, 331-348.

Daarnaast zijn er ook theoretici die de nieuwe onderwijsbedeling afwijzen. Hun kritiek varieert. Sommigen beweren dat het nieuwe onderwijssysteem in die mate op economische groei en de eisen van de economie is afgestemd dat het onmogelijk in de behoeftes van de bevolking kan voorzien (vgl. Greenstein, 1995), terwijl anderen menen dat multiculturalisme in het onderwijs het ontwikkelen van Afrocentrisch onderwijs tegenstaat (vgl. Asumah & Hlatshwayo, 1995). Vervolgens is er nog de mening van Enslin (1993/4) die vanuit een feministisch perspectief beweert dat het huidige politieke- en onderwijsstelsel seksistisch is en daarom impliciet het bereiken van de idealen van gelijkberechtiging blokkeert.

3 Synthese: eensgezindheid vanuit verscheidenheid na 1994. De lacune van een vergelijkend perspectief

Vanaf het moment dat de maatschappelijke wederopbouw en de onderwijshervormingen van start zijn gegaan, d.w.z. in 1994, bestaan er

drie verschillende paradigma's in de sociaal-wetenschappelijke literatuur over onderwijs in Zuid-Afrika. Waar de tegenstand vóór 1994 echter duidelijk het resultaat was van verzuiling zonder dialoog, zoals reeds besproken in 1.3, is er na 1994 brede eenstemmigheid over de fundamenteën en de idealen van de nieuwe onderwijsbedeling. Er wordt nu vanuit drie verschillende invalshoeken op hetzelfde studie-object gefocust. Vanuit drie verschillende perspectieven wordt nu met eendracht gewerkt aan hetzelfde project. Deze uitgangspositie is uiteraard meer geëigend voor een zinvolle interactie tussen onderwijspraktijk en -theorie dan de situatie van vóór 1994. Het probleem is dat de overgrote meerderheid van de literatuur over onderwijs in Zuid-Afrika (de research van vóór 1994) ofwel het toenmalige onderwijssysteem probeerde te rechtvaardigen óf geheel in beslag werd genomen door kritiek op dit systeem (zonder ooit een alternatief te formuleren). Nu dit systeem niet meer bestaat, is deze literatuur van geen enkele waarde voor de opbouw en verbetering van het huidige onderwijssysteem. Daarom zijn vergelijkende studies met onderwijssystemen in andere landen waar al langer ervaring bestaat met het nastreven van de idealen die Zuid-Afrika zichzelf op onderwijsgebied tot doel heeft gesteld (democratisering, multiculturalisme, desegregatie, gelijkberechtiging) van grote waarde voor het opzetten van een raamwerk van wetenschappelijke kennis en literatuur als gids voor de onderwijshervormingen in Zuid-Afrika. Bijvoorbeeld: Het nastreven van gelijkberechtiging via het opzetten van 'Comprehensive schools' in Westeuropa in de jaren '60 en '70 is van onschatbare waarde voor de manier waarop Zuid-Afrika denkt haar onderwijs te kunnen hervormen.

Literatuur

- Adler, J. (1995) Dilemmas and Paradox – Secondary Mathematics Teachers' Knowledge of Their Teaching in Multi-lingual Classrooms. *Teaching and Teacher Education*, 7, 263-274.
- African National Congress. (1994)-1. *A Policy Framework for Education and Training*. (ANC Education Department)
- African National Congress. (1994)-2. *'n Beter Lewe*

- vir Almal: Span saam vir Werk, Vrede en Vryheid (Verkiessingsmanifes, 1994 Verkiessing) Marshalltown: ANC Department of Information and Publicity.
- Asumah, S.N. & Hlatshwayo, S. (1995). The politics of multi-cultural education in South Africa: Vogue, oxymoron or political paralysis. *Western Journal of Black Studies*, 19, 284-292.
- Barnard, S.S. (1984). *Inleiding tot die Vergelykende Opvoedkunde*. Durban: Butterworth.
- Behr, A.L. (1988). *Education In South Africa: 1652-1988: Origins, Issues and Trends*. Cape Town: Academica.
- Christie, P. (1991). *The Right To Learn: The Struggle for Education in South Africa*. (second edition). Braamfontein: Ravan Press.
- Claassen, J.C. (1995). The education system of South Africa. In Dekker, E. & Schalkwyk, O.J., van (eds), *Modern Education Systems* (second edition) (pp. 447-455). Durban: Butterworths.
- Cross, M. (1986). A Historical Review of education in South Africa: towards an assessment. *Comparative Education*, 22, 185-200.
- Department of Education. (1995) *Curriculum 2005. Lifelong Learning for the Twenty-First Century*. Pretoria: Department of Education.
- Department of Education (1997). *Outcomes-Based Education in South Africa: Background Information for Educators*. Pretoria: Department of Education.
- Department of Education (1998). *Green Paper on Further Education and Training: Preparing for the Twenty-First Century through Education, Training and Work*. Pretoria: Department of Education.
- Enslin, P. (1993/4). Education for Nation-Building: A Feminist Critique. *Perspectives in Education*, 15, 13-26.
- Fataar, A. (1997) Access to Schooling in a Post-apartheid South Africa: linking concepts to context. *International Review of Education*, 43, 331-349.
- Greenstein, R (1995) Education Policy Discourse and the New South Africa. *Perspectives in Education*, 16, 193-204.
- Hartshorne, K.B. (1992). *Crisis and Challenge: Black education: 1910-1990*. Cape Town: Oxford University Press.
- Harricombe, L.J. & Lancaster, F.W. (1995). *Out of the cold: academic boycotts and the isolation of South Africa*. Arlington: Information Resources Press.
- Horrell, M (1968). *Bantu education to 1968*. Johannesburg: South African Institute of Race Relations.
- Jaarsveld, F.A. van (1971). *100 Basiese dokumente by die studie van die Suid-Afrikaanse geskiedenis 1848-1961*. Kaapstad: NASOU.
- Johnson, D. (1995). The challenges of educational reconstruction and transformation in South Africa. *Comparative Education*, 31, 131-140.
- Kahn, M.J. (1996). Five years gone: A case study of education policy implementation in the transition to democracy in South Africa. *International Journal of Educational Development*, 16, 281-289.
- Karis, T. & Gerhart, G.M. (1977). *From Protest to Challenge: A Documentary history of African politics in South Africa 1882-1964, volume 3: Challenge and Violence: 1953-1964*. Stanford University: Hoover Institution Press.
- King, E.J. (1979). An educational way ahead for South Africa? *International Review of Education*, 25, 478-499.
- Labuschagne, A.J.J., Kok, J.C. & Myburgh, C.P.H. (1997). Vryheid, Verdraagsaamheid en die regte van minderhede as opvoedingsaangeleentheid in 'n demokrasie. *Suid-Afrikaanse Tydskrif vir Opvoedkunde*, 17, 136-143.
- Lemmer, E.M. (1993). Educational Research in South Africa: Problems and prospects. *Compare*, 23, 53-62.
- Lewin, K.M. (1995). Development Policy and Science Education in South Africa: Reflexions of Post-Formalism and Praxis. *Comparative Education*, 31, 131-140.
- Liebenberg, B.F. Die relevansie van kultuurgebonde onderwys vir behoorlike volwassewording (dissertatie). Universiteit van Suid-Afrika, 1998.
- Mackenzie, C.G. (1994). Black students in 'white' universities: the character and provision of liberal higher-education institutions in post-apartheid South Africa. *Compare*, 24, 67-78.
- Moodie, C.G. (1994). The state and liberal universities in South Africa: 1948-1990. *Higher Education*, 27, 1-40.
- Mphahlele, M.C.J. & Mminele, S.P.P. (1997). *Education through the ages* (book 3). Pretoria: Kagiso.
- Naidoo, K. (1990). The politics of student resistance in the 1980's. In Nkomo, M. (ed.), *Pedagogy of Domination: Toward a democratic education in South Africa*. (pp. 121-146). Trenton, New Jersey: Africa World Press.
- Nasson, B. (1988). Bitter Harvest: Farm Schooling for Black South Africans. *Perspectives in Education*, 10, 13-41.

- Nasson, B. & Samual, J. (1990). *Education: From Poverty to Liberty*. Cape Town: David Philip.
- Pampallis, J. (1991). *Foundations of the New South Africa*. London: Zed Books.
- Raad vir Geesteswetenskaplike Navorsing. (1981). *Onderwysvoorsiening in die RSA: Verslag van die Hoofkomitee van die RGN ondersoek na onderwys*. Pretoria: Raad vir Geesteswetenskaplike Navorsing.
- Republic of South Africa. (1996). *South African Schools Act (Act number 84 of 1996)*. Pretoria: Government Printer.
- Republic of South Africa. (1997). *The Constitution of the Republic of South Africa, 1996. Annotated Version*. Wynberg: HN Communications
- Republiek van Suid-Afrika. (1995). *Wet op die Suid-Afrikaanse Kwalifikasie Owerheid (Wet 58 van 1995)*. Pretoria: Staatsdrukker.
- Schriewer, J.K. & Keiner, E. (1992). Communication Patterns and Intellectual Traditions in Educational Sciences. *Comparative Education Review*, 36, 25-51.
- Steyn, H.J. & Vanderstraeten, R. (1998). Minority group education in South Africa: A Comparative Perspective. *South African Journal of Education*, 18, 89-96.

Auteur

C. C. Wolhuter (1957) studeerde pedagogiek aan de Rand Afrikaans Universiteit en aan de universiteiten van Pretoria en Stellenbosch. Hij is momenteel werkzaam als 'senior lecturer' in vergelijkende pedagogiek aan de Universiteit van Zoeloeland. Zijn onderzoek richt zich voornamelijk op raakvlakken in onderwijs-tradities tussen verschillende landen en op vergelijkingen tussen het onderwijs in Latijns-Amerika en Zuid-Afrika

Adres: Universiteit van Zoeloeland, Departement Vergelijkende Pedagogiek, Privaatsak X1001, Kwa-Dlangezwa 3886, Zuid-Afrika

Abstract

The state of the social-scientific literature depicting and evaluating education in South Africa: From ineffective pillarisation to constructive synergetic convergence

C.C. Wolhuter. *Pedagogische Studiën*, 1999, 76, 361-370.

The aim of this research is to obtain clarification about the state of the social-scientific literature depicting and evaluating education in South Africa, especially in the wake of the comprehensive post-1994 societal reform. In the pre-1994 literature, three schools of thought are identifiable: a conservative pro-status quo paradigm, a liberal paradigm, and a neo-Marxist conflict paradigm. Whereas the education system of the time was for the conservative pro-status quo paradigm an inseparable part of a non-negotiable and non-compromiseable societal order, the neo-Marxists rejected the education system outright, and demanded a complete political, social, economic and educational reconstruction. Such a situation rendered impossible a critical interaction between science and education practice, – science in service of the improvement of education practice. Since the post-1994 educational and societal reforms, the three paradigms in the social-scientific literature persist, but there now is consensus as to the foundation stones and ideals of the new education system, and meaningful interaction between science and education practice is possible. Comparative studies with countries which have an experience of the educational goals South Africa are pursuing, probably are at present the most promising way for the provision of a scientific theory guiding South Africa on its road of educational reform.

Onderwijsresearchdagen 1999

Inleiding (P. Sleegers en L. Verhoeven, Katholieke Universiteit Nijmegen)

Op 20 en 21 mei van dit jaar werd de 26^e editie van de OnderwijsResearchDagen (ORD) in Nijmegen gehouden. De ORD werd dit jaar in goede samenwerking georganiseerd door de sectie Onderwijs en Educatie (voorheen vakgroep Onderwijskunde), het Universitair Instituut voor de LerarenOpleiding (UNILO), het Instituut voor Toegepaste Sociale Wetenschappen (ITS) en het Instituut voor Onderwijskundige Dienstverlening (IOWO) van de Katholieke Universiteit Nijmegen. De deelnemers, ongeveer vierhonderd, hadden de keuze uit de traditionele en vertrouwde activiteiten: symposia, papersessies, fora en ronde tafelssessies. Daarnaast waren er vier buitenlandse onderzoekers (Walford, Smylie, Mezirow en Goldstein) als keynote speakers uitgenodigd, die aangevuld met twee Nederlandse onderzoekers (Rijsman en Kwakman) het aantal gastsprekers op een totaal van zes bracht. Verder was het programma gestructureerd volgens de bekende VOR-divisie-indeling. Het programma bevatte ook nieuwe elementen. Zo duurde het congres slechts twee volle dagen in plaats de gebruikelijke drie. In het programmaboek schrijft de voorzitter van de ORD '99 Simons daarover dat deze beslissing is genomen vanuit de ervaring dat de driedaagse ORD-programma's in de praktijk maar uit twee effectieve dagen bestonden: de eerste ochtend en de laatste middag werden niet of nauwelijks gebruikt.

Ook nieuw was dat in tegenstelling tot de voorafgaande jaren de ORD niet werd geopend met een plenaire zitting op het de eerste dag, gevolgd door een diner. In Nijmegen begon het congres met twee Nederlandse gastsprekers op de ochtend, gevolgd door presentaties met aansluitend de plenaire opening gekoppeld aan een inhoudelijke sessie met twee internationale gastsprekers halverwege de middag. De bedoeling van deze wijziging was om meer mensen

te betrekken bij de plenaire opening. Nieuw was ook de divisieoverstijgende themabijeenkomsten (symposia, papersessies en fora). Deze themabijeenkomsten waren parallel geprogrammeerd op de tweede dag (vrijdag 21 mei) van het congres. In deze divisie-overstijgende bijeenkomsten kwamen thema's aan de orde die vanuit verschillende divisies bestuurd worden. Voorbeelden van divisie-overstijgend activiteiten en thema's waren een forum over de pedagogische opdracht van de school (met bijdragen van Klaassen, Veuglers, Vedder en Wesseling) en een symposium over zelfstandig leren in het voortgezet onderwijs en implementatie van het studiehuis (met bijdragen van Alink, de Bruijn, Brand-Gruwel, Schellings en Vermetten, van der Waal en Vrijman-Putten). Deze themabijeenkomsten waren succesvol en verdienen het om een vervolg te krijgen bij de volgende ORD in Leiden.

Tot slot had de Nijmeegse organisatiecommissie een centraal thema bedacht: de ORD in Nijmegen stond in het teken van samenwerkend leren. Dit thema staat momenteel zowel nationaal als internationaal weer sterk in de belangstelling en staat bovendien in verschillende Nijmeegse onderzoeksprojecten centraal. Dit Nijmeegse thema was ook duidelijk zichtbaar afgedrukt op de tasjes die de deelnemers ontvingen bij de inschrijving. Op deze manier was op de Nijmeegse universiteitscampus de Nederlandse onderwijsonderzoeker te herkennen.

Zoals gezegd opende de ORD met twee keynote speeches die parallel in een collegezaal waren geprogrammeerd. Het gastcollege van Rijsman had als titel 'Sociaal Constructionisme en Onderwijs: Theorie en illustraties'. Zijn betoog bevatte een uiteenzetting van principes van het sociaal constructionisme, waarin kennis gezien wordt als de ogenschijnlijk individuele reproductie van de sociaal gecoördineerde interacties tussen mensen, die als betekenissen in taal en cultuur zijn opgeslagen. Dit sociaal constructionisme werd afgezet tegen de klassieke cognitieve psychologie waarin kennis opgevat wordt als het product van de individuele verwerking van informatie en in het verlengde daarvan communicatie gezien wordt als een uitwisseling van kennis tussen mensen. De consequenties van deze modellen voor het onderwijs werden door Rijsman in hoofdlijnen geschetst.

Kwakman ging in haar gastlezing in op de professionele docent die n haar ogen tegelijkertijd een lerende docent is. In feite rapporteerde zij over de bevindingen van haar recente promotie-onderzoek naar het leren van docenten. Daarbij kwam het begrip professionaliteit aan de orde en maakte Kwakman duidelijk dat dit begrip sterk samenhangt met het leren van docenten: om professioneel te zijn en te blijven, moeten docenten zelf leren. Een belangrijke conclusie van haar onderzoek was dat leren op de werkplek niet automatisch plaats vindt en dat scholen leren van docenten moeten ondersteunen en stimuleren. Kwakman eindigde haar lezing met een collage van kopteksten van artikelen die in landelijke dagbladen en vaktijdschriften waren verschenen naar aanleiding van haar proefschrift. Het beeld dat daarbij naar voren kwam was dat van een onderzoeker die 'zo nodig iets moet zeggen over wat docenten moeten doen, zonder dat ze zelf voor de klas heeft gestaan of op een school heeft gewerkt'. Waar bemoeit ze zich mee. De kloof tussen wetenschap en de praktijk is er de afgelopen jaren blijkbaar niet kleiner op geworden. Na deze twee openingssessies begonnen de diverse presentaties van onderzoek in de vorm van symposia, paperpresentaties, fora en rondetafel-discussies.

De officiële opening van de ORD was, zoals gezegd, dit jaar op het einde van de middag van de eerste dag (donderdag 20 mei) geplaatst. Deze opening werd gevolgd door een inhoudelijke programma waarbij Walford en Smylie hun keynote speeches zouden geven. Na het welkomstwoordje van de voorzitter van organisatiecommissie Simons en namens de KUN door de heer Peters, lid van het College van Bestuur, slaagde Verloop met zijn afscheidsrede als voorzitter van de VOR erin de aandacht van de deelnemers vast te houden. Hij stond eerst stil bij de inner circle van vooraanstaande en invloedrijke onderwijsonderzoekers die elkaar steeds allerlei klussen toeschuiven, waardoor de werkbelasting van diezelfde kleine groep onderzoekers alleen maar toeneemt. In het verlengde hiervan hield hij een pleidooi om deze vicieuze cirkel te doorbreken door het scheppen van mogelijkheden voor jonge veelbelovende onderzoekers om meer dan nu het geval is te participeren aan deze gemeenschap van onderwijsonderzoekers. In heldere en niet

mis te verstane bewoordingen ging hij vervolgens in op het gebrek aan collegialiteit van een aantal niet nader genoemde collega-onderzoekers, die hij betitelde als 'parasieten'. Het betreffen onderzoekers die uitsluitend gericht zijn op het schrijven van artikelen en het binnenhalen van onderzoeksgelden. Andere activiteiten waarvoor ze gevraagd worden en die ook onderdeel zijn de wetenschappelijke beroepsuitoefening, zoals het beoordelen van artikelen of projectvoorstellen, vergaderen en het bijwonen van allerlei bijeenkomsten, het opknappen van een klusje voor een collega-onderzoeker worden door deze 'parasieten' consequent geweigerd. Alles draait bij deze onderzoekers om het schrijven van artikelen en het onderhouden van interessante buitenlandse contacten. In het verlengde van deze felle kritiek deed Verloop een beroep op het vergroten van de samenwerking tussen onderwijsonderzoekers, ook op terreinen die te maken hebben met de randvoorwaardelijke aspecten van onderwijsonderzoek. De toehoorders dachten na en keken soms wat onzeker om zich heen.

Na deze persoonlijke reflectie van Verloop, was er ruimte in het openingsprogramma gemaakt voor de presentatie van de VOR-website (www.ou.nl/open/VOR) door Ellen Rusman. Hoewel de start niet soepel verliep, bleef zij rustig en vertelde ze kort iets over het experiment, dat met hulp van de Open Universiteit Nederland mogelijk is gemaakt. De bedoeling is dat de website ook actief wordt gebruikt, hetgeen helaas op dat moment nog niet het geval was, maar hopelijk wel in de toekomst het geval zal zijn.

Vervolgens werden de VOR-prijzen uitgedeeld. Als eerste werd het VOR-erelidmaatschap uitgereikt aan prof.dr. W. Wijnen, voor zijn bijdrage aan en inzet voor het onderwijsonderzoek en de onderwijsvernieuwing in Nederland. Hierdoor is het aantal ereleden gegroeid tot vier (de Groot, Knoers, Warries en Wijnen). Daarna werd de VOR-dissertatieprijs 1998 uitgereikt door de voorzitter van de jury Meijnen. In totaal waren er negen dissertaties ingestuurd, die door de jury, die naast Mijnen bestond uit Verloop, Snijders en Simons, beoordeeld waren met behulp van de volgende criteria: theoretische inbedding, relevantie voor de praktijk, wetenschappelijke ambachtelijkheid, originaliteit en creativiteit. De VOR-prijs

ging dit jaar naar dr. Edith Hooge voor haar proefschrift getiteld 'Ruimte voor beleid. Autonomievergroting en beleidsuitvoering door basisscholen'.

Na deze openingsceremonie waren Walford en Smylie aan de beurt. Walford ging in zijn keynote in op de relatie tussen 'choice', 'competition', 'cohesion' en 'equality'. Centraal in zijn lezing stond de vraag naar de relatie tussen de toegenomen competitie tussen scholen en de (potentiële) effecten daarvan op de sociale cohesie, kwaliteit en gelijkheid van het onderwijs. Daarbij ging hij vooral uit van de Engelse situatie, maar maakte hij af en toe ook een vergelijking met de Nederlandse situatie. Zo gebruikte hij bijvoorbeeld het verschijnsel van vrijheid van schoolkeuze zoals we dat in Nederland kennen om aannemelijk te maken dat de verklaring die in Engeland wordt geformuleerd dat toename van competitie tussen scholen onder andere veroorzaakt wordt door een toegenomen keuzevrijheid, te simpel is.

Als laatste spreker van het middagprogramma ging Smylie in op het leren van docenten op de werkplek. In zijn presentatie ging hij in op de betekenis van sociaal kapitaal voor de ontwikkeling van menselijk kapitaal binnen de context van onderwijsvernieuwing. Met behulp van resultaten van onderzoek liet hij zien dat wederzijds vertrouwen, sociale steun, het ontwikkelen van collectieve mechanismen voor collegiale controle, samenwerking en het creëren van een professionele gemeenschap die gekenmerkt wordt door continue interactie, het leren van docenten op de werkplek en hun functioneren positief kunnen bevorderen. Aan het einde van zijn bijdrage benadrukte hij de cruciale rol die de schoolleider kan spelen voor de bevordering van de professionele ontwikkeling van docenten.

De lezingen van beide buitenlandse gast-sprekers waren interessant, maar kwamen door de wijze van programmering helaas niet helemaal tot hun recht. Ze waren namelijk op het einde van een twee uur durende plenaire sessie gepland. Hiermee een te groot beroep op de aandacht en concentratie van een groot aantal deelnemers gevraagd, waardoor de zaal steeds meer leegstroomde. Iets om bij de volgende ORD te Leiden rekening mee te houden.

Na deze algemene impressies, zijn hieronder de indrukken opgetekend van divisie-

specifieke thema's. Achtereenvolgens komen aan de orde Curriculum, Leren en Instructie, Onderwijs en Samenleving, Methodologie en Evaluatie, Lerarenopleiding en Leraarsgedrag, Hoger Onderwijs, Beroeps- en Beleidsopleidingen en Volwasseneneducatie, Beleid en Organisatie in het onderwijs en het Nijmegenese thema Coöperatief leren. Achter elk thema staan de namen van de themacoördinatoren vermeld, waarbij de eerstgenoemde persoon de samenvattende tekst heeft verzorgd.

Curriculum (C. Aarnoutse, Katholieke Universiteit Nijmegen, N. Nieveen, Universiteit Twente en R. Soetaert, Rijksuniversiteit Gent)

In het thema *Curriculum* was gelukkig weer leven in de brouwerij. Er werden maar liefst vier symposia, drie fora en een papersessie gehouden, die voor het merendeel ook echt over het curriculum of curriculumontwikkeling gingen.

In het eerste symposium dat georganiseerd was door het Centrum voor Didactiek van Wetkunde en Natuurwetenschappen uit Utrecht, voorzitter Van der Valk, stond het project *Beta Profielen in het Studiehuis* centraal. Dit project stelt zich tot doel om in samenwerking met scholen voor het voortgezet onderwijs een studiehuis-didactiek voor de bèta-vakken te ontwikkelen. Uit het onderzoek kwamen enkele interessante uitkomsten naar voren. Zo bleek dat het voorlopig beter is om lessen genetica en kansberekening op elkaar af te stemmen dan te integreren, dat de probleemstellende werkwijze mogelijkheden biedt om het leren van basale natuurwetenschappelijke begrippen te combineren met het ontwikkelen van zelfstandig leren, en dat docenten hun bereidheid tot samenwerken ook richten op het vormgeven aan de eigen rol en de zelfstandigheid van leerlingen.

Het tweede symposium had betrekking op het ontwikkelen en onderzoeken van multimedia cases op het gebied van *Natuur en Techniek* ten behoeve van de professionele ontwikkeling van Pabo-studenten. Dit symposium dat georganiseerd was door Van den Akker en Van den Berg van de universiteit Twente paste goed in het thema Curriculum. Uit de formatieve evaluatie van de multimedia casus 'Vloeistoffen in reageerbuisen' bleek dat Pabo-studenten

duidelijke ondersteuning nodig hebben om te kunnen profiteren van informatie in een multimedia casus. De multimedia casus 'Dat vogelen we samen wel even uit' op het gebied van natuuronderwijs en het curriculuminformatie systeem 'Op zoek in de natuur' vonden bij Pabo-docenten en voor een deel ook bij Pabo-studenten een goed onthaal.

Het derde symposium 'Een tussenbalans van didactisch handelen in de basisvorming' werd gepresenteerd door Roelofs en Houtveen van de universiteit Utrecht en Zwarts en Sips van de inspectie. Uit het uitgevoerde onderzoek bleek dat de basisvorming vrij sterk gericht is op de ontwikkeling van basiskennis en -vaardigheden en dat het beoogde didactische concept van 'toepassing, vaardigheden en samenhang' in de praktijk nauwelijks uit de verf komen. Inmiddels weten we dat het ideaal van de basisvorming nog verder weg ligt dan tijdens het symposium naar voren kwam.

Het vierde symposium *Het curriculum als symposium* stond onder leiding van Soetaert van de lerarenopleiding van de universiteit van Gent. Binnen deze opleiding wordt een opleidingsonderdeel ontwikkeld waarin een wetenschapsfilosofie wordt gepresenteerd voor leraren. Hierbij concentreert men zich op een aantal rollen van leraren: de leraar als onderzoeker, als designer en als postmoderne intellectueel. Interessant was dat tijdens het symposium ontwikkeling en onderzoek in de vorm van een dialoog werden gepresenteerd.

In de fora pleitten Van den Akker en Nieveen niet zonder flair voor een meer gericht gebruik van de beschikbare curriculaire denkaders en benaderingen in het debat over onderwijsvernieuwingen. Boersma verdedigde het standpunt dat een smal curriculum dat zich richt op het verwerven van breed toepasbare kennis en vaardigheden de voorkeur heeft boven een breed curriculum. Terwel beklemtoonde dat het constructivisme eerder een uitdaging is voor ontwerpers en onderzoekers dan een theorie van waaruit het curriculum deductief kan worden afgeleid.

In de papersessie werden vier onderzoeken gepresenteerd. Reezigt deed verslag van haar onderzoek naar de implementatie van twee educatieve programma's voor peuters en kleuters, namelijk Kaleidoscoop en Piramide. Het bleek dat veel leidsters en leerkrachten nog

moeite hebben met onder andere het stimuleren van talige interacties tussen kinderen onderling. Broer, Aarnoutse en Kieviet lieten zien dat kinderen van groep 8 met behulp van hun experimenteel programma heel goed in staat zijn om schema's te maken van informatieve teksten met een bepaalde tekststructuur. Uit het onderzoek van Smeets bleek dat de belangrijkste reden van docenten om geen gebruik van ICT te maken gezocht moet worden in het onbekend zijn met de mogelijkheden van ICT of in het zich niet zeker voelen ten aanzien van computergebruik. Voogt en Odenthal verzamelden informatie over het gebruik van ICT in 20 onderwijsleersituaties in verschillende landen en vertelden over de verschillende redenen die in het basisonderwijs, de basisvorming en de lerarenopleiding worden gehanteerd om ICT in te zetten.

Leren en instructie (L.Verhoeven, Katholieke Universiteit Nijmegen, J. Pieters, Universiteit Twente en L. Verschaffel, Katholieke Universiteit Leuven)

Het thema 'Leren en instructie' viel dit jaar uiteen in een tweetal symposia en een tweetal papersessies. Het eerste symposium had als onderwerp *inductief leren*. Veenman (RUL) trad daarbij als voorzitter op met Van Merriënboer (OU) als discussiant. In het symposium stond het door NWO in 1996 gestarte aandachtsgebied 'Inductief Leren' centraal. Inductief leren is daarbij gedefinieerd als het verwerven van kennis door het afleiden van principes en regels uit informatie in de (leer-)omgeving. Doel van dit aandachtsgebied is het in kaart brengen van initiële kennisverwervingsprocessen en de factoren die deze processen beïnvloeden. Tijdens het symposium is vanuit elk van de vier deelprojecten van dit aandachtsgebied verslag gedaan. Deelproject I (de Jong & Hulshof, UT) legt de nadruk op de rol van eventueel beschikbare voorkennis in het proces van inductief leren. In deelproject II (Veenman & Prins, RUL) wordt de relatie bestudeerd tussen intelligentie en metacognitieve vaardigheid tijdens inductief leren, en hoe deze relatie verandert onder invloed van toenemende expertise. Deelproject III (Beishuizen, Brinkman & Wilhelm, RUL) onderzoekt de inductieve vaardigheden van lerenden met verschillende leeftij-

den. In deelproject IV tenslotte, wordt inductief leren gemodelleerd in de ACT-R architectuur, waarbij de empirische data van de drie eerder genoemde projecten als input worden gebruikt. Voor de op empirische leest geschoeide deelprojecten werden de resultaten van een eerste serie experimenten gepresenteerd, terwijl vanuit deelproject IV de modellering van een eerste, eenvoudige inductieve taak is belicht. In de presentaties en de discussie kwam de onderlinge samenhang tussen de deelprojecten expliciet aan bod, onder meer bij het gebruik van gezamenlijke leeromgevingen, procesmaten en leermaten. De voorlopige resultaten van het programma wijzen op het belang van het gedetailleerd in kaart brengen van de relatie tussen voorkennis en leerprocessen, van de vaardigheden die bij uiteenlopende niveaus van expertise een faciliterende rol spelen, en van de contexten die transfer van inductieve vaardigheden kunnen bevorderen.

Het tweede symposium handelde over *ICT als gereedschap in realistisch wiskundeonderwijs* met Gravemeijer (RUU) als voorzitter. Binnen het kader van realistisch reken-wiskundeonderwijs wordt beoogd ICT in te zetten als middel om de eigen inbreng en construerende activiteit van de leerlingen te ondersteunen. Daarbij wordt gedacht aan toepassingen, zoals het gebruik van ICT als gereedschap bij het oplossen van wiskundige problemen, het gebruik van simulatieprogramma's die wiskundig onderzoek mogelijk maken en de inzet van ICT om informatie uit te wisselen. Dergelijk gereedschap kan worden ingezet uit efficiëntie-overwegingen. Het kan ook worden ingezet om taken die zonder hulpmiddel te lastig of zelfs onuitvoerbaar zijn binnen de mogelijkheden van de leerlingen te brengen. In het ene geval kan het gaan om complexe bewerkingen met weerbarstige getallen, in het andere geval kan het gaan om wiskundige bewerkingen die de leerling nog niet machtig is. Tenslotte kan zulk gereedschap worden ingezet voor exploratieve opdrachten in probleemgeoriënteerd onderwijs. In dit symposium zijn drie onderzoeken besproken die zich op deze kwesties richten. Het eerste onderzoek (Drijvers, RUU) betreft een verkenning naar de vraag of, en hoe je computer-algebra-systemen in de bovenbouw van het voortgezet onderwijs zou kunnen introduceren. Centraal stond het gebruik van de

symbolische rekenmachine in het voortgezet rekenonderwijs. Leerlingen blijken deze in het algemeen zonder problemen te adapteren, al vraagt dat aanvankelijk tijd en gewenning en stelt het gebruik ervan specifieke eisen aan voorkennis en vaardigheden van de gebruiker, zoals inzicht in de syntax van de machine, vermogen om te navigeren door de structuur van de machine en vaardigheid om de uitvoer van de machine te interpreteren. Het tweede onderzoek (Van Galen & Kanselaar, RUU) richt zich op het bieden van structuur bij probleemgeoriënteerde computertaken. Enerzijds is de vraag aan de orde, waar de gulden middenweg tussen te veel en te weinig structuur ligt, anderzijds richt het onderzoek zich op de vraag, hoe je structuur kunt bieden. Daartoe is in het onderzoek een onderwijspakket rond 'Measurement in Motion' ontwikkeld voor de brugklas. Er waren drie onderzoeksronden waarin steeds twee versies van taken werden vergeleken: een 'minder gestructureerde' versie met twee of drie onderzoeksopdrachten per taak, en een 'meer gestructureerde' versie met series korte deelvragen. De eerste resultaten in het project leveren concrete aanwijzingen op voor het structureren van open opdrachten in een ICT-leeromgeving. Bij het derde onderzoek (Gravemeijer, RUU) gaat het om specifiek voor een statistiek-onderwijsprogramma ontwikkeld gereedschap, dat het heruitvinden van statistische inzichten en representaties mogelijk moet maken. Centraal staat hier de vraag hoe de symbolische representaties op de computer dit proces kunnen ondersteunen. In een samenwerkingsproject met Vanderbilt University is gekozen voor het ontwikkelen van specifieke software die een integraal onderdeel uitmaakt van een leergang voor leerlingen van de eerste twee leerjaren van het voortgezet onderwijs. Het resultaat van het tot dusver verrichte ontwikkelingsonderzoek bestaat uit een experimentele leergang statistiek, aangevuld met empirisch en argumentatief onderbouwde aanwijzingen voor verbetering. In deze onderbouwing staat de rol die de grafische computer-representaties spelen in het leerproces centraal.

De eerste papersessie ging over *Leerprocessen in het primair onderwijs*. De sessie opende met een paper van Vanobbergen en Verhaeghe (Universiteit Gent) over sociale competentie

bij kleuters in Vlaanderen. De centrale onderzoeksvraag was in hoeverre met verschillen in manier van werken in het kleuteronderwijs ook verschillen in verwachte sociale competentie geassocieerd zijn. Eveneens is nagegaan in hoeverre er verschillen zijn in de wijze waarop en de mate waarin leer- en ondersteuningsmomenten in het werksysteem geïntegreerd worden. Tenslotte is ook de relatie tussen de moeilijkheidsgraad van de sociale participatiestructuur en de verwachte leerwinst voor de kleuters onderzocht. Om deze onderzoeksvragen te kunnen beantwoorden zijn in vier kleuterklassen observaties verricht en geanalyseerd. Een algemene ordening naar moeilijkheidsgraad qua sociale participatiestructuur van de vier kleuterklassen bleek echter onmogelijk. Een mogelijke verklaring voor deze vaststelling kan gezocht worden in het gegeven dat alle activiteiten ingebed liggen in een werksysteem dat steunt op een aantal onderliggende principes, die de sociale participatiestructuur van de ene activiteit vergemakkelijken en van de andere activiteit moeilijker maken.

In het tweede paper rapporteerden Aarnoutse en Manders (KUN) over een onderzoek naar de ontwikkeling van geletterdheid en gecijferdheid bij kinderen in de eerste drie leerjaren van de basisschool. In dit onderzoek stond de vraag centraal hoe zich bepaalde vaardigheden op het gebied van taal/ lezen en rekenen/ wiskunde in de onderscheiden jaargroepen ontwikkelen. Daartoe zijn instrumenten ontwikkeld voor afnames in kleine groepen van ongeveer vijf leerlingen en instrumenten voor individuele afnames. Voor de afnames in kleine groepen zijn in groep 1 en 2 voor de volgende vaardigheden instrumenten afgenomen: woordenschat, fonemisch bewustzijn, luisteren, seriëren en classificeren. Voor de individuele afnames zijn instrumenten afgenomen die betrekking hebben op: het onderscheiden van letterclusters, het benoemen van letters, het snel benoemen van letters, cijfers, plaatjes en kleuren, het fonologisch en fonemisch, begrippen kennen, tellen en getallen. Voor groep 3 zijn voor de volgende vaardigheden instrumenten ontwikkeld: woordenschat, luisteren, letters benoemen en het snel benoemen van letters, cijfers, plaatjes en kleuren, letterclusters onderscheiden, en fonemische analyse. De ontwikkelde instrumenten voldeden vrijwel allemaal aan de

normale psychometrische eisen. De betrouwbaarheid bleek over het algemeen vrij hoog. Per jaargroep zijn voor genoemde vaardigheden ontwikkelingscurves gepresenteerd.

In een derde presentatie beschreven Luwel, Verschaffel, de Corte en Onghena (Universiteit Leuven) een onderzoek onder tweede- en zesdeklassers naar het schatten van hoeveelheden in een tweedimensionale omgeving. Daarbij is de ontwikkeling van een adaptieve strategie voor het schatten van hoeveelheden in een tweedimensionale omgeving onderzocht. De taak bestond uit een serie trials waarbij telkens het aantal gekleurde blokjes in een rechthoekig rooster moet worden geschat. Deze taak laat minstens twee verschillende schatprocedures toe: een optelprocedure waarbij men herhaaldelijk (schattingen van) groepjes van gekleurde blokjes bij elkaar optelt, of een aftrekprocedure waarbij men het (geschatte) aantal lege vakjes aftrekt van het (geschatte) totaal aantal vakjes in het rooster (= het zgn. 'anker'). Een rationele taakanalyse maakt duidelijk dat de meest efficiënte schatstrategie bestaat uit het adaptief gebruik van beide procedures, rekening houdend met de verhouding tussen het aantal gekleurde blokjes en lege vakjes (= MIN strategie). Een andere mogelijke strategie is de stereotype ADD strategie waarbij men altijd de optelprocedure toepast, *ongeacht het aantal gekleurde blokjes*. In tegenstelling tot eerder onderzoek werd geen evidentie gevonden voor een geleidelijke ontwikkeling van de ADD naar de MIN strategie. Zowel tweede- als zesdeklassers gebruiken de MIN strategie met als grote verschil dat de meerderheid van de tweedeklassers het anker niet correct weet te bepalen. Dit doet vermoeden dat het correct kunnen toepassen van de MIN strategie bepaalde wiskundige vaardigheden vereist die nog niet bij alle tweedeklassers ontwikkeld zijn.

In het laatste paper van deze sessie deden De Bock, Verschaffel, Janssens en Rommelare (Universiteit Leuven) verslag van een onderzoek naar het optreden van de lineariteitsillusie. In deze studie is de wijze waarop een niet-proportionele opgave geformuleerd is, onderzocht door de factor 'opgavestructuur' experimenteel te manipuleren, terwijl de wiskundige kern van de opgave constant werd gehouden. Het onderzoek werd uitgevoerd onder 12-13- en 15-16-jarigen in twee Vlaamse

secundaire scholen. De studie bevestigde de enorm sterke impact van de lineariteitsillusie bij middelbare scholieren. Bovendien werd aangetoond dat deze tendens tot onterecht proportioneel redeneren mede kan verklaard worden door een opgavestructuur die door de leerlingen ten onrechte met proportionele opgaven wordt geassocieerd.

In de tweede papersessie vormden *Leren en instructie in het hoger onderwijs* het centrale thema. Hierin kwamen vier papers aan bod. Allereerst presenteerden Verheij, Veenman en Prins (RUL) een onderzoek naar het identificeren van risicostudenten aan een technische universiteit. In dit onderzoek stond de vraag centraal in hoeverre de eerder door Vermunt ontwikkelde Inventaris Leerstijlen (ILS) een valide en betrouwbaar instrument vormt voor het opsporen van risicostudenten. In een groot-schalige studie is het verband tussen scores op de ILS en studieresultaten nagegaan bij studenten aan de TU Delft. Daarbij bleek sprake van een marginaal verband. Als mogelijke reden voor het kleine verband wordt aangevoerd dat de ILS geen vaardigheden meet, maar kennis van studenten over studieactiviteiten of sociale wenselijkheid. Bovendien vraagt de ILS niet naar de kwaliteit van de door studenten verrichte studieactiviteiten maar naar de kwantiteit van de verrichte activiteiten. Het tweede paper van Elen en Clarebout (Universiteit Leuven) ging in op epistemologische opvattingen van leerlingen ten aanzien van het collaboratief leren in een ICT-leeromgeving. De doelgroep bestaat uit 144 leerlingen uit 5 Europese landen: België, Duitsland, Nederland, Spanje en Zweden. Al deze leerlingen volgen veelal 'gewoon' onderwijs gekenmerkt door een eclectisch gebruik van werkvormen. Gedurende een periode van 8 weken werkten deze leerlingen aan een open taak gerelateerd aan de werkzaamheden van het Europees Parlement in een probleem-georiënteerde, collaboratieve en hoog-technologische (telefoon, fax, e-mail, multimedia computer, internet en een specifiek ontworpen database) omgeving. Bij al deze leerlingen werd tweemaal met een tussentijdse pauze van minstens 10 weken een vragenlijst afgenomen waarin tevens 12 vragen werden opgenomen om epistemologische opvattingen van leerlingen vast te stellen. Uit een eerste exploratieve analyse blijkt een beperkt effect

van initiële epistemologische opvattingen op uiteindelijke leerresultaten en een geringe wijziging van initiële epistemologische opvattingen. Het derde paper van Limbach, Pieters en De Jong (UT) ging in op mogelijkheden om het instructieontwerpproces te ondersteunen. In een onderzoeksproject is een aanpak gevolgd waarbij de ondersteuning niet is gebaseerd op de volgorde van de activiteiten, maar op een type ondersteuning waarbij rekening wordt gehouden met een persoonlijke ontwerpbenadering van de ontwerper welke gebaseerd is op bijvoorbeeld ervaring. Geprobeerd is om inzicht te krijgen in het ontwerpproces van instructie voor simulatie-gebaseerde zelf-ontdekkende leeromgevingen. Er werd zowel gekeken naar het ontwerpproces van studenten als van docenten. Binnen het ontwerpproces is gekeken naar drie categorieën namelijk specifieke ontwerpactiviteiten en het gebruik van verschillende type kennis en operatoren. De verzamelde informatie uit de studie is mede gebruikt om de inhoud van een informatiesysteem te bepalen, welke de ontwerper tijdens het ontwerpproces kan raadplegen en bestaande uit o.a. tips, voorbeelden en achtergrondinformatie. Het vierde paper van Vos (UT) ging in op kennisstructuren bij experts-docenten. De centrale onderzoeksvraag is hoe de kennisstructuur van experts er uit ziet. Het doel van het onderzoek is te achterhalen hoe hun kennis van dat vak georganiseerd is. Acht docenten in het basisvak Netwerkanalyse in de Elektrotechniek hebben deelgenomen aan het onderzoek. Het onderzoek bestond uit een ongestructureerde sorteertaak van 350 kaartjes waarop voorbeelden van de belangrijkste kernbegrippen uit de Netwerkanalyse zijn weergegeven, in verschillende representaties. Het bleek dat docenten een zeer verschillende ordening aanbrachten. Het lijkt erop dat docenten niet een uniforme organisatie in hun hoofd hebben van de kennis uit het vak en dat de structurering van de kennis in dit vak tot stand komt op grond van ervaring.

Onderwijs en Samenleving (A. Wesselingh, Katholieke Universiteit Nijmegen, A. Dijkstra, Rijks Universiteit Groningen en Van Damme, Katholieke Universiteit Leuven).

De organisatoren van deze divisie kunnen

terugzien op een reeks geslaagde bijeenkomsten, die over het algemeen goed werden bezocht en een grote diversiteit van thema's hebben geboden, zoals van het brede thema Onderwijs en Samenleving kon worden verwacht. Er is geen aanleiding de formule zoals dit jaar toegepast grondig te wijzigen. De afwisseling van papersessies, symposia, forum en ronde tafel heeft redelijk gewerkt, al viel het kenmerkende verschil niet altijd goed te onderscheiden.

De divisie leverde op de eerste dag 's middags een van de key-note speakers, Dr. Geoffrey Walford, Reader in Education Policy aan de universiteit van Oxford, UK. Dr. Walford belichtte het Engelse onderwijssysteem vanuit het gezichtspunt van de toenemende marktwerking in het onderwijs en de gevolgen daarvan voor de keuze van scholen door ouders en leerlingen. De veelzeggende titel van zijn key-note address luidde: *'Choice, Competition, Cohesion and (E)quality in Education'*. Eerder die dag werd 's ochtends het symposium *'De relaties tussen schoolwelbevinden, prestaties en schoolloopbaan in het middelbaar onderwijs in Nederland en Vlaanderen'* gehouden. Een Vlaams-Nederlandse coproductie met papers van Van der Werf, Lubbers en Kuiper en Mertens en Van Damme. Aan de hand van een grootschalig longitudinaal Vlaams en Nederlands leerlingencohort werd stilgestaan bij de effecten van leerlingkenmerken als welbevinden en academisch zelfconcept op de prestaties in het voortgezet onderwijs. De discussie werd ingeleid door J. Roeleveld, die het gesprek onder meer bracht op het al dan niet bestaan van verschillen in de effecten die voor beide cohorten werden gerapporteerd.

Na de lunch volgde de papersessie *'Onderwijsachterstanden'* gepresenteerd in de vertrouwde vorm van losse papers over een traditioneel onderwijssociologisch onderwerp: achterstanden in het onderwijs, waarbij P. Jungbluth als referent optrad. Ook hier een Vlaams-Nederlandse samenwerking met interessante verschillen in aanpak van het achterstandenbeleid in Groningen (Knuver en Mulder), het taalbeleidsprogramma Deltaplan in Rotterdam (Westerbeek) en een Leuvens onderzoek naar de relatie tussen leerlingkenmerken en doelmatigheidsbeleving bij leerkrachten (Verhaeghe). Van de geplande paral-

lellessies op de ochtend van de tweede dag kwam het symposium *'Allochtone leerlingen in het basisonderwijs'* door ziekte te vervallen. De parallelle sessie had als thema *'De pro's en contra's van WSNS'*. Hier werden twee Amsterdamse papers gepresenteerd over verschillen in ontwikkeling van leerlingen in het basis- en speciaal onderwijs (Vergeer, Peetsma en Roeleveld) en de schoolloopbaan van risicoleerlingen (Jepma). Onder voorzitterschap van K. Koster werd levendig gediscussieerd. Het beleid rond Weer Samen naar School blijft discussie oproepen.

Na de middagpauze volgde een divisie-overstijgend thema *'De pedagogische opdracht van de school; aanknopingspunten voor onderzoek'*. Een thema dat leeft getuige de grote opkomst. Onder voorzitterschap van Mw. H. Dekkers presenteerden achtereenvolgens Roede (andere kwaliteiten van onderwijs), Klaassen (de pedagogische professionaliteit van docenten), Veugelers (docenten, ouders en leerlingen over de pedagogische opdracht) en Vedder (multiculturaliteit en pedagogische opdracht) hun standpunten, veelal uit onderzoek verkregen. Referent Wesselingh voegde het element burgerschap toe aan een levendige discussie. Duidelijk is dat hier ruime mogelijkheden liggen voor onderzoek, zij het dat het thema lijkt te lijden aan verschillende visies en definities en een terminologische uitzuivering gewenst is. Tegelijkertijd (de divisie heeft altijd veel aanbod en moet woeke- ren met de beschikbare tijd) vond rond de key-note speaker Geoffrey Walford een ronde tafel plaats onder voorzitterschap van F. van Wieringen over de marktwerking in het onderwijs onder de titel *'Marketization in Dutch Education: Help or Hinder?'*. De discussie werd ingeleid door twee Nederlandse bijdragen gebaseerd op recent afgesloten promotiestudies rond dit thema, waarin zowel de 'Risico's van de markt' (Teelken) als de 'Risico's van de overheid' (Waslander) aan de orde werden gesteld.

De reeks werd besloten met een papersessie rond het thema *'Schoolprestaties en toegevoegde waarde'* waarbij R. Bosker als referent optrad. Dijkstra en Veenstra (RUG) bespraken het effect van 'functional communities' op schoolprestaties, een kritische toetsing van het onderzoek van James Coleman in de VS. Dec-

kers en Dronkers (UvA) demonstreerden de stabiliteit van de toegevoegde waarde van basisscholen. Doolaard (UT) tenslotte liet zien dat gedurende een periode van acht jaar inderdaad stabiliteit optreedt in de werking van school- en klaseffecten op enkele afhankelijke variabelen.

Methodologie en Evaluatie (M. Voeten, Katholieke Universiteit Nijmegen, P. Sanders, Cito en P. Onghena, Katholieke Universiteit Leuven)

Voor het thema 'Methodologie en Evaluatie' was Harvey Goldstein van het Institute of Education in London uitgenodigd om een lezing te houden. Onder de titel *Modelling educational complexity* betoogde hij dat we voor het beschrijven van complexe onderwijsrealiteiten modellen en technieken nodig hebben die recht doen aan de complexiteit van de verschijnselen. De eenvoud van procedures of modellen vormt op zich geen rechtvaardiging voor het gebruik ervan. Factoranalyse en itemresponse modellen, waarvan het Rasch-model het eenvoudigste voorbeeld is, werden genoemd als voorbeelden van modellen die ondanks hun technische complexiteit te simpel zijn gebleken om recht te doen aan de complexiteit van de reële wereld. Dit werd geïllustreerd aan de hand van de International Adult Literacy Survey. Internationaal vergelijkende studies gaan gebukt onder het probleem of dezelfde set van toetsmaterialen wel gebruikt mag worden in verschillende culturele en maatschappelijke settings. Een ander terrein waar te simpele modellen kwaad kunnen aanrichten is het gebruik van 'league tables' voor het vergelijken van scholen op hun effectiviteit. Als voorbeelden van complexe modellen voor complexe systemen werden uiteraard de multiniveau-modellen genoemd, waar Goldstein en zijn medewerkers van het 'Multilevel Models Project' van het Institute of Education sinds het midden van de tachtiger jaren aan werken. Genoemd werden het 'value added'-model van schooleffectiviteit en de beperkingen daarvan voor het rangschikken van scholen naar effectiviteit. Het simpele hiërarchische multiniveau-model is inmiddels in diverse richtingen uitgebreid tot onder meer modellen met random kruisclassificaties (leerlingen die hun basis-

schoolperiode in een bepaalde school doorbrachten en voortgezet onderwijs genoten in een andere school) en meervoudige lidmaatschapsmodellen (leerlingen die van de ene naar de andere basisschool zijn verhuisd). Wanneer het multiniveau-model wordt toegepast op meta-analyse ontstaan nieuwe mogelijkheden, zoals de combinatie in één analyse van studies waarover alleen samenvattende statistieken beschikbaar zijn met studies waarvan de complete data gebruikt kunnen worden. Goldstein en zijn medewerkers pasten deze methode toe op studies naar de effecten van klassengrootte (inclusief het bekende STAR project). Experimenteel onderzoek naar klassengrootte diende als voorbeeld voor de stelling dat gebruik van randomisering tot foutieve causale conclusies kan leiden, als belangrijke effecten van de sociale context genegeerd worden.

In een andere sessie gaf Goldstein een demonstratie van de nieuwe versie van zijn computerprogramma MlwiN die eind 1999 uitkomt. In deze versie kunnen bijna alle commando's via het grafische Windows interface worden ingevoerd, hetgeen Roel Bosker (UT) ertoe bracht op te merken dat het doen van ingewikkelde multiniveau-analyses gevaarlijk eenvoudig wordt. De interpretatie van resultaten wordt gevaarlijk als de gebruiker niet de implicaties van de toegepaste modellen doorziet.

Hoe moeilijk het kan zijn een eenvoudig model te toetsen werd gedemonstreerd door Van Leeuwe (KUN), die verslag deed van pogingen om de 'simple view of reading' (begrijpend lezen is het product van decoderen en begrijpend luisteren) van Hoover en Gough te toetsen. Een van zijn conclusies was dat de predicties van Hoover en Gough onvoldoende onderscheidend zijn om de geldigheid van de 'simple view' te toetsen.

Veel aandacht ging uit naar problemen van meten en toetsen. Sanders (Cito en UvA) gaf een boeiend overzicht van aspecten van 'performance-based assessment' en de toepassing van generaliseerbaarheidstheorie. Uit zijn data bleek dat de interbeoordeelsovereenkomst geen probleem hoeft te zijn. Het grote probleem was de hoge mate van taakspecificiteit waardoor de generaliseerbaarheid van taakprestaties te laag bleek om van een betrouwbare meting te kunnen spreken. Sanders zocht de

oplossing in een nieuwe generaliseerbaarheids-theorie die niet langer uitgaat van één enkele trek, maar die een psychologische theorie omvat over intra- en interindividuele verschillen in oplossingsstrategie. Ook hier weer zien we het streven naar meer complexe modellen die beter de (psychologische) realiteit representeren. Sluijter (Cito) rapporteerde over een procedure voor het betrouwbaar, efficiënt en valide meten van tekenvaardigheid van leerlingen aan het einde van de basisschool. Per tekenopdracht werd een negenpuntsschaal geconstrueerd met geselecteerde kindertekeningen als ankerpunten.

Een belangrijke kwestie bij toetsing in het onderwijs is de vraag in welke mate de inhoud van een toets aansluit bij het door de leerkracht geïmplementeerde curriculum. Deze vraag naar toets-curriculum overlap (TCO) is al op diverse ORD's aan de orde geweest. Van Abswoude (Cito) deed verslag van de ontwikkeling van een instrument om TCO te meten in het kader van een project dat erop gericht is basisscholen in staat te stellen tot zelfevaluatie. Het instrument werd toegepast op een rekentoets uit het Cito leerlingvolgssysteem. Geconcludeerd werd dat een instrument voor TCO, naast een vraag naar de gebruikte lesmethoden, vragen moet bevatten naar de maakbaarheid van de toetsitems. Manders (KUN) rapporteerde over de ontwikkeling van een door de vakgroep Onderwijskunde van de KUN ontwikkeld leerlingvolgssysteem 'Leer in Zicht', gebaseerd op longitudinaal onderzoek waarin leerlingen zijn gevolgd van groep 3 tot en met 8 van het basisonderwijs. Het leerlingvolgssysteem bevat een computerprogramma dat ontwikkelingscurven voor individuele leerlingen schat op een aantal vaardigheden (technisch lezen, begrijpend lezen, woordenschat, spelling en technisch rekenen) en deze vergelijkt met het landelijk gemiddelde. Een symposium, verzorgd door medewerkers van Toegepaste Onderwijskunde van de Universiteit Twente onder leiding van Glas, was gewijd aan het oplossen van problemen bij het toepassen van computerondersteund adaptief toetsen.

Naast meting van leerprestaties en vaardigheidsmeting werd ook aandacht besteed aan het meten van welbevinden van leerlingen op school. Schepens en Deconinck (RU Gent) presenteerden een onderzoeksprogramma om een

instrument ten behoeve van de onderwijsinspectie te ontwikkelen voor leerlingen in het secundair onderwijs. Dit programma bestaat in de exploratiefase uit het ontwikkelen van een conceptueel kader mede op basis van literatuurstudie en het bevragen van leerlingen door middel van panelgesprekken om na te gaan welke beleavingsaspecten leerlingen zelf relevant achten en hoe zij deze verwoorden. Op basis van het gespreksmateriaal worden Likert-items voor de vragenlijst geconstrueerd. Helaas bood de ORD-opzet te weinig ruimte voor het discussiëren over de opzet van zo'n onderzoeksprogramma.

Tenslotte kwam de evaluatie van opleidingen in twee bijdragen aan bod. Een handelde over het constructivistisch maken van het evaluatieschema van Kirkpatrick, dat een algemeen aanvaarde standaard lijkt te zijn bij het evalueren van bedrijfsopleidingen (Schippers en Elen van de K.U. Leuven). Als het leren en opleiden vanuit constructivistische gedachten bekeken worden, dan moet dat met het evalueren ook maar. Of dat ook tot iets concreets kan leiden, bleef hier echter verborgen. De tweede bijdrage (Michels, KUN) behandelde een methode voor grootschalige evaluatie van bedrijfsopleidingen op werkgedragsniveau. Hierbij werd gebruik gemaakt van schriftelijke feedback naar de cursusdeelnemers om na te gaan of daarmee verdere opleidingsbehoeften bepaald kunnen worden.

Hoe realistisch zijn onze modellen en onderzoeksmethoden voor het beschrijven van de complexe onderwijswerkelijkheid? Is er nog een excuus voor de eenvoudige onderzoeker? Met die vragen zien wij de ORD het nieuwe millennium ingaan.

Lerarenopleiding en leraarsgedrag (P. Van de Ven, Katholieke Universiteit Nijmegen, J. Buitink, Rijks Universiteit Groningen en Libotton, Vrije Universiteit Brussel)

Het thema *Lerarenopleiding en leraarsgedrag* omvatte twee papersessies, drie symposia en één rondetafel-sessie. Tijdens de papersessie *Succesvol functioneren van leraar en opleiding* presenteerden Jager en Lokman onderzoek naar de rol van de docent in een leeromgeving met ICT. Tot op heden blijft de toepassing van ICT achter bij de verwachtingen. Uit literatuur-

studie en de interviews met docenten kan worden opgemaakt dat een implementatieproces niet alleen het gebruik van ICT moet bevatten, maar een vernieuwing van het totale onderwijsconcept. De lerarenopleidingen vervullen wat dat betreft op dit moment nog geen sleutelrol. Ze zijn evenzeer als de scholen bezig met een leerproces in dezen. De Jong en Van Leeuwen onderzochten of de samenstelling van de studentenpopulatie aan lerarenopleidingen in de afgelopen jaren is veranderd. De eindconclusie van het onderzoek is dat er weinig verschillen zijn naar objectieve kwaliteitskenmerken, maar dat de studenten een carrière in het onderwijs wel als minder veelbelovend beoordelen dan andere loopbanen. Het verschil tussen studenten in de onderwijssector en in de rest van het Hbo lijkt niet zo zeer te zitten in de objectieve kwaliteit (goede versus slechte studenten) maar meer in de gerichtheid. In een tweede paperpresentatie vroegen Senders en Vermeulen zich af of de praktijkschok voor lio's minder zou zijn indien zij een werknemersstatus verkrijgen. Er blijken verschillen tussen 'gewone' lio's en lio's als werknemer. De werknemers blijken een uitgebreider takenpakket te hebben met meer minder lesgebonden taken. De begeleiding wordt vaker gegeven door getrainde coaches, er worden met lio-'werknemers' meer functioneringsgesprekken gevoerd. In het derde paper beschreef De Jager een onderzoek naar de effecten van een training in instructie bij begrijpend lezen. Het betrof twee verschillende trainingen voor twee verschillende groepen: een in directe instructie en een training in *cognitive apprenticeship*. De trainingen hebben effect op het instructiegedrag van de leerkrachten. De kwaliteit van de instructie is toegenomen. Ook wordt meer tijd besteed aan het nabespreken, aan tekststructuren en metacognitieve vaardigheden. Er waren wel effect-nuances tussen beide trainingsgroepen.

In het vierde paper stelde Veugelers de vernieuwing in de tweede fase van het voortgezet onderwijs aan de orde. Deze veranderingen zijn beleidsmatig te zien als een voorbeeld van een ook internationaal geconstateerde samenwerking tussen het aansturen van bovenaf en het vergroten van de autonomie van docent en school. Het gaat om de verhouding tussen top-down stimulering en een bottom-up benadering. De vraag is dan wie doelen bepaalt, wie

de controle over doelen en processen heeft, hoe ver de autonomie van scholen/docenten/leerlingen gaat. Conclusie is dat het werk van de leraar er in het studiehuis gevarieerder op geworden is en dat de samenwerking tussen docenten is vergroot. Maar Veugelers constateert ook dat de bouwstenen van het studiehuis lijken te worden vernalen tussen de overladenheid van het programma en de nauwkeurigheid van de controle. De controle van de overheid naar de school, van de school naar de docent en van de docent naar de leerling vormt een belemmering voor de ontwikkeling van school, docent en leerling naar zelfstandigheid.

Tijdens het symposium *Wat vraagt leren onderzoeken en ontwerpen in de Tweede fase van docenten?*, presenteerden Dekkers en Van Rens onderzoek naar de bestaande lespraktijk en de achterliggende opvattingen van docenten ten aanzien van natuurwetenschappelijke practica binnen een Bèta-netwerk van vijf scholen. Docenten hebben vele bruikbare visies en ideeën om de lespraktijk zo aan te passen dat de nieuwe leerdoelen in de lespraktijk een plaats krijgen. Docenten zijn goed in staat activiteiten te ontwerpen en te begeleiden waarin leerlingen handelingen verrichten waarin ze vaardigheden oefenen. Maar het blijkt lastig voor docenten, leerlingen te laten expliciteren welke onderzoekshandelingen ze verrichten, wat daarvan de kwaliteit is en hoe die kwaliteit te verbeteren. In een tweede bijdrage presenteerde Van Tilburg een onderzoek naar de praktijkkennis van bèta-docenten in het kader van de ontwikkeling van een handleiding voor de begeleiding en de beoordeling van profielwerkstukken. Kennis en opvattingen van docenten over leren onderzoeken blijken bij de start van het project nog eclectisch en oppervlakkig. Het betreft dan ook een voor vele leraren nog niet eerder onderwezen onderwerp. Docenten geven aan hoofdzakelijk te leren van eigen ervaringen en van uitwisseling met collega's. Door gebrek aan ervaring is echter verspreiding van expertise binnen de beroepsgroep niet groot en is men aangewezen op expertise van buiten. Werkdruk enerzijds en het te weinig toegankelijk zijn van relevante literatuur bemoeilijkt daarenboven de ontwikkeling van leren onderzoeken. In een derde presentatie besprak Van der Waal verschillende visies op ontwerpen, en op de relatie tussen

ontwerpen en onderzoeken. Door de visie die ten grondslag ligt aan de geformuleerde exameneisen te analyseren en deze te koppelen aan de opvattingen die bèta-docenten blijken te hebben, worden suggesties voor opleiding en nascholing van docenten geïdentificeerd.

In het symposium *Een kennisbasis van docenten: verheldering en beschouwing van een onderzoeksprogramma* werden drie onderzoeken gepresenteerd. Meijer deed empirisch onderzoek naar praktijkkennis van ervaren talendocenten omtrent het onderwijzen van tekstbegrip in de bovenbouw van het VWO. Het onderzoek laat zien dat de praktijkkennis van docenten met betrekking tot het onderwijzen van tekstbegrip in de bovenbouw van het vwo overheerst wordt door variatie. De praktijkkennis van docenten blijkt deels individueel, deels groepsgebonden. Het onderzoek verschaft inzicht in de cognities die een rol spelen in de praktijk van het onderwijzen. Verder heeft het onderzoek informatie opgeleverd over instrumenten waarmee de praktijkkennis van ervaren docenten in kaart kan worden gebracht. Het tweede onderzoek van Van Driel richtte zich op praktijkkennis van ervaren bèta-docenten omtrent het onderwijzen van modellen en modelleervaardigheden. Uit interviews blijkt dat er een groep docenten is, die zich bij voorkeur richt op de inhoud van specifieke modellen, en daarbij vooral rekening houdt met onzekerheidsgevoelens en beperkte modelleervaardigheden van hun leerlingen. Daartegenover staat een groep die veel meer aandacht schenkt aan de aard en de ontwikkeling van natuurwetenschappelijke modellen. Zij hebben bovendien meer vertrouwen in de capaciteiten van hun leerlingen op dit gebied. Uit de analyses van vragenlijsten komt naar voren dat de twee groepen docenten met name verschillen in de mate waarin zij doceeractiviteiten toepassen waarin modellen centraal staan. In de derde bijdrage presenteerde Uhlenbeck de ontwikkeling van instrumenten voor de beoordeling van beginnende docenten. Uitgangspunt voor een dergelijke beoordeling moet de kennisbasis van de docent zijn. De beoordeling moet zoveel mogelijk plaats vinden in situaties die lijken op de werkelijke onderwijssituatie. De beoordeling moet zich ook richten op achterliggende overwegingen van docenten. Belangrijk is de overtuiging dat alleen door het gebruik van

meerdere beoordelingsinstrumenten tegemoet kan worden gekomen aan de complexiteit van het onderwijzen. Drie verschillende instrumenten werden ontwikkeld: een toets met gesloten vragen; simulaties van drie kenmerkende lessituaties; een portfolio. Deze instrumenten worden momenteel getest.

De papersessie *Leraar en lerarenopleiding: motieven en achtergronden* gaf vier papers te zien. Visser verschaftte in haar presentatie over 'de waardevolle leraar' antwoord op vragen aan leraren naar de motivatie voor en de verwachtingen van hun vak. Ze vroeg hen of die verwachtingen zijn uitgekomen en hoe die verwachtingen zich voor wat betreft de rol van de leraar verhouden tot de huidige ontwikkelingen in het onderwijs. Het belangrijkste motief voor het beroep van leraar is dat men het leuk vindt met kinderen te werken. Dit geldt sterker voor leraren basisonderwijs. Leraren in het voortgezet onderwijs zijn vooral leraar geworden omdat ze hun vak zo leuk vonden. Daarnaast komt uit het onderzoek naar voren dat alhoewel leraren vinden dat door de vele vernieuwingen in het onderwijs het werk zwaarder is geworden, ze wel positieve aspecten aan deze vernieuwingen onderscheiden. In het tweede paper presenteerden Jongmans, Biemans, De Jong en Sleegers een onderzoek naar de relatie tussen de participatie van docenten in het onderwijskundig beleid en het gebruik van actieve werkvormen door docenten. Het gebruik van actieve werkvormen door docenten wordt in de studie opgevat als een onderwijskundige vernieuwing. In de literatuur wordt participatie van docenten veelvuldig genoemd als een vorm van samenwerking die het implementeren van vernieuwingen in scholen kan bevorderen. Gebleken is dat er een relatie bestaat tussen de mate waarin docenten participeren in het onderwijskundig beleid van hun school en de mate waarin ze volgens hun leerlingen gebruik maken van actieve werkvormen: naarmate docenten in hogere mate participeren in het onderwijskundig beleid van hun school hanteren ze in de ogen van hun leerlingen meer werkvormen die zijn gericht op het stimuleren en activeren van de leerlingen. Geconcludeerd kan worden dat samenwerking tussen docenten op het niveau van de schoolorganisatie een positieve invloed heeft op de lespraktijk van docenten. In het derde paper

deed Janssens verslag van onderzoek naar hoe opleidingsinstituten denken over het opleiden van leraren. Een analyse van feitelijk aangeboden leerinhouden werd aangevuld met een interview met tenminste één docent uit elk opleidingsinstituut. Geconcludeerd kan worden dat opleidingsinstituten verschillen wat betreft de inhouden van het vakkencluster 'Opvoedkundige Wetenschappen'. Zo gaat in acht van de elf onderzochte instituten de mees- te aandacht naar de rubriek 'didactiek'. In twee instituten komt de rubriek 'psychologie' op de eerste plaats en in één instituut 'fundamentele pedagogiek'. In acht opleidingsinstituten wordt het onderliggend referentiekader gekenmerkt door een directe gerichtheid op het leerkrachtenberoep, terwijl in drie andere persoonlijke groei en ontwikkeling van de studenten op de voorgrond komt. Voorts is er een duidelijke tendens waarneembaar. Die betreft het instrumentele karakter van de inhouden waaraan extra aandacht wordt besteed of die door het docententeam zelf worden toegevoegd aan die uit het leerplan. Dit gaat ten koste van theoretische achtergronden en reflecties.

In het vierde en laatste paper bespraken Kelchtermans, Ballet, Berens en Janssens een onderzoek naar de micropolitiek in de professionele ontwikkeling van beginnende Vlaamse leraren basisonderwijs. De rapportage betrof vooral de interactie tussen de beginnende leerkracht en de andere actoren in de beroepscontext (i.c. de school als organisatie). De onderzoeksresultaten omvatten op de eerste plaats een inventaris van de verschillende soorten professionele belangen die in het professioneel functioneren van beginnende leerkrachten een rol spelen. Het onderzoek bevestigt de vroeger gemaakte categorisering van materiële, organisatorische, cultureel-ideologische, sociaal-professionele en zelfbelangen. Een tweede belangrijk resultaat betreft het proces van micropolitiek leren als onderdeel van de professionele ontwikkeling, nl. het ontwikkelen van micropolitieke geletterdheid, die gekenmerkt wordt door een kennisaspect, een belevingsaspect en een operationeel aspect.

Tijdens de rondetafel over *Reflectie van docenten* rapporteerden Longayroux, Verloop en Beijaard descriptief-exploratief onderzoek naar de inhoud en aard van de reflecties van de docenten-in-opleiding. Een vooronderzoek

beoogt te komen tot een categorieënsysteem om inhoud en aard van de reflecties te kunnen beschrijven. In het hoofdonderzoek zal dit categorieënsysteem gebruikt worden voor de analyse van portfolio's en diepte-interviews. Thijs en Van den Berg gingen in op de rol van peer coaching bij het implementeren van extern geïnitieerde curriculumvernieuwingen in de lespraktijk. De centrale vraag is hoe peer coaching, als onderdeel van een nascholingsprogramma met voorbeeldlesmateriaal, leraren kan ondersteunen bij het implementeren van leerlinggericht onderwijs in hun eigen onderwijssituatie in Botswana. Daarbij wordt voorbeeldlesmateriaal gebruikt als basis voor verschillende coachingsactiviteiten. De docenten blijken positief over deze coachingsactiviteiten, en noemen verschillende leeropbrengsten. Deze hebben vooral betrekking op nieuwe ideeën ten aanzien van de vakinhoud en didactische basisvaardigheden.

Het symposium *De pedagogische professionaliteit van de docent* bracht drie papers. Leerink en Klaassen verrichtten kwalitatief en kwantitatief onderzoek naar de vraag naar de gezamenlijke en aanvullende verantwoordelijkheid van gezin en school inzake de pedagogische opdracht. Ouders zowel als leraren zien geen contradictie tussen de doelen van conformeren en autonomie. Men ziet een duidelijke morele stellingname als belangrijke voorwaarde voor persoonlijke ontwikkeling en de vorming van acceptabel sociaal gedrag. Ouders en leraren zien in dat leraren een belangrijke rol spelen in conflicten binnen school, en geven aan dat leraren duidelijke regels moeten hantieren. In een tweede presentatie bespraken Theunissen, Van Veen, Klaassen, Slegers, Bergen en Hermans welke professionele en leven-beschouwelijke opvattingen van docenten een rol spelen in kritische situaties in de klas. Reactie-intenties van docenten werden in kaart gebracht. Docenten blijken sterk te hechten aan hun pedagogische professionaliteit. Ze zijn duidelijk op waarden en normen georiënteerd, al blijken ze in de beleving daarvan te verschillen: van lastige bijkomstigheid tot wezenlijke lestaak. De meeste docenten zien de pedagogische taak echter als een vanzelfsprekend onderdeel van hun functie. De doorwerking van de eigen levenbeschouwing blijkt problematisch, niet in het minst omdat docenten weinig vat op

die eigen levenbeschouwing blijken te hebben of twijfels daaromtrent koesteren. De derde bijdrage kwam van Luttenberg die onderzoek verricht naar normatieve aspecten in het beroepsmatig handelen van docenten? Het onderzoek richt zich op het ontwikkelen van zowel een conceptueel kader, als een methodologie nodig voor een zorgvuldige interpretatie van normatieve aspecten in dit handelen. Het onderzoek vormt een combinatie van literatuurstudie (filosofische invalshoek) en een kwalitatief-empirisch onderzoek onder docenten. Het onderzoek leidt tot een theoretisch-normatief model voor volwaardige rationaliteit in het beroepsmatig handelen. Het leidt tot een empirisch-normatief model van manieren van omgaan met normatieve aspecten in het beroepsmatig handelen van docenten. Dit model is tevens geschikt om te dienen als reflectiemodel voor docenten.

Beroepsopleiding en volwasseneneducatie

(D. Wildemeersch, Katholieke Universiteit Nijmegen, H. Baert, Katholieke Universiteit Leuven, Jeroen Onstenk, SCO-Kohnstamm, Amsterdam).

Het thema *Beroepsopleiding en volwasseneneducatie* is een jaarlijks weerkerend thema dat op een groeiende belangstelling mag rekenen. Dit jaar waren er 16 papersessies, drie symposia en twee ronde tafels. De papersessies werden gegroepeerd rond vier thema's. Het eerste thema *Werkplek leren* kende vier bijdragen. Een eerste bijdrage werd gepresenteerd door M. van Woerkom, W. Nijhof en L. Nieuwenhuis. Het ging over flexibel vakmanschap door leren op de werkplek. Hierin werd verslag uitgebracht over empirisch onderzoek in diverse arbeidsorganisaties over trends en ontwikkelingen inzake de lerende organisatie en het leren op de werkplek. Een tweede bijdrage ging over zoek- en leerprocessen in de beroepspraktijk en was van de hand van P. Gielen, L. Nieuwenhuis en A. Jager. Hierin werd verslag gedaan over een vierjarig onderzoeksprogramma binnen Stoas, waarbij de congruentie tussen leren in de beroepspraktijk en leren in het beroepsonderwijs centraal staat. De centrale vraagstelling hierbij is: Hoe kan de aansluiting onderwijs - arbeid verbeterd worden, uitgaande van een dynamische beroepspraktijk in de agrarische sector. Een derde thema, gepresenteerd door

T. Bastiaens en Y. Krul, ging over 'Performante en ondersteuning met on-line helpsystemen'. Er werd een effectiviteitsstudie ter sprake gebracht die de implementatie en effectevaluatie beschrijft van een on-line helpsysteem in een productiebedrijf. Dit systeem moet de werknemers ondersteunen bij processen van on-the-job-leren. Een vierde thema ging over werkstress en leren op de werkplek in het voortgezet onderwijs. Het was van de hand van K. Kwakman en R.J. Simons. In dit onderzoek werd het Job Demand-Control model van Karasek getoetst. Het model gaat ervan uit dat stress wordt veroorzaakt door een combinatie van taakeisen en de hoeveelheid controle die men bezit.

Een tweede papersessie had als thema *Kosten en baten*. Een eerste presentatie ging over de kosteneffectiviteit van levenslang leren. In het onderzoek van M. van Leeuwen werden de economische aspecten bestudeerd van enkele (belasting-)maatregelen ter bevordering van levenslang leren. Het onderzoek viel in drie delen uiteen: een literatuuronderzoek over het falen op de markt van permanente educatie, het verzamelen van gegevens rond werken en leren en de constructie van een rekenmodel. Een tweede presentatie was van de hand van M. de Steur en L. Borghans en had als thema *Werken buiten de zorgsector: push of pull-effect?*. Het ging over de ontwikkeling van een empirisch onderzoek over deze materie met als vertrekpunt volgende twee hypothesen. Zorgopgeleiden die niet werkzaam zijn in de zorgsector of een zorgberoep, beschouwen een zorgberoep als minder aantrekkelijk (pull-effect). Zorgopgeleiden die niet werkzaam zijn in de sector of een zorgberoep zijn minder geschikte arbeidskrachten voor zorgberoepen dan zorgopgeleiden die een zorgberoep uitoefenen (push-effect). Een derde presentatie ging over het ontwikkelen van curricula in de Faculteit Economie, Bedrijfs- en Bestuurskunde van de universiteit van Maastricht op basis van onderzoek naar competenties in de beroepspraktijk. Verslag over een kleinschalig kwalitatief empirisch onderzoek hierover werd gedaan door J. Boon, E. Bos en M. van der Klink. Tenslotte ging het in deze sessie over kosten en baten ook nog over de betekenis van modulaire systemen voor het secundair onderwijs in Vlaanderen. Daarbij werd vertrokken van een comparatieve

studie van ontwikkelingen in dit verband in Schotland en Nederland. Deze sterkte-zwakte analyse werd vervolgens in verband gebracht met nieuwe ontwikkelingen in Vlaanderen. Het onderzoek werd gedaan door K. van Valckenborgh, M. Douterlungne en L. Sels.

Het derde thema van de papersessies ging over *kwalificaties*. In een eerste presentatie gingen P. den Boer en T. Harms in op sociale beroepsvaardigheden. Het hoofddoel was het verhelderen van het begrip sociale beroepsvaardigheden. Daartoe werd een literatuuronderzoek opgezet en werden deskundigen geïnterviewd. Een tweede sessie ging over de vraag of specifieke kwalificaties nodig zijn om transities in loopbanen aan te kunnen. E. van Wenchem en M. Jellema knoopen hiervoor aan bij het begrip sleutelkwalificaties en bij de theoretische discussie hieromtrent. Vastgesteld werd dat de begripsvorming hieromtrent veel te weinig vanuit empirisch onderzoek wordt ingevuld. Dit leidde tot een empirisch onderzoek met als centrale vraagstelling: zijn er specifieke 'sleutel'kwalificaties nodig om adequaat te kunnen inspelen op veranderingen in de loopbaan? En zo ja: om welke kwalificaties gaat het dan, wat betekent dit voor de competenties waarover (aankomende) beroepsbeoefenaren moeten beschikken, en hoe kunnen die verworven worden? Een derde thema in dit verband was van de hand van S. van Zolingen. De titel was: '25 jaar sleutelkwalificaties. De analyse van een concept en consequenties voor het onderwijs'. In een uitgebreide literatuurstudie over dit thema werd onderzoek gedaan vanuit volgende vraagstelling: 'hoe is de ontwikkeling van het begrip sleutelkwalificaties en wat is de betekenis hiervan voor het onderwijs'? Tenslotte werd door J. van Loo, A. de Grip en M. de Steur verslag uitgebracht over kwalificatieveroudering en de rol van scholing en employability. De onderzoekers stelden vast dat rond dit thema hoofdzakelijk theoretische discussies worden gevoerd. Daarom werd gekozen voor een empirische testing van een aantal, uit de theorie afgeleide hypothesen, op grond van een aantal bestaande databestanden. De vierde en laatste papersessie ging over *Beroepsgerichte volwasseneneducatie*. Een eerste presentatie, vanwege A. de Grip en A. Jacobs, ging over de wijziging van het kwalificatieprofiel van mensen door het volgen van

niet-initiële scholing. Nagegaan werd in hoeverre in Nederland, na voltooiing van het initiële dagonderwijs, ook nog niet-initieel onderwijs wordt gevolgd, met als vraag in hoeverre hierdoor het kwalificatieprofiel veranderde. Voor het zoeken naar antwoorden op deze vragen werd gebruik gemaakt van data uit de onderwijsrekeningen van het CBS. Een tweede bijdrage binnen dit thema was van de hand van E. Legiest en D. Van Damme. Deze brachten verslag uit van een kwalitatief onderzoek naar de educatieve strategieën van participanten aan de volwasseneneducatie. Hierbij werd gebruik gemaakt van kwalitatief biografisch onderzoek bij een beperkt aantal cases. Daarbij werd o.m. vastgesteld dat de traditionele participatievariabelen in grote mate werden bevestigd, maar dat deze onvoldoende houvast bieden om participatiegedrag ten volle te begrijpen. Een derde presentatie ging over 'patronen van participatie in de volwasseneneducatie in Vlaanderen en Nederland' en was van de hand van E. Verhasselt. In dit onderzoek werd gebruik gemaakt van een bestaande dataset van de International Adult Literacy Survey (Ials). Zo werd onder meer vastgesteld dat volwassenen met een actieve vrijetijdsbesteding-stijl meer geneigd blijken te zijn om deel te nemen aan de volwasseneneducatie dan mensen met een passieve vrijetijdsbesteding-stijl. Een vierde paper, van de hand van G. Straetmans, ging over de wiscat: een adaptief toetspakket rekenen/wiskunde voor de BVE. Verslag werd uitgebracht over de constructie van een Wiskunde/rekenen Computergestuurd Adaptief Toetspakket. Het toetspakket bestaat uit een omvangrijke opgavenbank en een computerprogramma dat twee soorten toetsen kan afnemen en resultaten daarvan kan rapporteren naar de cursist.

De symposia waren uiteenlopend van aard. Een eerste symposium *Leren voor beroep* werd voorgezeten door C. Doets en becommentarieerd door J. Geurts. Binnen dit kader werden projecten gepresenteerd van achtereenvolgens, J. Onstenk, over de fundamenteën van concept van krachtige leeromgevingen voor de BVE-sector; van J. van den Berg en A. de Vries, over zeven nieuwe leerconcepten die binnen de BVE-sector worden gehanteerd; en vanwege Marijke Krol, over de invloed van virtuele omgevingen op het leren, met aandacht voor

het leren via internet. Een tweede symposium was gewijd aan het thema *'Een leven lang leren vanuit perspectief'* en werd voorgezeten door M. van der Kamp en becommentarieerd door B. Hövels. Er waren bijdragen van D. Wilde-meersch en T. Jansen over reflexieve modernisering en het toenemend belang van sociaal leren in deze context, vanuit een sociaal-pedagogisch perspectief. Jitte Brandsma presenteerde een inleiding over beroepsonderwijs, beroepsgerichte volwasseneneducatie en levensloopleven, vanuit een arbeidssociologisch perspectief. Hessel Oosterbeek presenteerde een inleiding vanuit onderwijs-economisch perspectief, met als sleutelbegrippen: 'human capital' en aanbod- en vraaggerichte financieringssystemen. Een derde symposium was gewijd aan de *'Learning history method'*. Dit bracht, onder begeleiding van D. Wilde-meersch, onderzoekers samen die onderzoek in voorbereiding of in uitvoering hebben waarin deze nieuwe interpretatieve methode, ontwikkeld aan het Massachusetts Institute of Technology, wordt gehanteerd. Het ging om F. Basten die verslag uitbracht van een onderzoek binnen de Nijmegen Business School naar de dynamiek van disciplinaire wetenschapstradities en onderwijspraktijken; om L. Verheijen die de learning history methode als sociaal leerproces duidde en om M. Coenders en R. Portengen die de learning history methode als instrument van kenniscreatie presenteerden.

Een eerste ronde tafel had als onderwerp *Een leven lang leren*. Hierbij was een inbreng van B. Hövels (discussiant), J. Fritman, E. Thomas, T. Eimers, F. van der Krogt, J. Onstenk, E. Voncken, R. Klarus, P. Gielen, K. Visser, W. Houtkoop en M. van der Kamp (voorzitter). Inzet van de discussie was de veranderingen die zich momenteel op macro-vlak in de samenleving voltrekken zoals individualisering, technologisering en internationalisering en daarmee samenhangend de revival van het concept 'levenslang leren'. In samenhang hiermee werden thema's besproken zoals werkend leren, erkenning van eerder verworven kwalificaties, voortijdig schoolverlaten en landelijke kwalificatiestructuur. Een tweede ronde tafel werd georganiseerd naar aanleiding van de key-note bijdrage van J. Mezirow over 'transformative learning van volwassenen'. De sessie werd voorgezeten door R. van der Veen.

Transformatief leren wordt opgevat als het zich bewust worden van de context van zijn eigen overtuigingen; de kritische reflectie over de assumpties die hieraan ten grondslag liggen, de verkenning van de justificaties en tenslotte de vertaling van deze inzichten in de richting van handelingsalternatieven.

Hoger onderwijs (J. Prins, Katholieke Universiteit Nijmegen, H. van Berkel, Rijks Universiteit Maastricht en J. Elen, Katholieke Universiteit Leuven)

In dit thema stonden centraal fundamenteel en toegepast onderzoek in relatie tot beleid aangaande het hoger onderwijs. De term 'beleid' is breed opgevat. Het kon zowel grootschalig kwantitatief als kleinschalig beleidsanalytisch onderzoek betreffen, zoals beleidsgericht onderzoek op instructieniveau (docenten, studeergedrag, leerprocessen), op institutioneel niveau (voor opleidingen en instellingen in WO en HBO) en op landelijk niveau (voor beleidsmakers bij overheden en koepelorganisaties). De gepresenteerde bijdragen aan dit thema zijn ingedeeld in enkele subthema's. Hieronder wordt in het kort de essentie van de verschillende bijdragen per subthema weergegeven.

In het thema *Macro-aangelegenheden in het hoger onderwijs* concentreerden drie bijdragen zich vooral op macro-aspecten in het hoger onderwijs. Cortens gaf een scherpe analyse over de relatie onderwijs-arbeidsmarkt. Ze ging in op trends in de Vlaamse arbeidsmarkt en samenleving welke uitdagingen vormen voor het onderwijs in Vlaanderen. Deze uitdaging zijn te vinden op het terrein van permanente vorming, en kwaliteitszorg in evaluatie en financiering. Universiteiten zullen moeten uitgroeien tot volwaardige expertisecentra. Van Damme ging in op de optimalisatie van het academisch opleidingsaanbod in Vlaanderen. Hij baseerde zich op twee rapporten dit aangaande van Dillemans. In zijn analyse richt hij zich op de conflicterende relatie van enerzijds deregulering en een toenemende autonomie van de Vlaamse universiteiten en anderzijds de planning van het opleidingsaanbod in Vlaanderen. Marktlogica inzake aanbodsplanning zou meer op zijn plaats zijn. In de Nederlandse situatie zien we meer en meer publicaties waar-

in opleidingen met elkaar worden vergeleken. Hulshof en Van den Broek gaan in op de betrouwbaarheid van deze ratings en rankings. Ze bespreken hoe opleidingen significant en substantieel van elkaar onderscheiden kunnen worden, waar rangordening inhoudelijk naar verwijst en wat de functie van rangordening is.

In het thema *Studiekeuze, selectie en instroom* hadden vijf bijdragen betrekking op selectie- en studiekeuzeaspecten bij de instroom van studenten. Van Leeuwen en Webbink presenteerden een onderzoek naar de mate waarin de studiekeuze van studenten te beïnvloeden zou zijn. Hun interesse ging met name uit naar de keuze voor een exacte studie. Uit een groot scala van stimulansen blijkt eigenlijk alleen dat een verhoging van de basisbeurs tot f 750,- en een stijging van het financiële voordeel van de student zijn belangstelling voor een exacte studie doet toenemen. Leerlingen moeten in een vroeg stadium van hun studieleven beslissingen nemen over hun gewenste toekomstige beroep en het daarbij behorende studiep pad. Van Leeuwen en Hop baseren hun resultaten op de Studiekeuzemonitor en concluderen dat de studiekeuze mede beïnvloed wordt door allerlei beleidsmaatregelen ten aanzien van bijv. studiefinanciering, toegestane studieduur, toelatingseisen, bindend studieadvies, invoering Studiehuis en keuze-profielen. Verder blijken van belang uniciteit opleiding, nabijheid bij ouderlijk huis en aansprekende voorlichting. Schittekatte en Koetsier zijn op zoek naar het profiel van de instromende studenten Psychologie en Pedagogische Wetenschappen. Naast het bestuderen van klassieke biografische gegevens hebben ze kwalitatieve parameters als intelligentie, persoonlijkheid, motivationele aspecten en zelfvertrouwen onderzocht. Persoonlijkheidskenmerken evenals resultaten op proefexamens blijken sterke predictoren voor studiesucces bij eerstejaars te zijn. Ook Lievens vraagt aandacht voor het toelatingsexamen, in dit geval voor 'arts en tandarts' in Vlaanderen. Hij onderzocht de predictieve, nomologische en face validiteit van het toelatingsexamen. Intelligentieproeven bleken een iet wat betere voorspeller te zijn dan casusonderdelen van het toelatingsexamen. Het casusgedeelte van het examen wordt door kandidaten als relevanter voor studie en beroep beschouwd dan intelli-

gentieproeven. De Open Universiteit experimenteert met formatieve voorkennistoetsen via internet. Martens en Hermans rapporteren onderzoek naar de bevindingen van deze voorkennistoets. Ze concluderen dat de inschrijvers de instrumenten nuttig vonden, de bedoeling duidelijk vonden, maar dat het maar bij weinig studenten de keuze had beïnvloed.

Het subthema *Het beïnvloeden van studiegedrag en studieloopbaan* omvatte vier bijdragen op het gebied van studieloopbaan en hoe deze te beïnvloeden. Van der Hulst en Jansen richtten zich in hun onderzoek op de effecten van curriculumveranderingen op studievoortgang en studiestaking en constateren dat er vooralsnog geen duidelijke positieve relatie is tussen curriculumverandering en studiesucces, maar mogelijk wel tussen curriculumverandering en vervroegde studieuitval. Otting, Stuurwold en Zwaal beschrijven een onderzoek naar de effecten van entreesselectie. Ze constateerden dat VWO'ers aanmerkelijke betere studieprestaties leveren in een drietal opleidingen in de economische dienstverlening sector dan havisten en MBO'ers. Een structurele selectie hierop bij één opleiding leidde tot hogere slaagpercentages dan bij de andere twee opleidingen waar de studenten via een lotingsstelsel toegelaten werden. In Leiden probeert men volgens Elsen door de invoering van een bindend negatief studieadvies en de invoering van het Leids Studiesysteem invloed uit te oefenen op studievoortgang en zelfselectie bij studenten. Uit interviews met studiestakers kwamen een aantal aspecten naar voren waarmee het studiekeuzep proces bij studenten te beïnvloeden lijkt. Beekhoven, De Jong en Van Hout hebben een model (Amsterdams model) ontwikkeld voor het verklaren van studievertraging en studieuitval, gebaseerd op de theorie van Tinto en de Human Capital theorie. Gepresenteerd werd onderzoek naar de effecten van het samenvoegen van de factoren academische en sociale integratie op het verklaren van studievoortgang.

In het symposium *Kwaliteitsaspecten van academische vorming en kritische beroepsvaardigheden* stond het begrip 'academische vorming' centraal. De bijdragen besteedden aandacht aan de belangrijkste knelpunten en aandachtspunten in onderzoek en ontwikkeling van academische vorming en kritische

beroepsvaardigheden. Allereerst het werkveld waarvoor het onderwijs moet voorbereiden: gaat het primair om het opleiden van onderzoekers of betreft het eerder de vorming van intellectuele experts voor een grote diversiteit van functies? Ook van belang daarbij is de vraag naar de maatschappelijke context waarvoor opgeleid wordt. Borghans presenteerde een onderzoek onder afgestudeerde economen naar de vraag of ze dezelfde opleiding opnieuw zouden doen, terugkijkend op hun loopbaan. Welke vaardigheden hebben ze tijdens hun opleiding daadwerkelijk gemist? Holleman, Admiraal, Oost, Milius en Van den Berg hebben op basis van een uitgebreide literatuurstudie een vormingsprofiel opgesteld dat passend zou zijn voor iedere academicus. Verrijt en Mulders onderzochten afgestudeerden met een bèta-opleiding. Het doel van het onderzoek was te achterhalen welke kennis en competenties vereist zijn om effectief te kunnen functioneren in de beroepsomgeving. In de tweede plaats gaat het om de leerlijnen in een curriculum, waarlangs de student systematisch en consistent wordt gestimuleerd deze vaardigheden te ontwikkelen. Terlouw presenteerde de 'rode draad' in een HO-curriculum. Welke ontwerp-opdrachten moeten afgestudeerden van een opleiding kunnen aanpakken en waaruit bestaat een model van een ontwerpaanpak die kan worden gezien als de kern van een opleiding en die zou kunnen dienen als een 'rode draad' voor het curriculum. Hij kon vier niveaus van complexiteit onderscheiden. Van Til, Blom en Holen deden onderzoek naar de ontwikkeling van beroepsvaardigheden tijdens ontwerpprojecten. Ze vroegen studenten de eigen beheersing van vaardigheden in te schatten en het verloop hierin in relatie tot het werken aan een project aan te geven. De onderzoekers konden geen verbetering in de beheersing van vaardigheden aantonen. De redenen hiervoor konden gezocht worden in de context van het curriculum.

In een ronde tafel discussie over *ICT in het hoger onderwijs* zijn twee papers ingebracht over het subthema ICT. Van der Wende, Beerkens, Collis en Moonen voerden een internationale studie uit naar het gebruik van ICT in het hoger onderwijs. Ze richtten hun onderzoek op drie aspecten: gehanteerde onderwijskundige modellen, kosten en efficiëntie, en de beleids-

implicaties er van in enkele landen rondom Nederland. Van Tartwijk, Veen, Lam, Pilot, Geloven, Peters en Moonen hebben een grootschalig onderzoek uitgevoerd naar het educatief gebruik van ICT in het Nederlandse hoger onderwijs. Men probeerde een beeld te krijgen van de schaal waarop ICT wordt ingezet ten behoeve van leeractiviteiten van studenten, wat als belangrijkste meerwaarde van ICT-gebruik wordt gezien en welke knelpunten zich voordoen bij de invoering en het gebruik van ICT in het onderwijs.

Tenslotte bleken er enkele papers ingestuurd te zijn die moeilijk te plaatsen bleken te zijn in voornoemde thema's. Van Dijk presenteerde een experimenteel onderzoek naar de effecten van activering in colleges. Studiepresentsaties zouden positief beïnvloeden kunnen worden indien we studenten kunnen activeren tijdens de colleges, was haar uitgangspunt. Het bleek dat activering met behulp van een elektronisch stemkastje en aangevuld met peer instruction de betrokkenheid en motivatie van studenten verhoogd. Jansen deed onderzoek naar de effectiviteit van instructie op het studeergedrag en de studievoortgang in het wetenschappelijk onderwijs. De effectiviteit van instructie werd in beeld gebracht door analyse van studiegidsteksten en interviews met docenten. De resultaten onderscheiden naar schalen op het gebied van werkcollege, beoordeling en studentgedrag bleken positief gerelateerd te zijn aan behaalde tentamencijfers. De bijdrage van Sluijsmans en Moerkerke richtte zich op peer assessment in probleemgestuurde onderwijsomgevingen. Ze gingen na hoe betrouwbaar studenten elkaar kunnen beoordelen en hoe groot de acceptatie van deze oordelen is. Studenten bleken inderdaad in staat te zijn tot betrouwbare oordelen te komen. Bovendien verhoogde het de betrokkenheid van de studenten. Wel hebben studenten meer behoefte aan instructie training in peer assessmenttechnieken.

Beleid en organisatie in het onderwijs (F.C.G. Smit, Katholieke Universiteit Nijmegen, J.F.A. Braster Erasmusuniversiteit Rotterdam en P. van Petegem, Universiteit Antwerpen)

De activiteiten van de divisie *Beleid en Organisatie in het Onderwijs* hebben tijdens de

ORD'99 betrekking op vier papersessies en een ronde tafeldiscussie over innovaties in het hoger onderwijs. Daarnaast zijn enkele leden van de divisie betrokken bij de organisatie van thema-overstijgende bijeenkomsten. In deze bijdrage concentreren wij ons op de inhoud van de papersessies.

De eerste papersessie laat zien dat onderzoek naar de verwevenheid van levensbeschouwing en onderwijsbeleid nog steeds een belangrijk thema is in Nederland en België. In Nederland en België hebben ouders het recht om hun kinderen aan te melden voor de school van hun voorkeur. Die voorkeur kan zijn grond vinden in uiteenlopende motieven. De levensbeschouwelijke grondslag van de school is daarbij niet meer het belangrijkste keuzemotief, zo blijkt onder andere uit een recent onderzoek van Van Petegem en Brandt (Universiteit Antwerpen) over schoolkeuzes in het Vlaamse basisonderwijs. Motieven die blijkens dit onderzoek een belangrijker rol spelen zijn nabijheid en bereikbaarheid van de school. Nederlands onderzoek levert dezelfde conclusie op.

In België is de vrije schoolkeuze grondwettelijk geregeld, terwijl dat in Nederland niet het geval is. Bijzondere scholen in Nederland én 'vrije gesubsidieerde scholen' in België kunnen in beginsel leerlingen de toegang weigeren. Uit vergelijkend juridisch onderzoek van Laemers (ITS/KUN) blijkt dat niet kan worden geconcludeerd dat de schoolkeuzevrijheid in België in de praktijk groter is dan in Nederland. Dat komt door het ontbreken van de financiële gelijkstelling tussen 'openbaar' en 'bijzonder' onderwijs in België. In schijnbare tegenspraak met trends in de richting van secularisering werden Braster, Hofman en Schenkeveld (RISBO/EUR) in Rotterdam geconfronteerd met een toegenomen belangstelling van ouders voor levensbeschouwelijk vormingsonderwijs binnen het openbaar onderwijs. Een verklaring daarvoor wordt gezocht in een combinatie van factoren: revitalisering van het belang van waarden en normen in onze samenleving en aandacht voor de rol die de school daarin kan spelen, een hernieuwde aandacht voor levensbeschouwing los van formele structuren, een toenemende open houding van ouders in het openbaar onderwijs tegenover godsdienstonderwijs en de sterk veranderde instroom in de

openbare scholen, met name de toename van kinderen met een islamitische achtergrond.

Smit, Pelkmans en Slegers (ITS/KUN) richten zich in hun paper op de levensbeschouwelijke identiteit van Regionale Opleidingen Centra. De directies van dergelijke instellingen slagen er nog maar ten dele in om een draagvlak te creëren voor hun identiteitsplannen. Dit komt omdat allerlei managementactiviteiten, zoals het organiseren, regelen en voorwaarden scheppen voor het functioneren van identiteitscommissies en -werkgroepen, de meeste energie opeisen en er weinig tijd overblijft voor andere activiteiten, zoals het coachen, stimuleren en corrigeren van (godsdienst)docenten om het identiteitsbeleid daadwerkelijk van de grond te krijgen. Daardoor blijft het levensbeschouwelijk element op de meeste scholen een 'hobby' van enkelen in plaats van een 'professie' voor velen.

De tweede papersessie, met als onderwerp *Effecten en opvattingen over onderwijsbeleid*, trok meer mensen dan er stoelen aanwezig waren in het collegezaaltje. Deze relatief sterke belangstelling kan te maken met het feit dat de onderzoekers binnen deze sessie de praktische consequenties aan de orde stelden van politiek actuele issues als basisvorming, kwaliteitszorg en lokaal onderwijsbestuur. Zo signaleren wat betreft de basisvorming Cremers-Van Wees en Brandsma (UT) een trend in de richting van meer modernisering en harmonisering. Rutgers (EUR) doet vervolgens verslag van een casestudy naar de implementatie van effecten van bestuurlijke vernieuwing binnen het lokaal onderwijsbestuur in een gemeente. Eén van zijn voornaamste conclusies is dat in een omgeving waar traditioneel weinig behoefte bestaat aan overleg en verandering 'bestuurlijke vernieuwing' niet verder reikt dan beheersmatige aandacht voor verplichte zaken als schoolverzuim, huisvesting, schoolbegeleiding en onderwijsachterstanden. Het wachten is nu op de dissertatie waarin de analyses van meerdere cases bij elkaar worden gebracht.

In de bijdrage van Devis e.a. (Katholieke Universiteit Leuven) worden de consequenties van overheidsbeleid ook doorgetrokken naar het niveau van de school. Duidelijk wordt dat centrale beleidsmaatregelen inzake kwaliteitszorg in het Vlaamse basisonderwijs op schoolniveau verstrikt kunnen raken in processen van interpretatie, onderhandeling of zelfs strijd

vanuit de professionele belangen van directies en leerkrachten.

De Maeyer en Van Petegem (Universiteit Antwerpen) confronteren vervolgens de onderwijsprioriteiten van schoolleiders met die van het grote publiek. Ook in dit Vlaamse onderzoek, dat is verricht in het kader van het groot-schalig INES-project (Indicators of Education Systems) van de OESO, is sprake van een spanningsverhouding. Analyses met behulp van LISREL tonen b.v. aan dat het meetmodel voor onderwijsopvattingen er anders uit ziet in de afzonderlijke subpopulaties.

In de derde papersessie over *Schoolmanagement* maakt de bijdrage van Vermeulen, Evers en Teurling (IVA/KUB) duidelijk dat het Human Resources Management in het primair onderwijs nog maar in de kinderschoenen staat, terwijl het paper van Doolaard (UT) laat zien dat er in het afgelopen decennium niet veel is veranderd wat betreft de invulling van het schoolleiderschap. De uitdaging waar directies van instellingen in het (voorbereidend) beroepsonderwijs de komende jaren voor staan, is het opbouwen van een organisatie die flexibel en snel haar expertise kan inzetten op zeer diverse vragen naar kennisoverdracht en vaardigheidstraining. Nieuwe onderwijskundige ontwikkelingen binnen deze sector vragen om verdere professionalisering van leraren: functies zouden enerzijds kunnen worden verbreed en anderzijds in beperkte mate worden gedifferentieerd. En er moet vooral meer worden samengewerkt. Met betrekking tot dat laatste omarmen steeds meer opleidingen met wisselend succes het concept van zelfsturende teams, aldus Pelkmans, Smit & Slegers (ITS/KUN).

In de laatste papersessie over ontwikkelingen in beleid en organisatie werd een tipje opgelicht van de sluier van wat de school als organisatie bijdraagt aan de realisatie van vernieuwingen in de lespraktijk door docenten. Uit onderzoek van Geijssels, Slegers en Van den Berg (KUN) naar condities voor de invoering van vernieuwingen in de agrarische opleidingscentra blijkt dat grootschalige innovaties doorgaans in beperkte mate de lespraktijk van docenten blijken te beïnvloeden. Het gehanteerde model maakt duidelijk dat daarbij rekening gehouden moet worden met de subjectieve percepties van docenten ten aanzien

van hun werkomgeving. Tenslotte gaf Polder (SCO/UvA) een beschrijving van het bestuurlijk decentralisatieproces op het gebied van volwasseneneducatie. Er is een algemene tendens dat gemeenten meerjarige raamovereenkomsten afsluiten met regionale opleidingen centra. Van de beoogde concurrentie tussen instellingen komt in de praktijk echter weinig terecht. Als een gemeente besluit over te stappen naar een andere instelling loopt zij het risico te moeten opdraaien voor de schade die de in de steek gelaten instelling hiervan ondervindt.

Coöperatief leren (S. Veenman, Katholieke Universiteit Nijmegen, F. de Jong, Universiteit Wageningen en J.P. Verhaege, Universiteit Gent)

Het lokaal gekozen thema voor de Onderwijs Research Dagen van dit jaar bestond uit drie papersessies, twee symposia, een postersessie en een ronde tafeldiscussie. Op de eerste dag stond vooral het samenwerken in coöperatieve groepen centraal waarbij de interactie tussen de groepsleden wordt gemedieerd door de computer (Computer Supported Collaborative Learning), op de tweede dag het samenwerken in coöperatieve groepen waarbij de groepsleden direct in een face-to-face relatie met elkaar samenwerken.

De papersessie *Computer-ondersteunde samenwerkende leeromgevingen in het hoger onderwijs* werd geopend met een presentatie waarin Ronteltrap en Eurelings (UM) ingingen op de aansluiting van ICT bij de uitgangspunten van probleemgestuurd onderwijs (PGO). In twee experimenten waarbij de POLARIS-leeromgeving werd ingezet, werden implementatiegegevens verzameld door observaties, interviews met studenten en tutores en de server logdata. Analyse van deze gegevens en inhoudsanalyses van de rapportage-database leidden tot de volgende conclusies: Studenten besteedden meer tijd aan het zelfstandig verzamelen en schriftelijk verwerken van informatie; de interactie in de onderwijsgroep veranderde van een 'rondje verslag doen' naar een bespreking van de inhoudelijke database; de betrokkenheid van studenten werd positief beïnvloed door de mogelijkheid van het asynchroon samenwerken. Tutores hadden meer tijd nodig omdat ze meer producten van stu-

denten moesten lezen. In de tweede presentatie ging Veldhuis-Diermanse (LUW) in op hoe studenten leren binnen een CSCL-omgeving. Zij stelde dat onwetendheid bij (universiteit)-docenten omtrent het leerrendement en de meerwaarde van CSCL debet is aan het geringe gebruik ervan in het academisch onderwijs. Met haar onderzoek wil zij derhalve meer inzicht geven in het (samenwerkend) leerproces in CSCL-leeromgevingen en de invloed van verschillen in studentkenmerken op dat proces. Binnen de opleiding Bodemkunde en Geologie introduceerde zij het kennisoverleg platform Web-Knowledge Forum (WebKF) dat ontwikkeld is door Scardamalia en Bereiter (OISE, Toronto). Zowel voorkennis als leerstijlen van studenten werden geanalyseerd. Uit de resultaten blijkt dat naarmate studenten meer voorkennis hebben zij minder aantekeningen (zgn. 'notes') creëren en lezen, maar dat deze wel meer tekst bevatten. Wat betreft leerstijlen blijkt dat wanneer studenten meer een stapsgewijze verwerking hanteren zij meer woorden in hun notes gebruiken. Studenten die meer georiënteerd zijn op een diepere verwerking besteedden minder tijd aan het schrijven van hun notes. Studenten die sterk gericht zijn op externe sturing besteedden juist weer minder tijd in het WebKF. Zoals ook discussant Van der Meijden opmerkte is een analyse naar de relatie met de kwaliteit van de notes een noodzakelijke volgende stap. Dit commentaar geldt ook voor de bijdrage van Strijbos (KUN) die in zijn onderzoek de relatie naging tussen de samenstelling van groepen op basis van hun voorkennis en de participatie in WebCSILE (voorganger van WebKF). Bij een cursus bleken studenten Onderwijskunde in een groep die heterogeen was samengesteld qua voorkennis het meest constructief. Dit omdat zij de meest progressieve ontwikkeling vertoonden wat betreft het creëren van en reageren op notes. Ook hun leerprestatie nam sterker toe in relatie tot de homogene groepen met veel voorkennis.

In het symposium *Implementatie van ICT als collaboratieve leeromgeving: CL-Net* onder leiding van F. de Jong presenteerde zich het Nederlands/Belgische cluster van het Europese onderzoekproject Collaborative Learning Networks (CL-Net). Samen met onderzoekpartners in Griekenland, Italië en Finland houden zij zich bezig met de implementatie van CSCL

in het basis-, voortgezet en beroepsonderwijs. Centraal staat het onderzoek naar de didactische en cognitieve aspecten bij het bevorderen van intentioneel leren en kennisconstructie in leergemeenschappen met behulp van educatieve multimedia en GroupWare in het primair, secundair en beroepsonderwijs. Hoewel elke partner vanuit zijn eigen expertise, onderwijssector en onderwijsinhoud opereert worden gemeenschappelijke vraagstellingen, metacognitie en motivatie vragenlijsten, pre- post-test onderzoekopzet voor cognitieve toetsen en het format van actieonderzoek gehanteerd. De participanten binnen dit symposium gebruiken allen WebKF als CSCL-leeromgeving. WebKF is een Internetsysteem met een centrale database die (mede) opgebouwd wordt door de leerlingen zelf. Dhert (KU Leuven) presenteerde het Belgische onderzoek dat gericht is op de ontwikkeling van een (meta)cognitieve strategie voor het vaardig oplossen van wiskundige toepassingsproblemen bij leerlingen van de bovenbouw van de basisschool. Gebruikmakend van WebKF leren leerlingen samen wiskundige toepassingsproblemen te stellen, op te lossen en te reflecteren op hun oplossingsprocessen. In de eerste fase van kennismaking was de plaats van KF beperkt, welomlijnd en sterk gestuurd door de leerkracht. Naarmate de vaardigheid van de leerlingen in het gebruik ervan vorderde, nam de intensiteit, de veelzijdigheid en de zelfstandigheid van het gebruik WebKF door de leerlingen gradueel toe. Het Amsterdamse onderzoek (Slighte en Lauret) dat is ingebed in CIAO maakt gebruik van het Leuvense wiskunde-materiaal. De basisschoolleerkracht gebruikt WebKF echter in een veel minder strak gestuurde context. Ook richt dit onderzoek zich meer op (internationale) uitwisseling van concepten met betrekking tot alledaagse onderwerpen zoals 'Waarom geen pet op in de klas'. Het Wageningse onderzoek (Lutgens) is ingebed in het middelbaar agrarisch onderwijs. Ingegaan werd op het WebKF gebruik in afstandsituaties en op de aspecten van het coachen op afstand. Hier was sprake van een nog lossere sturingscontext dan in het Amsterdamse onderzoek. Uit de analyse van geschreven en gelezen notes blijkt op basis van vier experimenten dat er een consistente relatie bestaat tussen de mate van interventie (goede opdrachten, op het juiste moment reageren etc.)

door de docenten en de mate van participatie door de studenten. Studenten zeggen ook behoefte te hebben aan ondersteuning door de docent. Van der Meijden en Trimpe presenteerden het Nijmeegse WebKF gebruik in het middelbaar onderwijs bij de vakken biologie, natuurkunde en geschiedenis. Zij sluiten nauw aan bij onderwerpen in de reguliere leerboeken en benaderen deze vanuit diverse invalshoeken en complexiteit. Alle vier CL-Net partners maken gebruik van data uit systematische klasobservaties, audio-opnames en analyse van de KF-notes, logboekanalyse en interviews met leerkrachten. Aangezien de meeste onderzoeken nog in volle gang zijn werden slechts summiere en meestal kwantitatieve resultaten gepresenteerd. Als discussiant merkte Kanseelaar dan ook op dat in de toekomst voorgenomen, kwalitatieve analyses van de leerlingbijdragen meer dan de kwantitatieve indicatoren inzicht moeten geven in de invloed van de CL-N's op het intentioneel leren, waarbij het leerproces gezien wordt als het construeren van kennis en de klas als een leergemeenschap. Hij onderstreepte het belang van een dergelijke internationale aanpak en bundeling van expertise binnen het project en gaf tevens aan dat twee jaar te kort is om nu al effecten te laten zien. Continuering is nodig.

In het symposium *Samenwerkend leren om samen te leren werken: het trainen van teams* onder leiding van A. Schaafstal (TNO) werd onderzoek besproken naar methoden voor de training van teamvaardigheden zoals dat wordt uitgevoerd door de Afdeling Training en Opleiding van de TNO Technische Menskunde. Door technologische, organisatorische en politieke ontwikkelingen neemt de complexiteit binnen werkomgevingen toe. Als gevolg daarvan kunnen allerlei taken niet meer door één individu worden uitgevoerd, maar moeten mensen in teamverband samenwerken. Het louter groeperen van een aantal experts leidt niet automatisch tot een goede teamprestatie; een goede coördinatie en communicatie zijn van cruciaal belang voor een effectief functioneren van team. Het symposium richtte zich derhalve op methoden voor de verwerving van team taakvaardigheden. In de bijdrage van Bots en Schaafstal stond de methode van crosstraining centraal, een trainingsstrategie waarin teamleden worden getraind in de taken,

rollen en verantwoordelijkheden van de overige teamleden. Uit de resultaten van een tweetal experimenten bleek dat expliciete instructie in de relaties tussen eigen taken en andermans taken en veelbelovende strategie is om de communicatie en prestatie van teams te verbeteren. Een oefening op elkaars posities bleek niet te leiden tot een betere teamprestatie. Schaafstal en Van Berlo beschreven vervolgens een praktische toepassing van crosstraining in het trainen van een leidinggevende aan boord van marine-fregatten, die, om zijn taak goed uit te kunnen voeren, enerzijds goed moeten kunnen opereren in zijn eigen team, en anderzijds, voldoende moet weten over andere dienstvakken aan boord. Van Berlo analyseerde in zijn bijdrage verschillende modellen en richtlijnen ter ondersteuning van het analyseren van teamtaken en het ontwerpen van trainingsscenario's, terwijl Riemersma, in het verlengde daarvan zich richtte op de vraag hoe teamprestaties moeten worden vormgegeven, zodat goede feedback kan worden gegeven aan teams over hun prestaties tijdens de training. Wil het taakgedrag gecoördineerd en soepel kunnen verlopen dan eist dit een zekere mate van voorkennis bij de teamleden in de vorm van op elkaar afgestemde (mentale) modellen, zoals een gemeenschappelijk situatiemodel, een gemeenschappelijk teamtaakmodel, model van de taak van het ander teamlid en het model van de eigen taak. Gestimuleerd door de discussiant Van der Krogt ontstond zich tot slot een discussie over de verschillen tussen enerzijds teamleren en samenwerkend leren en anderzijds in een samenwerkingsverband leren rekening houden met de taken van een ander.

In de papersessie *Peer tutoring en de toepassing van coöperatief leren in het basisonderwijs* werden vier papers gepresenteerd. Uit een onderzoek naar hoe leerlingen samenwerken en uitleg geven in kleine groepen (Terwel, VUA en UvA), bleken coöperatief gedrag (samenwerken) en verbale interactie (uitleg geven) sterk gelieerd te zijn. Zowel aanleg als training in communicatieve vaardigheden hadden een positief effect. Sterkere en getrainde leerlingen scoorden hoger op samenwerking en gaven meer uitleg op verzoek. Dit kan bijkomend helpen verklaren waarom uitleg op verzoek effectiever is: ze wordt eerder door sterkere leerlingen gegeven. Ook blijkt dat in

klassen met een hoog niveau leerlingen elkaar vaker uitleg geven. Een ontwikkeling doorheen de tijd werd niet vastgesteld. Wat leerkrachten die zeggen coöperatief leren te hanteren, precies doen, blijkt niet volledig overeen te stemmen met wat in de literatuur als kenmerken van CL wordt vooropgesteld, zo bleek uit de paper voorgesteld door Veenman (KUN). Aan gedragsregels voor effectief samenwerken bijvoorbeeld wordt weinig aandacht besteed. Voor de werkverdeling binnen de coöperatieve groepen, het luisteren naar elkaar, het besluitvormingsproces werden slechts lage scores toegekend. Uit het onderzoek voorgesteld door Vosse (UvA) bleek dat pull out peer tutoring voor rekenen leidt tot betere leerprestaties bij de tutees (= leerlingen met rekenproblemen uit groep 4), maar niet tot een verbetering van de scores voor 'self-efficacy' voor rekenen in de gewone klascontext. Vedder (RUL) presenteerde de resultaten van een externe evaluatie van het cross age peer tutoring programma 'Stap Door!'. Er is sprake van een positief effect op het welbevinden van de tutors (groep 7), in het bijzonder bij de Marokkaanse leerlingen. Ten aanzien van leesvaardigheid, woordenschatontwikkeling en leesplezier werd bij tutors noch tutees enig positief effect gevonden. Bij de tutees was er bij woordenschat zelfs een negatief effect. Oorzaken van de beperkte effectiviteit van het programma moeten volgens de onderzoeker gezocht worden in de manke implementatie van het programma (te weinig intensief, te weinig coaching door leerkrachten, leesboekjes van te laag niveau). Discussiant Van der Linden verbaasde zich over het wijd verbreide enthousiasme voor CL, dat volgens hem algemeen te simpel voorgesteld wordt als 'situaties waarin leerlingen van en met elkaar leren', wat op veel verschillende wijzen kan worden ingevuld. Men kan dus moeilijk spreken van DE processen of DE effecten van CL. Inzake het invoeren van CL zou meer terughoudend volgens hem gepast zijn. De verklaringen voor de (evt.) effectiviteit van CL blijven vooralsnog zwak te noemen. In de daaropvolgende discussie wordt zowel gewezen op de noodzaak van een goede en correcte implementatie als op de beperkte meetbaarheid van heel wat positief ervaren effecten, in het bijzonder op het vlak van de sociaal-emotionele ontwikkeling. Vosse meent dat zelfs bij nulwinst op cognitief vlak de winst op

sociaal-emotioneel vlak PT en CL nog de moeite waard maakt.

In de papersessie *Coöperatief Leren in het voortgezet onderwijs* werden vier papers gepresenteerd. In het eerste paper van Luttik en Erkens (RU) stond de vraag centraal hoe in het kader van het studiehuis in de bovenbouw van het voortgezet samenwerkend leren vorm krijgt. Hierover is nog weinig bekend. Onderzocht werd hoe de vervulling van verschillende leerfuncties plaats vond. Uit de resultaten van het survey-onderzoek bleek dat de docenten de taken die te maken hebben met het voorbereiden, het reguleren en het motiveren van het samenwerkingsproces vaak zelf in de hand houden. Als samenwerkend leren wordt toegepast dat gebeurt dit meestal in de vorm van zelfstandig samenwerken. Gezamenlijk leren komt af en toe voor, terwijl zelflerende teams zelden voorkomen in het voortgezet onderwijs. Docenten vonden samenwerkend leren over het algemeen een nuttige, maar moeilijke werkvorm, die past binnen alle vakgebieden. In het tweede paper van Van Boxtel (RU) werd de ontwikkeling van natuurkundige begrippen door participatie in interactie en het effect van het gebruik van schoolboeken op de kwaliteit van de interactie aan de orde gesteld. Het onderzoek dat werd uitgevoerd in 4 Havo. In een experiment werkten de leerlingen in tweetallen een concept-mapping opdracht en een poster-opdracht over elektriciteit uit. Uit de analyse van de gesprekken van de leerlingen bleek dat concept-mapping opdracht, waarbij de leerlingen een woordenweb maakten van de relaties tussen de natuurkundige begrippen, meer dan de poster-opdracht resulteerde in het praten over natuurkundige begrippen, elaboratie van conflicten en redeneren. In een experiment waarin de leerlingen bij de uitwerking van de opdracht gebruik mochten maken van schoolboeken bleek dat het raadplegen van boeken negatief correleerde met het voorkomen van elaboratie in de interactie. In de tijd dat er gebladerd en gelezen werd in het boek was er nauwelijks sprake van elaboratie van het gelezene, de boekkennis werd meestal zonder meer overgenomen in het groepsproduct. In het derde paper stelde Raemdonk (Universiteit Gent) de relatie tussen klasmanagement en de leervoorkeurstijlen bij leerlingen in het zesde jaar secundair onderwijs in Vlaanderen aan de

orde. Zij stelde zich onder meer de vraag of leerlingen onder invloed van een op samenwerken georiënteerd onderwijs een voorkeur ontwikkelen voor coöperatief leren. Zij verwachtte in haar onderzoek een verband te vinden tussen de manier van werken in de klas en de leervoorkeurstijlen van de leerlingen. Uit de resultaten kon niet geconcludeerd worden dat er een positief verband bestond tussen de variabelen 'coöperatief klasmanagement' en 'coöperatieve leervoorkeur'. Wel werd een negatieve relatie gevonden tussen 'een op individueel leren gericht klasmanagement' en de 'voorkeur van leerlingen voor individueel leren'. Een toename van een op individueel leren gericht klasmanagement ging gepaard met een afname van de voorkeur van leerlingen voor individueel leren. Verder bleek een toename in individueel klasmanagement samen te gaan met een toename in een coöperatieve leervoorkeur. In het laatste en afsluitende paper stelden Jörg en Hollander (RU) het thema 'samenwerkend leren tussen theorie en praktijk' aan de orde. Volgens hun opvatting is er tot nu toe weinig aan echte theorievorming gedaan. Vanuit een psychologische invalshoek ontwikkelden zij aanzetten voor de modellering van samenwerkend leren als de interactie tussen twee participanten waarbij wordt uitgegaan van een reciproke relatie met bijbehorende interactie die gezien kan worden als een proces van causale beïnvloeding. Leren binnen samenwerkend leren wordt geconstitueerd in de interactie; de effecten van dit proces worden niet alleen geconstrueerd in, maar ook gemedieerd door de interactie. Dit maakt het leerproces complex. Als discussiant wees Terwel er op dat dit model verdere aandacht verdient en dat het mogelijkheden biedt om de complexe interacties tussen tweetallen in beeld te brengen.

In de ronde tafeldiscussie introduceerden de coördinatoren Van der Linden en Kanselaar het nieuwe ICO-thema *Interactief Leren*. Leden van de landelijke onderzoeksschool en belangstellenden discussieerden met elkaar over drie vragen: hoe uitdagend en bindend is het thema voor toekomstig onderzoek, op welke thema's is samenwerking mogelijk en waartoe dienen onze gecoördineerde acties te leiden? Nadat de participanten zich hadden voorgesteld met hun onderzoek ontspon zich een discussie over interactief leren. Enerzijds kent het thema het

sociale face-to-face gemedieerde leren, anderzijds de computer ondersteunde gemedieerde kant van het leren. Bijzondere aandacht kregen de kenmerken van de taaksituatie bij interactief leren. Besloten werd de samenwerking te zoeken in de richting van het gezamenlijk formuleren van onderzoeksvorstellen, het uitwisselen van ervaringen ook met betrekking tot de gebruikte software en het werken aan verdere theorievorming op het vlak van interactief leren.

Aan deze kroniek werkten mee:

C. Aarnoutse, H. Baert, H. van Berkel, J.F.A. Braster, J. Buitink, J. van Damme, A. Dijkstra, J. Elen, F. De Jong, Libotton, N. Nieveen, P. Onghena, J. Onstenk, P. Van Petegem, J. Pieters, J. Prins, P. Sanders, P. Slegers, F.C.G. Smit, R. Soetaert, S. Veenman, P. Van de Ven, J.P. Verhaege, L. Verhoeven, L. Verschaffel, M. Voeten, A. Wesselingh, D. Wildemeersch

De eindredactie werd verzorgd door *P. Slegers* en *L. Verhoeven*.

Mededelingen

Oraties

Prof. dr. P.J.C. Slegers

Op vrijdag 25 juni sprak Peter Slegers zijn intrede Leiding geven aan leren uit. Daarmee aanvaardde hij het hoogleraarschap 'Onderwijskunde, in het bijzonder onderwijs- en opleidingsmanagement' aan de Katholieke Universiteit Nijmegen (Faculteit der Sociale Wetenschappen).

Prof. dr. J.J.G. van Merriënboer

Op donderdag 1 juli sprak Jeroen van Merriënboer zijn intrede Cognitie en multimedia-ontwerp voor complex leren uit. Daarmee aanvaardde hij het hoogleraarschap 'Onderwijstechnologisch onderzoek' aan de Open Universiteit Nederland (Onderwijstechnologisch Expertisecentrum OTEC).

Prof. dr. A.J.M. de Jong

Op donderdag 30 september om 16.00 uur, sprak Ton de Jong zijn intrede De proef of de som uit. Daarmee aanvaardde hij het hoogleraarschap 'Instructietechnologie' aan de Universiteit Twente (Faculteit Toegepaste Onderwijskunde).

Prof. dr. G.T.M. ten Dam

Op donderdag 4 november om 14.30 uur, spreekt Geert ten Dam haar intrede uit. Daarmee aanvaardt zij het hoogleraarschap 'Onderwijskunde' aan de Universiteit van Amsterdam (Instituut voor de Lerarenopleiding).

Afscheidscollege prof. dr. J. Dronkers

Op donderdag 25 november om 15.00 uur houdt Jaap Dronkers zijn afscheidscollege bij de faculteit der pedagogische en onderwijskundige wetenschappen aan de Universiteit van Amsterdam.

Promoties

Rijksuniversiteit Groningen

Promovendus: mw. M.J. (Marie-José) Verkroost

Titel proefschrift: Onderzoekend handelen in het leeronderzoek

Promotor: prof. dr. J.J. Peters

Datum: 3 mei

Promovendus: mw. M. Winkel

Titel proefschrift: Niet-verbale Intelligentietest SON-R 2112-7. Constructie, plaatsbepaling en validering van deze algemene niet-verbale intelligentietest voor jonge kinderen

Promotor: prof. dr. W.K.B. Hofstee

Datum: 14 juni 1999

Universiteit Twente

Promovendus: M.B.M. (Marcel) de Croock

Titel proefschrift: The Transfer Paradox. Training Design for Troubleshooting Skills

Promotoren: prof. dr. J.M. Pieters en prof. dr. J.J.G. van Merriënboer (OUN)

Datum: 18 juni

Katholieke Universiteit Nijmegen

Promovendus: N.A. Broer

Titel proefschrift: Leren schema's maken. Een onderzoek naar de effecten van een lessenserie 'Schema's maken' voor de hoogste groep van het basisonderwijs

Promotoren: prof. dr. C.A.J. Aarnoutse en prof. dr. F.K. Kieviet (UL)

Copromotor: dr. J.F.J. van Leeuwen (KUN)

Datum: 7 juni 1999

Promovendus: J.G.G. Zuylen

Titel proefschrift: Professionalisering van docenten – een instrumentarium voor lesobservatieprojecten.

Promotor: prof. dr. P.R.J. Simons

Datum: 7 juni 1999

Promovendus: mw. W. Droop

Titel proefschrift: Effects of Linguistic and Cultural Diversity on the Development of Reading Comprehension. A Comparative Study of Dutch, Turkish and Moroccan Children Living in the Netherlands

Promotor: prof. dr. L. Verhoeven

Datum: 16 juni Nijmegen

Vrije Universiteit Amsterdam
Promovendus: J. Bokhorst
Titel proefschrift: Kinderen met teruggetrokken gedrag op de basisschool
Promotor: prof. dr. P.A. de Ruyter
Datum: 17 juni 1999

Universiteit van Amsterdam
Promovendus: mw. S. Kef
Titel proefschrift: Outlook on relations. Personal networks and psychosocial characteristics of visually impaired adolescents
Promotoren: prof. dr. van Gennep, prof. dr. Swarnborn
Copromotor: mw. dr. R. Habekothél
Datum: 22 juni 1999