



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

<https://hdl.handle.net/1874/214867>

# Inhoud

## THEMANUMMER OVER DE EVALUATIE VAN DE BASISVORMING

*Gastredactie: M.W.H. Peters-Sips, W.J. v.d. Linden en W.J.C.M. v.d. Grift*

Inleiding op het onderwerp van dit themanummer:  
de evaluatie van de basisvorming / **pag. 3**

*W.J. v.d. Linden, W.J.C.M. v.d. Grift en M.W.H. Peters-Sips*

Beleidsmatige voorgeschiedenis en opzet / **pag. 6**

*M.W.H. Peters-Sips*

Kwaliteit van het onderwijsleerproces / **pag. 20**

*H. v.d. Bergh, M.A. Zwarts en M.W.H. Peters-Sips*

Kwaliteit van het vakspecifieke aanbod / **pag. 40**

*M.W.H. Peters-Sips, M.A. Zwarts, H. v.d. Bergh en M.W.A. Schuurmans*

Kwaliteit van het vakoverstijgende aanbod:  
het voorkomen van de algemene vaardigheden / **pag. 53**

*M.A. Zwarts, H. v.d. Bergh en M.W.H. Peters-Sips*

Kwaliteit van de opbrengsten: Worden de kerndoelen gehaald? / **pag. 64**

*F. Kleintjes, M.A. Zwarts en M.W.H. Peters-Sips*

Kwaliteit van de opbrengsten: Zijn de prestaties bij Tekstbegrip Nederlands  
en Wiskunde vooruit gegaan? / **pag. 76**

*H. v.d. Bergh, R.J. Bosker en H.P. Brandsma*

## Reacties en discussiebijdragen

Onderwijs verdient compliment en stimulans / **pag. 91**

*Interview met W.J. Deetman door P.L.A. Helmholt*

Evaluatie basisvorming van grote kwaliteit / **pag. 93**

*Interview met J. Wallage door P.L.A. Helmholt*

De evaluatie van de basisvorming geëvalueerd / **pag. 97**

*R.J. Bosker*



Voor herhaling vatbaar / *pag. 103*

*G. Rijlaarsdam*

Oorzaken en aanpak van overladenheid van het operationele curriculum in de basisvorming / *pag. 110*

*K.Th. Boersma*

Vorm aan de basis, een terugblik op invoering en evaluatie van de basisvorming / *pag. 118*

*M. Verhagen en A. Harteveld*

Evaluatierapport Basisvorming uniek document / *pag. 124*

*M.A.M. Barth*

Ik houd van basisvorming / *pag. 129*

*W.G.J.M. v.d. Camp*

Regeringsstandpunt stelt teleur / *pag. 135*

*E.D.C.M. Lambrechts*

Hervorming van de basisvorming / *pag. 137*

*C.G.A. Cornielje en L. Middelhof*

Redactionele medewerkers aan het themanummer over de basisvorming / *pag. 141*

**Notities en Commentaren / *pag. 143***

## Inhoud

Overpeinzingen bij een fusie / **pag. 151**

*Van de redactie*

Ontwikkelingen in het Nederlandse onderwijsonderzoek / **pag. 155**

*B.P.M. Creemers*

### DE BESTE ARTIKELEN UIT 25 JAAR TIJDSCHRIFT VOOR ONDERWIJSRESEARCH

Metacognitie van lerenden in onderwijsleerprocessen / **pag. 169**

*M. Elshout-Mohr*

Training voor transfer van statistische vaardigheden: toepassing van een vier-componenten instructie-ontwerpmodel / **pag. 187**

*F.G.W.C. Paas en J.G. van Merriënboer*

Teaching effectiveness: an overview of studies / **pag. 199**

*J. Lowyck*

Effecten van klasseconsultatie op curriculumimplementatie en leerresultaten / **pag. 209**

*J. Snippe*

Aanpak van kwalificatieproblemen in organisaties: een onderzoek naar strategieën van opleidingskundigen / **pag. 219**

*F.J. van der Krogt, A.A. Vermulst en I.E.C. ter Woerds*

Relevante variabelen bij het doorverwijzen na de lagere school; een structureel model / **pag. 231**

*H. Blok en W.E. Saris*

Deregulering en de marktpositie van scholen / **pag. 249**

*H.W.C.H. van Amelsvoort en B. Witziers*

Onderwijs en maatschappelijke ongelijkheid, een overzicht van 25 jaar theorievorming en onderzoek in Nederland / **pag. 267**

*J.C.C. Rupp en A.A. Wesselingh*

**Lijst met huidige werkplek auteur / pag. 289**

## Inleiding op het onderwerp van dit themanummer: de evaluatie van de Basisvorming

Wim van der Linden, *Universiteit Twente*

Wim van de Grift, *Inspectie van het Onderwijs*

Marielle Peters-Sips, *Inspectie van het Onderwijs*

Van de resultaten van de evaluatie van de basisvorming is door de Inspectie in 1999 uitvoerig verslag gedaan in een algemeen rapport en 19 vakrapporten (Inspectie van het Onderwijs, 1999). Deze resultaten worden in dit themanummer niet herhaald.

Dit themanummer heeft twee andere doelen: In de eerste plaats wordt een aantal zaken betreffende de beleidsmatige context, het aanbod, het onderwijsleerproces en de opbrengst van de basisvorming verder uitgediept. In de tweede plaats wordt een aantal deskundigen uit wetenschap, politiek en praktijk aan het woord gelaten om hun oordeel te geven over de resultaten van de evaluatie van de basisvorming.

Alle bijdragen in dit themanummer op de laatste twee na zijn geschreven ruim vóórdat de beleidsvoornemens naar aanleiding van de evaluatie van de basisvorming bekend waren. Inmiddels (op 29 september 2000) zijn de beleidsvoornemens in de nota 'Ruimte voor kwaliteit in de basisvorming' door staatssecretaris Adelmund naar de Tweede Kamer gezonden.

### 1. CONTEXT, AANBOD, ONDERWIJSLEERPROCES EN OUTPUT VAN DE BASISVORMING

De evaluatie van de basisvorming is geen 'clean and tidy' wetenschappelijk experiment geweest. De evaluatie is ontstaan en uitgevoerd binnen een politiek-maatschappelijk krachtenveld, waarin vele spelers aan bal waren die ieder vanuit hun eigen optiek doelpunten gemaakt hebben en daarmee de ruimte bepaalden waarbinnen deze evaluatie kon worden uitgevoerd. De gevolgen van dit krachtenveld voor de evaluatie worden uitgediept door Marielle Peters-Sips. Zij heeft leiding gegeven aan de in dit themanummer beschreven evaluatieprojecten die door de Inspectie van het Onderwijs zelf werden uitgevoerd.

De kwaliteit van het onderwijsleerproces wordt beschreven door Huub van den Bergh, Michel Zwarts en Marielle Peters-Sips. Zwarts heeft als methodoloog van de Inspectie leiding gegeven aan de opzet van de evaluatieprojecten die door de Inspectie zelf werden uitgevoerd en aan de analyses van de gegevens uit deze projecten. Van den Bergh, werkzaam aan de Universiteit Utrecht, heeft een aantal belangrijke onderdelen van de statistische analyses van de gegevens uit de inspectieprojecten uitgevoerd. In hun bijdrage wordt verslag gedaan van de resultaten van de observaties van 66 inspecteurs in ruim 7000 lessen van 120 scholen voor voortgezet onderwijs. Het blijkt, dat deze lessen vaak hoog scoren op 'klassenklimaat', 'klassenmanagement' en 'instructie' (kwaliteitseisen die ook al vóór de basisvorming golden). Op 'vakdidactisch handelen' en 'actief leren' blijken deze lessen op de grens tussen sterk en zwak te scoren. In het 'differentiëren naar de capaciteiten van de leerlingen' blijkt in deze lessen zwakte te overheersen. De auteurs concluderen derhalve dat vijf jaar na de invoering van de basisvorming de twee vernieuwende aspecten van de basisvorming: 'actief leren' en 'differentiatie' en de vernieuwde vakdidactiek nog onvoldoende uit de verf komen. Met name blijken er grote verschillen tussen scholen en vooral tussen leraren binnen dezelfde scholen te bestaan. Juist deze verschillen brengen tot uitdrukking, dat er duidelijk mogelijkheden tot verbetering aanwezig zijn.

De kwaliteit van het vakspecifieke aanbod vormt het onderwerp van de bijdrage van Marielle Peters-Sips, Michel Zwarts, Huub van den Bergh en Lisette Schuurmans. Schuurmans, werkzaam bij de Inspectie, heeft de statistische analyses die nodig waren voor de evaluatie mee voorbereid en uitgevoerd. In hun bijdrage wordt voor 120 scholen voor voortgezet onderwijs voor alle vakken in de basisvorming nagegaan in hoeverre de kerndoelen in het aanbod zijn opgenomen. Geen enkele school blijkt aan de verplichting te voldoen om alle kerndoelen aan alle leerlingen aan te bieden. Gemiddeld genomen komt een leerling met iets minder dan twee van de drie kerndoelen in voldoende mate in aanraking om daar iets van op te steken. De vakken Muziek, Informatiekunde, Wiskunde, Frans, Engels, Lichamelijke opvoeding, Beeldende vorming en Duits komen daarbij beter uit de bus dan Techniek, Verzorging, Geschiedenis, Economie, Natuur-/scheikunde, Nederlands, Biologie en Aardrijkskunde.

De kwaliteit van het vakoverstijgende aanbod wordt behandeld in de bijdrage van Zwarts, Van den Bergh en Peters-Sips. Bij dit vakoverstijgende aanbod door scholen gaat het om het aanleren van algemene vaardigheden als: 'eenvoudig onderzoek doen', 'eigen standpunten kunnen verwoorden', 'samenwerken', 'relaties leggen met de beroepspraktijk', 'eigen mogelijkheden en interesses ontdekken' en 'criteria hanteren bij het zelf beoordelen van werkstukken'. Alleen 'samenwerken aan opdrachten' komt op meer dan de helft van de 120 geëvalueerde scholen in voldoende mate aan bod. 'Eigen standpunten verwoorden' en 'onderzoek doen' komen enigermate aan de orde. De overige algemene vaardigheden: 'relaties met de beroepspraktijk', 'eigen mogelijkheden en interesses ontdekken' en 'criteria hanteren' komen bij minder dan een kwart van de scholen in voldoende mate aan bod.

De vraag of de kerndoelen worden gehaald wordt behandeld in de bijdrage van Frans Kleintjes, Michel Zwarts en Marielle Peters-Sips. De eerstgenoemde is als methodoloog werkzaam bij het Cito en heeft een belangrijke bijdrage geleverd aan de opzet en verslaglegging van het onderzoek naar de leerlingresultaten, dat de Inspectie en het Cito gezamenlijk uitvoerden.

Voor het beantwoorden van de vraag naar de resultaten van de leerlingen is gebruik gemaakt van de door het Cito ontwikkelde toetsen voor de basisvorming. Voor de opbrengstmeting zijn de prestaties van bijna 65.000 leerlingen uit ongeveer 3.000 klassen gebruikt. Bij zes vakken (Engels, Nederlands, Duits, Biologie, Muziek en Natuur-/scheikunde) blijkt het percentage leerlingen dat boven het minimumniveau presteert hoger dan de gestelde grens van 90 procent. Bij vier vakken (Economie, Frans, Geschiedenis en Wiskunde) ligt dit dicht tegen de 90 procent aan. Bij vijf vakken (Aardrijkskunde, Techniek, Informatiekunde, Beeldende vorming en Verzorging) blijft dit meer dan 5 procent onder de 90 procent-grens. De vier laatstgenoemde vakken zijn met de basisvorming geïntroduceerd.

Jammer genoeg is het door de gebruikte techniek van steekproeftrekking in dit onderzoek uitsluitend mogelijk om per vak vast te stellen hoeveel leerlingen boven het minimumniveau presteren. Het is helaas niet mogelijk om inzicht te krijgen in het percentage leerlingen dat alle kerndoelen voor alle vakken beheerst.

De vraag of de prestaties bij tekstbegrip Nederlands en bij Wiskunde zijn vooruitgegaan wordt behandeld in de bijdrage van Huub van den Berg, Roel Bosker en Hennie Brandsma. De beide laatstgenoemden zijn vanuit het OCTO (Universiteit Twente) betrokken geweest bij het beschreven evaluatieonderzoek, dat op verzoek en met subsidie van de Inspectie via het NWO opgedragen is aan het GION (Universiteit Groningen) en het OCTO.

Deze drie onderzoekers rapporteren over de prestaties van drie cohorten leerlingen in het derde jaar van het voortgezet onderwijs. Het eerste cohort betrof leerlingen die in 1989, dus vóór de basisvorming, startten met het voortgezet onderwijs. Het tweede cohort betrof leerlingen die in 1993, op het moment dat de basisvorming was ingevoerd, begonnen met het voortgezet onderwijs. Het derde cohort betrof leerlingen die in 1995, twee jaar na de start van de basisvorming, aan het begin van hun voortgezet onderwijs stonden.

De gemiddelde prestaties bij tekstbegrip bleken in het cohort van 1993 lager te zijn dan in de cohorten van 1989 en 1995. De verschillen tussen de gemiddelde prestaties bij wiskunde bleken bij de drie cohorten verwaarloosbaar te zijn. Verder bleken de verschillen tussen de prestaties van de leerlingen in de loop van de tijd, tegen de verwachting in, eerder groter dan kleiner geworden te zijn.

Ondanks het feit dat in dit onderzoek gekampt werd met diverse steekproeftechnische en meettechnische problemen concluderen de onderzoekers dat de hoge verwachtingen van de basisvorming, algehele prestatieverhoging en verkleining van de verschillen tussen de prestaties van de leerlingen, niet bereikt zijn.

## 2. DISCUSSIEBIJDRAGEN EN REACTIES

Dit themanummer wordt afgesloten met oordelen uit de belangrijkste spelersgroepen in en rond de basisvorming.

Allereerst komt een tweetal oud-bewindslieden aan het woord die aan de wieg van de basisvorming hebben gestaan. Wim Deetman en Jacques Wallage waren respectievelijk als minister en staatssecretaris verantwoordelijk voor de voorbereiding van het traject respectievelijk de besluitvorming over de invoering van de basisvorming. Zij geven nu, vanuit een andere functie, namelijk burgemeester van respectievelijk Den Haag en Groningen, hun terugblik op de stand van zaken in de basisvorming.

Roel Bosker, Gert Rijlaarsdam en Kerst Boersma bieden vanuit het perspectief van de onderwijswetenschap een aantal lessen aan voor toekomstige evaluaties en gaan dieper in op enkele problemen, zoals het niet aanbieden van alle kerndoelen.

De schoolleiders Mia Verhagen en Anke Harteveld geven een terugblik op de invoering en evaluatie van de basisvorming. In hun bijdrage bespreken ze diverse knelpunten die scholen zowel bij de ontwikkeling van het onderwijsaanbod en de schoolontwikkeling, als bij de evaluatie van de basisvorming hebben ervaren.

Het laatste woord wordt gegeven aan de volksvertegenwoordigers. Woordvoerders van de vier grote politieke partijen zijn in de gelegenheid gesteld om terug te kijken en vooruit te zien. Marleen Barth en Wim van de Camp hebben de handschoen direct opgepakt. Ursi Lambrechts en Clemens Cornielje hebben hun bijdrage pas geleverd ná het verschijnen van de beleidsvoornemens in de nota 'Ruimte voor kwaliteit in de basisvorming'. Met de bijdragen van deze onderwijsspecialisten uit de Tweede Kamer besluiten we dit themanummer met het perspectief van enkele spelers uit hetzelfde politieke krachtenveld waarmee de basisvorming ook is begonnen.



# Beleidsmatige voorgeschiedenis en opzet

Marielle Peters-Sips, *Inspectie van het Onderwijs*

## ABSTRACT

*Traditionally, secondary education in the Netherlands is characterised by a wide range of school types for pupils from the age of 12 onwards. After a long political discussion an amendment to the Secondary Education Act introduced 'basisvorming' (basic secondary education) for all 12 to 15 year old pupils on 1 August 1993.*

*The 1992 amendment contains an evaluation statement, which is unique in Dutch educational legislation. It took several years to prepare the evaluation. The evaluation is implemented by the Inspectorate of education between 1997 and 1999.*

*The article outlines the evaluation questions and the design of the evaluation. It introduces other articles on aspects of the evaluation concerning the same issue.*

## 1. DE WETTELIJKE EVALUATIE-OPDRACHT

*"De doelstellingen van invoering van de basisvorming zijn ambitieus en de gekozen middelen en wegen zijn ten dele nieuw voor het Nederlandse onderwijs. Het ligt daarom in het voorplan in tweeërlei opzicht te evalueren, nl. gericht op de effecten van invoering van de basisvorming na afloop van de invoeringsperiode, en gericht op verschillende factoren en effecten daarvan tijdens het invoeringsproces." Uit dit citaat uit de Memorie van Toelichting (p. 33) bij het wetsvoorstel Basisvorming (TK, 1987-1988, 20 381, nr.3) blijkt de behoefte om de effecten van het nieuwe onderwijsprogramma te evalueren. Dat bij het ontwerp van een onderwijswet al rekening wordt gehouden met een evaluatie is tamelijk uniek. Hoewel in het voorstel nog vele wijzigingen worden aangebracht, is de evaluatieparagraaf behouden. De 'Wet van 27 mei 1992, houdende 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 het voortgezet onderwijs en de invoering van het voorbereidend beroepsonderwijs' (Staatsblad 1992, 270) bevat nog steeds een evaluatiebepaling. In de genoemde Wet is in artikel IV. Overgangsbepalingen basisvorming onder B. Evaluatie, de evaluatie-opdracht als volgt geformuleerd: "Onze Minister van Onderwijs en Wetenschappen draagt er zorg voor dat voor 1 augustus 1999 evaluatie plaatsvindt van de invoering van de basisvorming in het voortgezet onderwijs en van de invoering van kerndoelen in het basisonderwijs en voortgezet onderwijs."*

## 2. DE VOORBEREIDING VAN DE EVALUATIE DOOR CCE, PEB EN INSPECTIE

In de periode die aan het wetsvoorstel voorafging, werden al voorbereidingen getroffen voor een evaluatie van de vernieuwing van de eerste fase van het Voortgezet Onderwijs (Ministerie van Onderwijs en Wetenschappen, 1984). Op 12 februari 1985 werd door minister Deetman de Coördinatiecommissie Evaluatieplan Voortgezet Onderwijs (CCE) geïnstalleerd. Het

werk van deze, in de woorden van de instellingsbeschikking, *"onafhankelijke ministeriële adviescommissie, wier taak het is de informatieverzameling te coördineren ten behoeve van het verkrijgen van de evaluatiegegevens welke ten grondslag zullen dienen te liggen aan de toekomstige besluitvorming omtrent de structuur van het voortgezet onderwijs"*, overlapte het werkterrein van de Programma Advies Commissie Voortgezet Onderwijs Eerste Fase (PAC-VO-1), een bestuurscommissie van de Stichting voor Onderzoek van het Onderwijs (SVO), die meerjarenprogramma's en jaarlijkse onderzoeksprogramma's voor de eerste fase VO opstelde voor de Bestuursraad van SVO.

Het is de CCE die in haar vierde advies (1987) aan de Minister voorstelde de evaluatie van de basisvorming 'onafhankelijk' te laten uitvoeren. Hiermee ging minister Deetman akkoord.

Op basis van dit vierde advies werd de nota 'Evaluatie van de basisvorming. Kader voor het uitvoeringsplan' opgesteld (Peschar, 1988), die integraal in het zesde advies werd opgenomen. Daarnaast werd een analytische evaluatie van de basisvorming uitgevoerd door een viertal wetenschappers (Beem, 1988).

Op 1 oktober 1990 bracht de CCE haar tiende en laatste advies uit. Een van de aanbevelingen van de CCE aan staatssecretaris Wallage was om overleg over de uit te voeren evaluatie te laten plaatsvinden tussen SVO, CBS en de Inspectie *"met het oog op afstemming van beschrijvende informatie tussen schoolkenmerkonderzoek, statistiek, indicatorenontwikkeling en leerplangegevens"*<sup>1</sup>.

Vanaf 1991 maakte de evaluatie van de basisvorming, en dan met name de evaluatie van de voorgenomen invoering van de basisvorming (weer) deel uit van het reguliere SVO-programma ex artikel 6 van de Wet op de Onderwijsverzorging (WOV).

In 1993 stelde het Ministerie van Onderwijs en Wetenschappen (OenW) het Platform Evaluatie Basisvorming (PEB) in. Het was de bedoeling dat de betrokken instanties (SVO, Cito, PMB, Inspectie van het Onderwijs en OenW) gezamenlijk de evaluatie van de basisvorming in de steigers zouden zetten. Tijdens de vijfde vergadering van het PEB (september 1994) werd de concept-notitie 'Evaluatieprogramma basisvorming' goedgekeurd met daarbij de aantekening dat aan enkele punten nog nadere aandacht zou moeten worden besteed. Zo werd het idee gelanceerd om snel een beleidsanalytische studie te laten uitvoeren om de vraag te beantwoorden of de basisvorming op alle punten evalueerbaar was. Ook werd voorgesteld de notitie te actualiseren door de evaluatie-opdracht weer in lijn te brengen met de uitwerking van de basisvorming zoals door deze staatssecretaris Wallage in het wetsvoorstel was neergelegd.

Eind december 1994 ging een werkgroep uit het PEB aan de slag om een nieuwe versie van het evaluatieprogramma op te stellen. In de werkgroep bleek de SVO veel waarde te hechten aan het handhaven van het onderdeel (lange termijn) 'effecten', terwijl de Inspectie van mening was, dat dit onderdeel niet in deze evaluatie kan worden meegenomen, aangezien de tijd van invoering te kort nog was. De werkzaamheden van de werkgroep van het PEB resulteerden in april 1995 in een concept 'Evaluatieprogramma Basisvorming'. Met het opleveren van dit programma, kwam tevens een einde aan de activiteiten van het PEB.

Een van de eerste activiteiten die vervolgens wordt uitgevoerd was de aanbesteding via de SVO van het onderzoek 'beleidsanalyse basisvorming' (Karstanje, 1996).

Het evaluatieprogramma van het PEB bevat het voorstel dat het Ministerie van Onderwijs, Cultuur en Wetenschappen (OCenW) het definitieve evaluatieprogramma zou opstellen en in nauwe samenspraak met de Inspectie de uitvoering van het evaluatieprogramma zou coördineren in de periode 1995-1997. In de definitieve versie van de aanbestedingsnota werd de volgende taakverdeling voorgesteld: de Inspectie doet de eindevaluatie en de directie Voortgezet Onderwijs (VO) van OCenW coördineert de uitvoering van het evaluatieprogramma in

nauwe samenspraak met de Inspectie. In juni en juli 1995 vond hierover overleg plaats met staatssecretaris Netelenbos. De Staatssecretaris besloot dat de Inspectie verantwoordelijk is voor de inhoudelijke coördinatie én voor de voortgangsbewaking van de uitvoering van het evaluatieprogramma basisvorming. De Staatssecretaris koos voor de Inspectie, omdat zij van mening was dat deze taken onafhankelijk van het departement moesten worden uitgevoerd. De afspraak werd gemaakt dat de Inspectie in het voorjaar van 1997 een concept evaluatieplan basisvorming aan de Staatssecretaris zou voorleggen.

Het 'Plan voor de evaluatie van de basisvorming' (OCW-97-598) werd tenslotte in mei 1997 door staatssecretaris Netelenbos naar de Tweede Kamer gezonden. De Vaste Kamer Commissie voor onderwijs stelde 17 schriftelijke vragen, die de Staatssecretaris in augustus van dat jaar beantwoordde (TK, 1996-1997, 23 085, nr.15). Na dit schriftelijk overleg volgde geen mondeling debat meer.

### 3. DE DOELEN VAN DE BASISVORMING

Voor een adequate evaluatie is het van belang de te bereiken doelen te expliciteren. Met de verschuivingen die in de loop der jaren optreden in de doelen van de basisvorming, wijzigen ook de evaluatievragen. Allereerst wordt ingegaan op deze doelen.

#### *Doelen volgens het WRR-advies*

Na een jarenlange discussie over de inrichting van het onderwijs in de eerste leerjaren van het VO verzocht de regering in december 1983 de Wetenschappelijke Raad voor het Regeringsbeleid (WRR) om een advies uit te brengen over de wenselijke inhoud, duur en structuur van de basisvorming. Het advies 'Basisvorming in het onderwijs' (WRR, 1986) verscheen in januari 1986.

In zijn eerste aanbeveling omschreef de Raad de doelen van de basisvorming als volgt: "Het streven om te komen tot een algemene basisvorming in de eerste fase van het voortgezet onderwijs verdient blijvende ondersteuning omwille van:

- a. een verhoging van het onderwijspeil van alle leerlingen en een deelname van alle (toekomstige) leden van de samenleving aan het culturele erfgoed;
- b. de cultuurpolitieke noodzaak van co-existentie in een pluriforme samenleving;
- c. eisen van economische en technische modernisering;
- d. vermindering van ongelijkheid van kansen in het onderwijs door vermijding van een te vroege definitieve school- en/of beroepskeuze;
- e. verbetering en herwaardering van het beroepsonderwijs, dat in de huidige situatie tot een restcategorie in het onderwijs dreigt te worden."

De Raad was van mening dat de basisvorming in een periode van tien jaar gestalte zou moeten krijgen. Als beleidsinstrumenten bij de invoering noemde hij de invoering van eindtermen en examenprogramma's (op te stellen en te bewaken door een Centrale Examen Commissie) en de bewaking van de kwaliteit van de basisvorming door periodiek peilingsonderzoek. De Inspectie zou volgens de Raad de scholen vervolgens moeten stimuleren en adviseren en "in het jaarlijkse verslag van de Inspectie dienen de voortgang van de invoering van de basisvorming en de kwaliteit ervan uitdrukkelijk aan de orde te worden gesteld." (pag. 231).

De voorstellen van de Raad werden in korte tijd politiek gemeengoed en vormden de basis voor het concept-wetsvoorstel dat eind 1987 door minister Deetman bij de Tweede Kamer werd ingediend.

#### *Doelen volgens de Wet Basisvorming*

In augustus 1990 stuurde staatssecretaris Wallage het tweede (herziene) wetsontwerp naar de Tweede Kamer. In de uitwerking van het basisvormingsconcept waren enkele wijzigingen aangebracht:



- de eindtermen op twee niveaus waren vervangen door kerndoelen;
- voor bijzondere scholen werd de mogelijkheid om de kerndoelen door eigen kerndoelen te vervangen uitgebreid en werd het mogelijk om de landelijk ter beschikking te stellen toetsen door eigen toetsen te vervangen;
- aan het einde van het tweede leerjaar zou verplicht advies over het verder te volgen onderwijs moeten worden uitgebracht;
- de mogelijkheid om regels te stellen voor de overgang van de basisvorming naar het vervolgonderwijs (MBO, MAVO, HAVO en VWO) was vervallen.

Ook het invoeringstraject (via experimenten) werd gewijzigd. Toch kwamen de doelen van de basisvorming in grote lijnen overeen met die uit het advies van de WRR. In de considerans van de wet is het zo geformuleerd:

*“Alzo Wij in overweging genomen hebben, dat een algehele verhoging van het peil van het jeugdonderwijs noodzakelijk is;*

*dat het gewenst is het moment van de verplichte studie- en beroepskeuze uit te stellen;*

*dat het met oog daarop gewenst is het onderwijsprogramma in de eerste leerjaren van het voortgezet onderwijs te moderniseren en gedeeltelijk te harmoniseren tot een basisvorming voor alle leerlingen, en kerndoelen in te voeren voor het basisonderwijs en het voortgezet onderwijs;*

*dat het gewenst is het beroepsvoorbereidend karakter van het lager beroepsonderwijs te versterken;*

*dat het met het oog op het voorafgaande gewenst is te komen tot breder samengestelde en grotere scholen .....”*

In juni 1991 aanvaardde de Tweede Kamer de Wet Basisvorming, in mei 1992 deed de Eerste Kamer dat.

Volgens de considerans van de wet kent de Basisvorming dus de volgende vijf doelstellingen:

- de algehele verhoging van het peil van het jeugdonderwijs;
- uitstel van het moment van verplichte studie- en beroepskeuze;
- modernisering en gedeeltelijke harmonisering van het onderwijsprogramma in de eerste leerjaren VO;
- versterking van het beroepsvoorbereidend karakter van het LBO;
- streven naar breder samengestelde en grotere scholen.

Centraal in deze doelen staat een brede vorming voor alle leerlingen. Om dit te bereiken dient in de onderwijspraktijk het volgende gerealiseerd te worden (Karstanje, 1996):

- modernisering en gedeeltelijke harmonisering van het onderwijsprogramma;
- begeleiding bij studie- en beroepskeuze;
- ontplooiën van talenten.

De vorming van breder samengestelde en grotere scholen beschouwde de wetgever als een middel om de verschillende doelen te realiseren. Scholen hebben de primaire verantwoordelijkheid voor het leveren van goed onderwijs en hebben een grote mate van autonomie gekregen. Bij de invulling van hun verantwoordelijkheid zouden scholen zich rekenschap moeten geven van de doelen van de overheid en dienen ze hun onderwijs mede daarop te richten.

De kenmerken die van belang zijn voor goed onderwijs in de leerjaren waarin de basisvorming wordt gerealiseerd, zijn door de Inspectie ten behoeve van de evaluatie als volgt geoperationaliseerd (Ministerie van OCenW, 1997).

### *Brede vorming voor alle leerlingen*

De veranderende maatschappelijke omstandigheden maken een verhoging van het peil van het jeugdonderwijs nodig. De maatschappelijke ontwikkelingen stellen zwaardere eisen aan de leden van de samenleving. Daarom wordt het noodzakelijk geacht dat de leerlingen een brede vorming krijgen aangeboden en dat het gemiddelde opleidingsniveau op een hoger

niveau wordt gebracht. Met betrekking tot het einde van de basisvorming bestaat de 'verhoging/verbreding' uit vier elementen:

- inhoudelijk: modernisering van het onderwijsaanbod voor alle leerlingen en uitbreiding van het aanbod tot (ten minste) de vijftien vakken van de basisvorming met de bijbehorende kerndoelen.
- didactisch: samenhang in het onderwijsaanbod, aandacht voor ontwikkeling van vaardigheden, toepassingsgericht, het bevorderen van de zelfstandigheid van de leerlingen.
- het resultaat: de school hanteert ten minste de kerndoelen als door de leerlingen te bereiken doelstellingen aan het eind van de basisvorming.
- de aard van het resultaat bij de leerlingen: zowel kennis en inzicht als vaardigheden.

#### *Modernisering en gedeeltelijke harmonisering van het onderwijsprogramma*

Met ingang van de basisvorming is er gedurende de eerste jaren van het Voortgezet Onderwijs sprake van een gemeenschappelijk programma, het kerncurriculum, dat wordt bepaald door de voorgeschreven vakken en de kerndoelen. Voor de invoering van de basisvorming was het onderwijsprogramma in de onderbouw van het LBO anders dan in de onderbouw van HAVO en VWO. Alle leerlingen krijgen na de invoering in de eerste leerjaren van het VO een breder, actueler en samenhangend onderwijsaanbod. Toegevoegd zijn de vakken: Informatiekunde, Techniek en Verzorging en voor meer leerlingen een tweede Moderne vreemde taal. Bovendien zijn de volgende vakken uitgebreid: Geschiedenis en staatsinrichting, Economie en Aardrijkskunde.

Ook leiden de inhoud van de kerndoelen tot veranderingen in de vakinhouden. Zo ligt bij de Moderne vreemde talen een zwaar accent op het communicatieve element en bij Wiskunde op het toepassingsgerichte. Niet alleen de verschuiving in het inhoudelijk accent in de vakken, de versterking van de samenhang tussen de vakken, maar ook de algemene vaardigheidsdoelen vragen een aanpassing van de didactiek. Van de leerling wordt gevraagd enerzijds een grotere verantwoordelijkheid en zelfstandigheid voor het eigen leren, anderzijds kunnen samenwerken met anderen waarmee een beroep wordt gedaan op zijn/haar sociale vaardigheden. Er moet meer aandacht gegeven worden aan differentiatie om beter aan te sluiten bij het leervermogen van de leerling en zo de motivatie te vergroten. De kerndoelen dienen voor scholen tevens als referentiekader voor beoogde leerprestaties.

Tijdens de leerjaren waarin de basisvorming wordt onderwezen, moet een leerling zoveel kennis, inzicht en vaardigheden ontwikkelen, dat hij in staat is met succes een opleiding in de vervolgfase af te ronden.

De school moet een deel van de beschikbare onderwijstijd besteden aan het kerncurriculum en een ander deel -de vrije ruimte- kan besteed worden aan andere onderwijsactiviteiten. Deze andere onderwijsactiviteiten hebben enerzijds betrekking op de noodzakelijke programma-onderdelen van het aansluitende schooltype in de bovenbouw; anderzijds kan de school die vrije ruimte besteden aan extra uren basisvorming, Godsdienst, Levensbeschouwelijke vorming, Mentoruren, Begeleidingslessen, Studielessen of Oriëntatie op studie en beroep. Juist deze differentiatiemogelijkheid in aanvulling op het kerncurriculum, biedt leerlingen een goede ontplooiingskans.

Er zijn echter drie beperkingen ten aanzien van de 'vrije ruimte':

- in het Gymnasium moeten alle leerlingen in ieder geval Grieks en Latijn krijgen;
- in het HAVO en het VWO moet in de tijd van de vrije ruimte een derde Moderne vreemde taal worden gegeven;
- in het VBO moet een deel van de vrije ruimte worden gebruikt voor beroeps voorbereidende vakken.

De 'gedeeltelijke harmonisering' verwijst naar de verdeling tussen kerncurriculum en vrije ruimte: geharmoniseerd wat betreft het kerncurriculum, niet geharmoniseerd aangaande de invulling van de vrije ruimte.

### *Begeleiding bij studie- en beroepskeuze*

Een betere studie- en beroepskeuze in de eerste leerjaren van het VO, zo is de verwachting, tot het beter benutten van het aanwezige talent. Hierdoor zal de ongelijkheid van kansen in het onderwijs te verminderen. Een betere keuze wordt mogelijk gemaakt door vijf elementen (Karstanje, 1996):

- harmonisering van het onderwijsprogramma in de eerste leerjaren van alle schooltypen;
- niveauverhoging;
- algemene invoering van nieuwe vakken zoals Techniek en Verzorging;
- vorming brede scholengemeenschappen;
- advisering aan het eind van het tweede leerjaar.

Een belangrijk keuzemoment ligt aan het eind van het tweede leerjaar als de leerlingen een advies krijgen over verder te volgen onderwijs. Daarnaast moet de leerling de vaardigheid ontwikkelen om een verantwoorde keuze te maken. Studie- en beroepskeuzebegeleiding vormt een belangrijk onderdeel ter bevordering van een succesvolle onderwijsloopbaan.

### *Ontplooien van talenten*

In de Tweede nota van wijziging (TK, 1989-1990, 20 381, nr 17, p.24) wordt gesteld: "Naast de noodzaak van een algemene verhoging van het peil van het jeugdonderwijs is een lange periode van een gemeenschappelijk onderwijsprogramma nodig om het aanwezige talent beter te benutten en zo de ongelijkheid van kansen in het onderwijs te verminderen. De ongelijkheid bestaat in onderwijsdeelname naar type en duur, waarbij nog altijd factoren als sociaal-economische positie, sociaal-culturele positie en onderscheid jongens/meisjes een te grote rol spelen." Voorkomen moet worden dat talenten van meisjes en jongens eenzijdig worden ontwikkeld. In de basisvorming zijn nieuwe vakken ingevoerd die kunnen bijdragen aan het maken van een meer evenwichtige keuze. Leerlingen moeten worden voorbereid op het vervullen van een volwaardige rol in onze samenleving. Het gaat hier niet alleen om het verwerven van nieuwe kennis en vaardigheden, maar ook om het ontwikkelen van een besef van de verscheidenheid aan normen in de samenleving.

Karstanje geeft in zijn beleidsanalyse van de basisvorming aan dat zowel het gemeenschappelijke onderwijsprogramma als het uitstel van het moment van de verplichte studie- en beroepskeuze een sleutelrol vervullen in relatie tot benutting van talenten. In de evaluatie door de Inspectie wordt betrokken in hoeverre een gemeenschappelijk onderwijs programma en een zorgvuldige studie- en beroepskeuze bijdragen aan het ontplooien van talenten en het verminderen van kansongelijkheid.

## 4. VAN DOELEN NAAR EVALUATIEVRAGEN

### *Evaluatievragen op basis van het WRR-advies*

Zowel in het concept-wetsvoorstel van eind 1986 als, zoals gezegd, in het wetsontwerp dat in december 1997 door minister Deetman werd ingediend, is een evaluatie voorzien tijdens de invoeringsperiode van de basisvorming. Al in 1988 (Peschar, 1988) is een reeks evaluatievragen opgesteld, aangepast aan het nieuwe wetsontwerp (Tabel 1). Deze vragen sluiten aan bij de door de WRR geformuleerde doelen.

Dit geldt (ten dele) ook voor de vragen die ten grondslag liggen aan het 'Voortgezet Onderwijs Cohort Leerlingen 1989' (VOCL'89). De cohortstudies moeten antwoord geven op de volgende vragen (Peschar, 1988):

- 1 In hoeverre verandert het niveau van de leerprestaties als gevolg van de invoering van de basisvorming?
- 2 In hoeverre veranderen de samenhangen tussen leerlingkenmerken en leerlingprestaties als gevolg van de invoering van de basisvorming?
- 3 Welke verschijnselen doen zich voor op schoolniveau bij de invoering van de basisvorming?

Tabel 1. Evaluatievragen uit het 'Kader voor het uitvoeringsplan' (1988).

- 
- A Evaluatievragen over de verhoging van het onderwijspeil.*
- A1 Zijn, bij leerlingen met vergelijkbare kenmerken, de prestaties in de basisvorming hoger dan in het huidige onderwijs?
- A2 Worden geconstateerde verschillen in prestaties tussen Basisvorming en het huidige onderwijs bij leerlingen met vergelijkbare kenmerken in de loop van de tijd groter?
- B Vragen over de versterking van de gemeenschappelijke culturele basis.*
- B1 Is het inderdaad zo dat het leerpakket in de Basisvorming breder en homogener is samengesteld voor grotere groepen leerlingen dan in het huidige onderwijs?
- B2 Hebben leerlingen in de Basisvorming in vergelijking met soortgelijke leerlingen in het huidige onderwijs inderdaad een bredere culturele basis?
- B3 Wordt de gemeenschappelijke culturele basis in de Basisvorming in de loop van de tijd groter, in vergelijking met soortgelijke leerlingbestanden uit het huidige onderwijs?
- C Vragen met betrekking tot de technologische en sociaal-culturele ontwikkeling.*
- C1 Is het inderdaad zo dat het curriculum in de Basisvorming meer aandacht besteedt aan technologische en sociaal-culturele ontwikkeling?
- C2 Is de arbeidsmarktpositie van leerlingen na de Basisvorming beter dan die van vergelijkbare leerlingen uit het huidige onderwijs?
- D Vragen naar de vermindering van ongelijkheid van kansen door vermindering van te vroege studie- en beroepskeuze?*
- D1 Vindt de keuze voor studie en beroep na afsluiting van de Basisvorming op een later moment plaats dan in het huidige VO?
- D2 Neeft het boven onderzochte verschil in ongelijke kansen in studie- en beroepskeuze in de loop van de tijd meer af bij leerlingen met Basisvorming, dan bij soortgelijke leerlingen uit het huidige VO?
- E Vragen over de effecten op schoolorganisaties.*
- E1 Welke ervaringen worden opgedaan in de experimentele scholen bij de vormgeving van de Basisvorming?
- E2 Welke effecten heeft de invoering van de Basisvorming op de structuur van het onderwijs, zowel voor de organisatie als voor de leerlingen?
- F Vragen die specifiek betrekking hebben op ontwikkelingen binnen de Basisvorming.*
- F1 Welke resultaten worden er op de verschillende toetsen in de 14 vakken op de 2 niveaus in de Basisvorming gehaald?
- F2 Zijn er belangrijke verschillen aan te wijzen in het tempo waarmee de leerlingen de toetsen Basisvorming afleggen?
- F3 Worden de verschillen in resultaten op de toetsen van de Basisvorming tussen leerlinggroepen, scholen en organisatievormen in de loop van de tijd geringer?
- F4 Hoe functioneren de eindtermen? Kunnen scholen er mee werken: welke effecten heeft dit voor curriculum en leerlingen?
- F5 Welke verschillende lessentabellen worden gehanteerd en hoe wordt de vrije ruimte ingevuld: varianten en strategieën van scholen?
- F6 Welke verschillende varianten en routes worden in het systeem van Basisvorming gehanteerd? In welke omvang komen de verschillende leerlingstromen voor?
- 

De cohortonderzoeken maken deel uit van het meerjarige onderzoeksprogramma dat gericht is op lange-termijneffecten over de leerlingpopulatie en over de mogelijke oorzaken van die effecten. De lange-termijneffecten hebben betrekking op een viertal begrippen: rendement, onderwijspeil, efficiëntie en gelijkheid van kansen. Door de SVO was het begrip 'rendement' benadrukt naar aanleiding van een advies van de Sociaal Economische Raad (SER) over het VO. Analooq aan het VOCL'89 is in 1993 weer een leerlingenonderzoek en een schoolkenmerkenonderzoek uitgevoerd (Wijnen, Overtoom en Van Tilborg, 1993 en Van der Werf,



Lubbers en Kuiper, 1999). In beide cohortonderzoeken is de vraag gesteld naar de gelijkheid van onderwijskansen (Van der Werf, 2000). In aanvulling op de cohortstudies zijn verschillende dieptestudies uitgevoerd (onder andere door Creemers-van Wees, Akkermans & Brandsma, 1999).

#### *Evaluatievragen in het definitieve evaluatieplan*

Hiervoor is geschetst hoe de Inspectie uiteindelijk het 'Plan voor de evaluatie van de basisvorming' opstelde. In dit plan is het object van evaluatie verruimd tot "het onderwijs in de eerste leerjaren van het VO, inclusief de basisvorming".

Om adequaat te kunnen evalueren, is het van belang te expliciteren aan welke kwaliteitseisen het onderwijs moet voldoen. Voor de Inspectie vormden de hiervoor genoemde doelen die de overheid met het onderwijs in de eerste leerjaren van het VO nastreeft, een belangrijk referentiekader. Daarnaast spelen ook het perspectief van de wetenschap en het perspectief van de direct betrokkenen een rol bij het bepalen van de eisen die aan 'goed onderwijs' in de eerste leerjaren van het VO gesteld moeten worden (zie het artikel 'Kwaliteit van het Onderwijsleerproces' elders in dit themanummer).

In het evaluatieplan waarmee de Tweede Kamer instemde, zijn negen evaluatievragen onderscheiden (Tabel 2). Deze vragen zijn in te delen in vragen naar:

- 1 de onderwijsinhoudelijke vernieuwing: het onderwijsaanbod, het onderwijs-leerproces en de opbrengsten;
- 2 de schoolorganisatorische omgeving waarbinnen de onderwijsinhoudelijke vernieuwing van het primaire proces gestalte wordt gegeven.

## 5. VAN EVALUATIEVRAGEN NAAR EVALUATIEONDERZOEKEN

Naast de reeds, door onder andere de SVO en het Procesmanagement Basisvorming (PMB), geprogrammeerde onderzoeken initieerde de Inspectie evaluatieonderzoek ter beantwoording van de in het evaluatieplan gestelde vragen (Tabel 2). Het ging, kort gezegd, om de beoordeling van de kwaliteit van het aanbod, het onderwijsleerproces, de schoolorganisatie, de aansluiting met voorafgaand en volgend onderwijs en de opbrengsten, zoals deze door de basisvorming gerealiseerd zijn. Van belang is hier op te merken dat het primair om 'evaluatie' gaat en niet om onderzoek. Binnen evaluatie speelt onderzoek een belangrijke rol, maar uiteindelijk gaat het om een oordeel en niet om de toetsing van een hypothese of een beschrijving van

*Tabel 2. Evaluatievragen uit het 'Plan voor de evaluatie' (1997).*

- 
- 1 Voldoen het onderwijsaanbod en de vormgeving van het onderwijsleerproces aan de kenmerken van goed onderwijs?
  - 2 Is er een gunstig pedagogisch klimaat?
  - 3 Is er in de eerste leerjaren van het voortgezet onderwijs een passende begeleiding voor alle leerlingen?
  - 4 Is de keuze voor studie en beroep voldoende voorbereid?
  - 5 Zijn de opbrengsten aan het einde van de basisvorming voldoende hoog?
  - 6 Is het rendement in de eerste leerjaren van het VO goed?
  - 7 Is de schoolorganisatie gericht op het adequaat realiseren van het onderwijsaanbod, het onderwijsleerproces en van de opbrengsten?
  - 8 Hebben de scholen voorzieningen getroffen om tot een goede afstemming te komen tussen basisonderwijs en de eerste fase?
  - 9 Hebben de scholen voorzieningen getroffen om tot een goede afstemming te komen tussen het onderwijsprogramma in de leerjaren waarin de basisvorming wordt gerealiseerd en het onderwijs in de bovenbouw/tweede fase?
-

de stand van zaken. De cyclus bij een evaluatie omvat dan ook meer stappen dan bij onderzoek gebruikelijk is (Janssens, 1997). Het gaat om:

- het vaststellen van kwaliteitskenmerken;
- het operationaliseren van kwaliteitskenmerken;
- het vaststellen van de normen voor de beoordeling van de kwaliteit;
- het verzamelen en analyseren van de gegevens;
- het beoordelen van de realisatie van de kwaliteitskenmerken.

*Tabel 3. Kwaliteitskenmerken voor de organisatorische aspecten uit vier deelstudies (1999).*

*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 keuze voor de determinatie van een schoolsoort zijn goed geregeld en vinden zorgvuldig plaats
- V6 De keuze 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 groepeeringsvormen
- 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 zorgbreedtebeleid van de school is gericht op het bereiken van aantoonbare effecten van onderwijs-op-maat

Om antwoord te kunnen geven op de vragen met betrekking tot de volgende organisatorische aspecten van de basisvorming: 'Schoolbeleid', 'Pedagogisch klimaat', 'Voorbereiding op studie

en beroep' en 'Zorgbreedte', zijn vier -via deels overlappende steekproeven gekoppelde-evaluatieonderzoeken uitgevoerd. Van de resultaten is eerder verslag gedaan (Van den Bergh, Peters-Sips & Zwarts, 1999). Deze evaluaties leverden met oordelen over 25 kwaliteitskenmerken (Tabel 3) bijdragen aan het beantwoorden van de evaluatievragen 2, 3, 4 en 7 in Tabel 2.

Veruit de grootste inspanning leverde de Inspectie met de evaluatie van het 'Onderwijsleerproces in de basisvorming' (Inspectie van het Onderwijs, 1999 en Peters-Sips & Zwarts, 1999). Voor de kwaliteit van het onderwijsleerproces in de eerste leerjaren waren 14 kwaliteitskenmerken vastgesteld (Tabel 4). Het evaluatieonderzoek leverde bijdragen aan het beantwoorden van alle evaluatievragen in Tabel 2, met uitzondering van vraag 5. In verschillende artikelen in dit themanummer worden aspecten van dit onderzoek nader belicht. Het onderwijsleerproces in enge zin komt aan de orde in 'Kwaliteit van het onderwijsleerproces'. Het onderwijsaanbod wordt uiteengezet in 'Kwaliteit van het vakspecifieke aanbod' en 'Kwaliteit van het vakoverstijgende aanbod: de algemene vaardigheidsdoelen nader bekeken'.

Aan de opbrengsten van de basisvorming zijn afzonderlijke evaluatieonderzoeken gewijd. Al vanaf 1989 worden via de hiervoor genoemde cohort onderzoeken opbrengstgegevens verzameld (zie ook het artikel 'Kwaliteit van de opbrengsten: Zijn de prestaties bij Tekstbegrip Nederlands en Wiskunde vooruit gegaan?'). Daarnaast is in de jaren 1997-1999 onderzoek gedaan naar de mate waarin de leerlingen aan het einde van de basisvorming de kerndoelen beheersen. Van dit laatste onderzoek wordt verslag gedaan in het artikel 'Kwaliteit van de opbrengsten: Worden de kerndoelen gehaald?' elders in dit themanummer.

*Tabel 4. Kwaliteitskenmerken van goed onderwijs uit de 'Evaluatie van het Onderwijsleerproces' (1999).*

- 
- 1 Het feitelijk leerstofaanbod omvat de vakspecifieke kerndoelen
  - 2 Het feitelijk leerstofaanbod omvat de algemene vaardigheidsdoelen
  - 3 Het pedagogisch handelen van de leraar draagt bij aan een goed onderwijsleerproces
  - 4 De leraar hanteert directe instructie en klassenmanagement adequaat en organiseert de verwerking goed
  - 5 Het vakdidactisch handelen van de leraar is in overeenstemming met de algemene doelstelling van het vak en met de overige eisen die aan het vak worden gesteld
  - 6 De leraar bevordert in zijn/haar les een actief leerproces bij de leerlingen
  - 7 De leraar houdt rekening met relevante verschillen tussen leerlingen
  - 8 De school treft adequate maatregelen voor de doorwerking van het onderwijskundig beleid in het onderwijsleerproces
  - 9 De vaksectie voert actief beleid ten aanzien van de inhoud en de (vak)didactische aanpak van het vak
  - 10 De school/vestiging plaatst en begeleidt de leerlingen goed en geeft hen binnen de lessentabel voldoende gelegenheid de kerndoelen van de basisvorming te bereiken
  - 11 De vaksectie registreert en evalueert systematisch de prestaties en vorderingen van de leerlingen
  - 12 De school/vestiging volgt de prestaties van de leerlingen systematisch
  - 13 De school/vestiging besteedt zorg aan de aansluiting tussen basisvorming en basisonderwijs
  - 14 De school/vestiging besteedt zorg aan de aansluiting tussen basisvorming en bovenbouw
- 

## 6. OPZET VAN DE EVALUATIE VAN HET ONDERWIJSLEERPROCES

De Inspectie hanteert bij het uitvoeren van een evaluatie een zogenaamd toetsingskader. Ook bij het voor de evaluatie van de basisvorming belangrijke onderzoek, 'de evaluatie van het onderwijsleerproces', is dat het geval. In een toetsingskader wordt beschreven wat belangrijke kwaliteitskenmerken zijn, hoe deze waargenomen kunnen worden en welke beslisseregels gevolgd moeten worden om tot een voldoende of onvoldoende oordeel te komen. In het door

de Inspectie ontwikkelde toetsingskader voor de evaluatie van de basisvorming wordt dan ook onderscheid gemaakt tussen kwaliteitskenmerken (door de Inspectie meestal standaarden genoemd), indicatoren (waar op gelet moet worden) en normen (de beslisregels waarin de relatie tussen de standaarden en indicatoren is vastgelegd). Het hanteren van normen leidt tot oordelen over het al dan niet in voldoende mate gerealiseerd zijn van de standaarden (Janssens, 1997).

Voor het ontwikkelen van een adequaat toetsingskader moeten de kenmerken voor 'goed onderwijs' gedefinieerd worden. Het gaat dan om kenmerken die - anno 1997/1998 - van toepassing zijn op het onderwijs in de eerste leerjaren van het VO. De criteria voor goed onderwijs ontleent de Inspectie, zoals gezegd, aan het perspectief van de wetgever, het perspectief van de wetenschap en van deskundigen en aan het perspectief van de directbetrokkenen bij de eerste leerjaren van het VO. Analyse vanuit deze drie invalshoeken levert kwaliteitskenmerken op op het niveau van de les (algemeen geldende en vakspecifieke), op het niveau van het vak (voor het onderwijsaanbod, de vakdidactiek en het vaksectiebeleid) en op het niveau van de school (voor het onderwijsaanbod en het schoolbeleid).

*Tabel 5. Voorbeeld van een kwaliteitskenmerk met indicatoren en normering.*

**Kenmerk:**

De leraar houdt rekening met relevante verschillen tussen leerlingen

**Indicatoren:**

- 1 De leraar hanteert een gedifferentieerde onderwijsstrategie
- 2 De leraar houdt bij de instructie rekening met niveauverschillen tussen leerlingen
- 3 De leraar verstrekt - indien nodig - extra ondersteuning voor zwakke leerlingen
- 4 De leraar verstrekt - indien nodig - goede leerlingen extra opdrachten
- 5 De leraar houdt - indien nodig - qua taalniveau bij de instructie en verwerking rekening met anderstalige leerlingen.

**Normering:**

Voor de kwalificatie 'sterkte overheerst zwakte' moeten de inspecteurs ten minste indicator 1 óf 2 én één andere indicator positief hebben gescoord.

Om te kunnen beoordelen of de kwaliteitskenmerken voor een goed onderwijsleerproces (Tabel 4) gerealiseerd zijn, is een verdere operationalisatie noodzakelijk. Iedere standaard wordt uiteengelegd in (vijf tot tien) indicatoren. Deze indicatoren worden zo veel mogelijk in termen van observeerbare gedragingen geformuleerd en maken het mogelijk om na te gaan welke aspecten van de standaarden wel, niet of nauwelijks voorkomen. Omdat de observatoren hooggekwalificeerde onderwijsdeskundigen zijn (inspecteurs van het onderwijs) kunnen de indicatoren van hoog inferentieniveau zijn. De uitwerking van de eis met betrekking tot het rekening houden door de leraar met relevante verschillen tussen leerlingen is als voorbeeld van een algemene standaard op lesniveau opgenomen in Tabel 5. Een grotere mate van specificiteit is te zien bij de uitwerking van de standaard voor het vakdidactisch handelen van leraren voor de Wiskunde (Tabel 6).

Evalueren is het bepalen van de waarde. Als dit op een systematische wijze gebeurt, zijn expliciete normen nodig die aangeven op welke wijze een kwalificatie wordt toegekend aan een aangetroffen toestand. Voor het ontwikkelen van normen is een uitgewerkte procedure nodig. De gehanteerde normen kunnen alleen door een beroep op deze procedure verdedigd worden.



Tabel 6: *Kwaliteitskenmerk 'Vakdidactiek' voor Wiskunde.*

**Kenmerk:**

Het vakdidactisch handelen van de leraar is in overeenstemming met de algemene doelstelling van het vak en met de overige eisen die aan het vak worden gesteld

**Indicatoren:**

- 1 De leraar stimuleert leerlingen gegevens en uitkomsten kritisch te beoordelen
- 2 De leraar stimuleert leerlingen creatief te zijn in het bedenken van oplossingen
- 3 De leraar schenkt aandacht aan het proces van wiskundig generaliseren
- 4 De leraar schenkt aandacht aan de ontwikkeling van een goed wiskundig taalgebruik
- 5 De leraar vraagt van leerlingen een bepaalde omgang met problemen (analyseren, plan van aanpak maken, schriftelijk neerslaan)
- 6 De leraar controleert regelmatig of de leerlingen de essentie van het probleem begrepen hebben
- 7 De leraar geeft niet direct zelf de aanzetten voor de oplossing van problemen
- 8 De leraar bevordert de communicatie over wiskunde problemen tussen leerlingen onderling
- 9 De leraar maakt een goed gebruik van didactische hulpmiddelen (w.o. concreet materiaal, wiskundige modellen, computer, rekenapparatuur)
- 10 De leraar verwijst naar ander vakken/leergebieden of naar leerstof uit andere domeinen
- 11 De leraar motiveert de leerling voor het vak en wekt interesse voor de inhoud

**Normering:**

Voor de kwalificatie 'sterkte overheerst zwakte' moeten de inspecteurs tenminste de indicatoren 3, 5, 6 en 9 positief scoren en twee andere indicatoren uit 1, 2, 4, 7 en 8.

(De indicatoren 10 en 11 gelden voor alle vakken, maar zijn niet in de normering voor de vakdidactiek opgenomen. Zij behoren bij de standaard over de algemene vaardigheidsdoelen)

Voor de beoordeling van de kwaliteit voor de standaarden is een vierpuntsschaal gehanteerd:

- 1 vooral zwak;
- 2 zwakte overheerst sterkte (ook aangeduid met 'meer zwak dan sterk' of 'zwakte overheerst');
- 3 sterkte overheerst zwakte (ook aangeduid met 'meer sterk dan zwak' of 'sterkte overheerst');
- 4 vooral sterk.

Bij de kwalificatie 'vooral zwak' en 'zwakte overheerst sterkte' is er sprake van 'onvoldoende' realisatie van de standaard. Bij de andere twee kwalificaties is sprake van 'voldoende' realisatie.

In het artikel 'Kwaliteit van het onderwijsleerproces' elders in dit themanummer wordt uiteengezet hoe de geschetste opzet concreet is uitgewerkt en hoe aan de aan de evaluatie te stellen kwaliteitseisen is voldaan.

## 7. TEN SLOTTE

Aan de Wet Basisvorming zijn verschillende voorlopige voorstellen en ontwerpen voorafgegaan. Het verlengen van het eindpunt van de periode van gemeenschappelijk onderwijs van de leeftijd van 12 naar 15 jaar vormt de rode draad in de voorstellen. De wens om deze nieuwe onderwijsvorm na enige jaren te evalueren is een andere constante in de voorstellen vanaf de eerste helft van de tachtiger jaren.

In het evaluatieplan zoals dat medio 1997 door de Tweede Kamer is geaccepteerd, staan vooral de ontwikkelingen in de scholen centraal en niet de rol van de overheid. In vergelijking met de evaluatievragen uit 1988 zijn de vragen van 1997 algemener van aard, zijn ze niet gericht op een evaluatie van het gevoerde onderwijsbeleid en houden ze meer rekening met de inschatting dat de ontwikkeling van de basisvorming vijf jaar na de invoering niet voltooid kan zijn.

In het evaluatierapport van de Inspectie worden de ontwikkelingen in de scholen geplaatst tegen de context van de andere ontwikkelingen in het VO in de negentiger jaren. De complexe omgeving en de (rand-)voorwaarden waarbinnen de basisvorming zich ontwikkelde zijn, zoals blijkt uit het 'Plan voor de evaluatie van de basisvorming' dat staatssecretaris Netelembos naar de Tweede Kamer stuurde, geen onderdeel van het object van de evaluatie.

De tijd die scholen hebben gekregen om de basisvorming in te voeren, staat in schril contrast tot de tijd die nodig was om de Wet Basisvorming tot stand te brengen en de tijd die is besteed aan de voorbereiding van de evaluatie. Dit neemt niet weg dat geconstateerd kan worden dat van een diepgaande en uitvoerige voorbereiding van de evaluatie sprake is geweest, die uitzonderlijk is in de Nederlandse onderwijsontwikkeling en -wetgeving. Ook de uitvoering van de evaluatie is diepgaand en zorgvuldig geschied.

De wettelijk vastgelegde evaluatieverplichting heeft in zekere zin, ondanks de wisselende voorbereiders en regievoerders van de evaluatie, gezorgd voor een permanente en gecoördineerde aandacht van velen binnen en buiten het onderwijs voor de ontwikkelingen in de eerste leerjaren van het VO.

Een landelijke evaluatie van de evaluatie is aan te bevelen. De huidige vernieuwingen in het VO: de nieuwe tweede fase HAVO en VWO met haar profielen en studiehuis, het VMBO met het leerwegondersteunend onderwijs en het praktijkonderwijs zullen kunnen profiteren van een (verbeterde) versie voor de evaluatie, zoals deze werd uitgevoerd bij de basisvorming. Tot op heden is bij deze vernieuwingen geen sprake van een centraal gecoördineerde evaluatievoorbereiding zoals dat bij de basisvorming het geval was.

## NOOT

1. De ontwikkeling van onderwijsindicatoren staat op dat moment in de belangstelling (Scheerens, Stoel, Vermeulen & Pelgrum, 1988) en heeft de werkwijze van de Inspectie zeker beïnvloed.

## LITERATUUR

- Beem, A.L. (red.) (1988). *Weten Redeneren Raden. Een analytische evaluatie van de basisvorming*. Amsterdam: Swets & Zeilinger.
- Bergh, H. van den, Peters-Sips, M. & Zwarts, M. Deelstudies in het kader van de evaluatie van de Basisvorming. *Pedagogische Studiën 1999 (76)* pag. 224-236.
- Creemers-van Wees, L.M.C.M., Akkermans, L.M.W. & Brandsma, H.P. (1999). *Ontwikkelingen in de BAVO in de periode 1990-1998: Modernisering en harmonisering in de basisvorming*. SVO nr. 96108. Universiteit Twente: OCTO.
- Inspectie van het Onderwijs (1999). *Werk aan de basis. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het Onderwijs.
- Janssens, F.J.G. (1997). *Van onderzoek naar evaluatie. De methodologie van de Inspectie*. 's-Gravenhage: Delwel.
- Karstanje, P.N. (1996). *Beleidstheorie basisvorming: een proeve van reconstructie*. Amsterdam: SCO-Kohnstamm Instituut.
- Ministerie van Onderwijs en Wetenschappen (1984). *Evaluatieplan voor het Voortgezet Onderwijs*.
- Ministerie van Onderwijs, Cultuur en Wetenschappen (1997). *Plan voor de evaluatie van de basisvorming*. OCW-97-598.
- Peschar, J.L. (1988). *Kader voor het uitvoeringsplan Evaluatie basisvorming*. 's-Gravenhage: SVO.
- Peters-Sips, M. & Zwarts, M. (1999). *Onderwijsleerproces basisvorming. Evaluatierapport*. Utrecht: Inspectie van het Onderwijs.
- Scheerens, J., Stoel, W.G.R., Vermeulen, C.J.A.J. & Pelgrum, W.J. (1988). *De haalbaarheid van een onderwijsindicatorenstelsel voor het basis- en voortgezet onderwijs*. Universiteit Twente: OCTO.

Werf, M.P.C. van der (2000). (On)gelijkheid van onderwijskansen voor en na invoering van de basisvorming. In M. Peters-Sips, J. van der Linden & A. Wald (red.). *Verder werken aan de basis. Basisvorming bundelt krachten*. Den Haag: SDU.

Werf, M.P.C. van der, M.J. Lubbers en Kuyper, H. (1999). *Onderwijsresultaten van VOCL '89 en VOCL '93 leerlingen*. GION: Groningen.

Wetenschappelijke Raad voor het Regeringsbeleid (1986). *Basisvorming in het onderwijs*. 's-Gravenhage: Staatsuitgeverij.

Wijnen, W., Overtoom, R. & Tilborg, L. van (1993). *Verkenning van een aantal vraagstukken en de mogelijke inrichting van een cohortonderzoek voortgezet onderwijs*. 's-Gravenhage: SVO.

# Kwaliteit van het onderwijsleerproces

Huub van den Bergh, *Utrechts Instituut voor Linguïstiek (UU)*

Michel Zwarts, *Inspectie van het Onderwijs*

Marielle Peters-Sips, *Inspectie van het Onderwijs*

## ABSTRACT

*This paper deals with aspects of the design and implementation of the evaluation of the 'basisvorming'. So, testing framework, norms, samples of schools and lessons, execution, and reliability of scores are touched upon.*

*The quality of lessons in the first three years of secondary education was investigated. Inspectors (N=66) rated the quality of lessons (N= 7268) of schools (N = 120) for secondary education on five aspects: 'class climate', 'class management', 'didactics', 'active learning' and 'differentiation'.*

*It was shown that the mean ratings on 'class climate' and 'class management' can be interpreted as relatively high. The mean ratings on 'didactics', 'active learning' and 'differentiation' on the other hand can be interpreted as rather weak.*

*Large differences were obtained due to the subjects. Subjects like gymnastics and music were rated (relatively) positively, whereas subject matters like economics, the modern languages and geography were rated (relatively) poorly on all five aspects. However, there were large differences due to teachers and, to a lesser degree, to the school. As differences between schools are to be interpreted as aggregated differences between teachers, these results only highlight the importance of the teacher for the quality of education.*

## 1. INLEIDING

In het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer zijn de vragen die in het kader van de evaluatie van de basisvorming beantwoord moeten worden weergegeven. Bij die evaluatie gaat het, kort gezegd, om de beoordeling van de kwaliteit van het aanbod, het onderwijsleerproces, de schoolorganisatie, de aansluiting en de opbrengsten, zoals deze in de basisvorming gerealiseerd zijn (Peters-Sips & Zwarts, 1999).

In dit artikel staat het onderwijsleerproces centraal. Dit evaluatieproject is uitgevoerd door de Inspectie van het Onderwijs. Naast de lespraktijk zijn het onderwijsaanbod, het sectiebeleid, het schoolbeleid en de aansluiting met voorafgaand en erop volgend onderwijs geëvalueerd. In deze bijdrage concentreren wij ons op de kwaliteit van het onderwijsleerproces in engere zin: de kwaliteit van de lessen in de basisvorming.

Het is van belang op te merken dat het hier primair om *evaluatie* gaat en niet om onderzoek. Binnen evaluatie speelt onderzoek weliswaar een belangrijke rol, maar uiteindelijk gaat het om een *oordeel* en niet om de toetsing van een hypothese of (alleen) een beschrijving van de stand van zaken (Janssens, 1997). Vandaar dat, in vergelijking met onderzoek, een aantal stappen extra gezet moeten worden, om deze normering zo transparant mogelijk te maken (zie het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer).

Een belangrijk aspect van de evaluatie van de basisvorming is de kwaliteit van het onderwijsleerproces. Immers, een goed onderwijsleerproces is een noodzakelijke voorwaarde voor

goed onderwijs, en goede prestaties van leerlingen. Daarnaast is met de invoering van de basisvorming ook een duidelijke innovatie van het onderwijsleerproces beoogt. In de basisvorming is niet alleen het aantal vakken toegenomen, maar ook moet de leerling meer actief aan het onderwijsleerproces deelnemen.

In deze bijdrage wordt de kwaliteit van het onderwijsleerproces (in engere zin) geanalyseerd. Hierbij concentreren we ons op de vijf aspecten: 'Klassenklimaat', 'Klassenmanagement en instructie', 'Vakdidactiek', 'Actief leren', en 'Differentiatie'. De algemene vraag naar de kwaliteit van het gegeven onderwijs zoals dat in de basisvorming gegeven dient te worden, is daarbij uitgesplitst in zes deelvragen:

- Wat zijn de sterke en zwakke punten in het onderwijs?
- Hoe hangen de verschillende kwaliteitsaspecten samen?
- Hoe groot zijn de verschillen in kwaliteit tussen leraren en scholen?
- Wat zijn de kwaliteitsverschillen tussen de vakken?
- Wat zijn de kwaliteitsverschillen tussen anders samengestelde klassen?
- Is er verband tussen de kwaliteit van de lessen en de opbrengsten van het onderwijs?

## 2. METHODE

### *Constructie van het toetsingskader voor het onderwijsleerproces*

In overeenstemming met de gehanteerde Inspectiemethodologie (Janssens, 1997) heeft de Inspectie bij de evaluatie van het onderwijsleerproces een toetsingskader ontwikkeld, dat kwaliteitskenmerken (standaarden), indicatoren en normen bevat.

Bij het ontwerpen van het toetsingskader heeft de inspectie 'de' kenmerken voor 'goed' onderwijs in de eerste leerjaren van het voortgezet onderwijs gedefinieerd. De criteria voor 'goed onderwijs' heeft de Inspectie mede ontleend aan de wetgeving, de wetenschap en de onderwijspraktijk:

1. In de Wet op het Voortgezet Onderwijs (WVO), de op deze wet gebaseerde besluiten en het aan deze wet gekoppelde onderwijsbeleid komt het perspectief van de wetgever naar voren. Dit perspectief heeft vooral een rol gespeeld bij de kwaliteitskenmerken voor het onderwijsaanbod. Naast de wet, de besluiten en de vele beleidsstukken is ook gebruik gemaakt van de ten behoeve van de evaluatie van de basisvorming opgestelde beleidsanalyse (Karstanje, 1996). Verder zijn vertegenwoordigers van het Ministerie van Onderwijs, Cultuur en Wetenschappen (OCenW) en het toenmalige Procesmanagement Basisvorming (PMB) geraadpleegd;
2. Het perspectief van de wetenschap en van deskundigen is met name van belang geweest bij het bepalen van de kwaliteitskenmerken van de (vak)didactische aanpak en organisatie van het onderwijs. Hiervoor is een beroep gedaan op de betreffende literatuur, daarnaast zijn wetenschappers en onderwijsdeskundigen op verschillende momenten geraadpleegd over de concepten, zoals gebruikt en geoperationaliseerd in het toetsingskader. Daarnaast is het project begeleid door drie wetenschappelijke adviseurs;
3. Bij de verdere uitwerking van het toetsingskader heeft vooral het perspectief van de direct betrokkenen bij de eerste leerjaren van het VO een rol gespeeld. Deze zijn geraadpleegd over de concepten van het toetsingskader, en vertegenwoordigd in een subcommissie van het Onderwijsoverleg Primair Onderwijs Voortgezet Onderwijs (POVO) die het project begeleidde.

Voor de beoordeling van het onderwijsleerproces heeft dit geleid tot vijf standaarden, met elk een verschillend aantal indicatoren:

- 'Klassenklimaat': om leerlingen een omgeving te verschaffen waarin geleerd wordt die-



- nen leraren te beschikken over pedagogische kwaliteiten. Veel leerlingen in de leeftijd waarop zij deelnemen aan de basisvorming hebben behoefte aan enerzijds acceptatie/veiligheid structuur, anderzijds aan uitdaging. Als aan deze basisbehoeften is voldaan, draagt dat in positieve zin bij aan een goed onderwijsleerproces (zes indicatoren);
- 'Klassenmanagement en instructie': directe instructie, adequate leerstofverwerking en goed klassenmanagement zijn van invloed op de prestaties van de leerlingen. De indicatoren zijn onder te verdelen in drie categorieën: 'Structuur', met drie indicatoren (bij voorbeeld: De leraar zorgt door introductie en afsluiting dat doel van de les duidelijk zijn), 'Proces', met vijf indicatoren (bij voorbeeld: De klassikale uitleg van de leerstof is duidelijk), en 'Huiswerk', met twee indicatoren (bij voorbeeld: Het huiswerk wordt zorgvuldig gecontroleerd);
  - 'Vakdidactiek': om leerlingen zo ver te brengen dat zij de vakspecifieke doelen (kunnen) bereiken, is tijdens het onderwijsleerproces een bepaalde vakdidactische aanpak van leraren gewenst. Dit vakdidactische handelen kan per vak of vakkencluster verschillen (aantal indicatoren verschilt per vak; gemiddeld zijn er acht indicatoren);
  - 'Actief leren': dit is één van de belangrijkste didactische vernieuwingen van het onderwijsleerproces in de basisvorming. De essentie van 'actief leren' is dat leerlingen zelf hun kennis en vaardigheden opbouwen, met behulp van informatie die geïnterpreteerd, bewerkt en geassimileerd wordt in samenhang met aanwezige kennis en vaardigheden (zes indicatoren; zie Tabel 1);
  - 'Differentiatie': één van de uitgangspunten van de WVO is dat het onderwijs aangepast dient te worden aan de capaciteiten van de leerlingen. Behalve door niveau- en tempodifferentiatie kan ook gedifferentieerd worden op basis van interesses en cultuurverschillen. Het voornaamste middel blijft echter het tijdens het onderwijsleerproces in de klas aanpassen van instructie en verwerking van leerstof aan niveauverschillen die er zijn tussen leerlingen en groepen van leerlingen (zes indicatoren).

### *Normering*

De realisering van de standaarden is beoordeeld op een vierpuntschaal (zie Tabel 1 en het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer). Voor de grens tussen 'zwakte overheerst' en 'sterkte overheerst' is een norm ontwikkeld. Deze grens geeft het verschil aan tussen voldoende en onvoldoende realisering van de standaard. Voor het opstellen van deze normen is gebruik gemaakt van een aangepaste vorm van de Angoff-methode (Van der Linden, Meijer & Vos, 1997). Deze methode is ontwikkeld voor het bepalen van de grens tussen voldoende en onvoldoende bij toetsen. Beoordelaars dienen dan aan te geven welke opgaven marginale leerlingen wel of niet goed maken. Hieruit kan de grensscore bepaald worden. Naar analogie hiermee hebben inspecteurs voortgezet onderwijs bij het vaststellen van de normen voor de verschillende standaarden aangegeven welke indicatoren minimaal gerealiseerd moeten worden om van een voldoende realisering van de standaard sprake te kunnen zijn. Deze antwoorden zijn geanalyseerd en omgezet in een concept beoordelingsvoorschrift. Eenzelfde procedure is toegepast bij leden van de schoolleiding en leraren van een tiental scholen. Hoewel deze over het algemeen hogere eisen stelden dan de inspecteurs is besloten de Inspectienormen niet te wijzigen.

Zo moet, om bij de standaard 'Actief leren' tot een voldoende oordeel over een les te komen tenminste de eerste ('de leraar stimuleert dat leerlingen zelf initiatieven nemen'), de derde ('de leraar laat leerlingen zelfstandig werken') en twee andere indicatoren positief beoordeeld zijn (Tabel 1). Slechts in uitzonderlijke situaties kan op beredeneerde wijze van deze normering afgeweken worden.

De norm geeft, zoals gezegd, een voorschrift voor het onderscheid tussen 'zwakte overheerst' en 'sterkte overheerst'. Het onderscheid tussen de andere kwalificaties is aan de inspecteurs overgelaten.

Tabel 1: Een overzicht van de indicatoren en normering bij de standaard 'Actief leren'.

| Indicator   | B                            | sd(B) |
|---|------------------------------|-------|
| 1 De leraar stimuleert dat leerlingen zelf initiatieven nemen;  | .67                          | .006  |
| 2 De leraar zorgt voor een taakgerichte interactie tussen leerlingen;                                 | .32                          | .009  |
| 3 De leraar laat leerlingen zelfstandig werken;   | .33                          | .008  |
| 4 De leraar laat leerlingen hun eigen oplossingen/handelen/gedrag expliciteren en daarop reflecteren; | .30                          | .005  |
| 5 Bij werken aan taken/opdrachten stimuleert de leraar leeractiviteiten van leerlingen;               | .29                          | .006  |
| 6 De leerlingen nemen actief deel aan het onderwijsleerproces.  | .35                          | .011  |
| Oordeel standaard   | 1. Vooral zwak               |       |
|   | 2. Zwakte overheerst sterkte |       |
|   | 3. Sterkte overheerst zwakte |       |
|   | 4. Vooral sterk              |       |

Norm voor oordeel 'sterkte overheerst zwakte': inspecteurs moeten tenminste de indicatoren 1 en 3 en twee andere indicatoren positief scoren

### Validiteit

Om de (gezichts)validiteit van de standaarden en hun operationalisatie in indicatoren te optimaliseren, zijn een twintigtal wetenschappers en onderwijsdeskundigen met algemene dan wel vakinhoudelijke expertise geraadpleegd. Zij hebben aangegeven in hoeverre de kenmerken (standaarden) van belang zijn voor het bepalen van de kwaliteit van de basisvorming en in hoeverre alle aspecten van de basisvorming zijn verdisconteerd. Voor de indicatoren hebben zij aangegeven in hoeverre deze een adequate operationalisatie vormen van de standaarden en in hoeverre alle belangrijk aspecten van de standaard in de indicatoren zijn meegenomen. Verder is hen gevraagd suggesties te doen voor verbetering of aanvulling. Deze opmerkingen zijn verwerkt en hebben op een aantal punten geleid tot bijstelling.

De normen (zie bijvoorbeeld Tabel 1) dienen als vuistregel waar de inspecteur alleen beargumenteed van kan afwijken. Als dit veel voorkomt, is dit een aantasting van de interpreteerbaarheid van het oordeel op een standaard. Achteraf is dan ook nagegaan in hoeverre inspecteurs van het voorschrift zijn afgeweken en welke argumenten zij daarvoor hanteerden. In het algemeen geldt dat als al van het voorschrift is afgeweken dat meestal in het voordeel van de school gebeurde.

Van de vijf standaarden voor het onderwijsleerproces is voor 'Klassenmanagement en instructie' het meest frequent van het normeringsvoorschrift afgeweken (bijna in twintig procent van de observaties). De belangrijkste reden daarvoor is dat inspecteurs een positief oordeel gegeven hebben, ook al werd één van de indicatoren die deel uitmaakte van de norm ('de structuur van de les is duidelijk') niet positief gewaardeerd. Bij de andere standaarden is minder van de voorschriften afgeweken; in het algemeen minder dan tien procent, meestal in het voordeel van de school. Er mag geconcludeerd worden dat de inspecteurs het normeringsvoorschrift overwegend gevolgd hebben. Dit betekent niet alleen dat de voorschriften in de praktijk hanteerbaar gebleken zijn, maar ook dat de gehanteerde argumenten deugdelijk zijn, en er dus geen (aantoonbare) afbreuk gedaan wordt aan de validiteit van de oordelen.

De kwaliteit van de beoordelingsprocedure zoals die in de praktijk is gebruikt, blijkt (achteraf) o.a. uit de relatie tussen de oordelen en de indicatoren. Er dient een positieve relatie te zijn tussen de indicatoren en de oordelen op de betreffende standaard. Deze relatie is geanalyseerd met behulp van meerniveau-modellen, zodat rekening wordt gehouden met de hiërarchische structuur van de data. Voor de kwaliteitskenmerken van de les is een model met drie niveaus gebruikt, waarbij een onderscheid gemaakt wordt tussen lessen genest binnen docenten, en docenten genest binnen scholen. Over het geheel wordt vijftig of meer procent van de variatie in de oordelen verklaard door de scores op de indicatoren (zie Tabel 2). Een opvallend uitschieter zijn de oordelen over 'Vakdidactiek' bij lichamenlijk opvoeding, waar duidelijk minder variatie in de oordelen verklaard kan worden met behulp van de indicatoren (slechts 34 procent)<sup>1</sup>.

Tabel 2: Proportie verklaarde variatie in de oordelen op de standaarden door de indicatoren (aantal indicatoren tussen haakjes).

| Standaard                              | Proportie verklaarde variatie |
|--|-------------------------------|
| Klassenklimaat (6)                     | 0.62                          |
| Klassenmanagement (10)                 | 0.63                          |
| Actief leren (6)                       | 0.72                          |
| Differentiatie (6)                     | 0.60                          |
| Vakdidactiek (2):                      |                               |
| Nederlands (+ 10)                      | 0.63                          |
| Engels (+ 8)                           | 0.73                          |
| Frans (+ 8)                            | 0.62                          |
| Duits (+8)                             | 0.66                          |
| Wiskunde (+ 9)                         | 0.70                          |
| Natuur- en scheikunde (+ 6)            | 0.74                          |
| Biologie (+ 5)                         | 0.60                          |
| Techniek (+ 7)                         | 0.68                          |
| Aardrijkskunde (+ 8)                   | 0.51                          |
| Geschiedenis en staatsinrichting (+ 7) | 0.62                          |
| Economie (+ 5)                         | 0.70                          |
| Informatiekunde (+ 11)                 | 0.70                          |
| Verzorging (+ 7)                       | 0.66                          |
| Beeldende vorming (+ 9)                | 0.63                          |
| Lichamenlijk opvoeding (+ 6)           | 0.34                          |
| Muziek (+ 9)                           | 0.69                          |

Om de bijdragen van de afzonderlijke indicatoren na te gaan zijn ook de regressiecoëfficiënten van de indicatoren op de standaard berekend. Op één uitzondering na zijn alle 183 coëfficiënten zowel significant als substantieel. Er is derhalve geen reden om aan de kwaliteit van het observatie-instrumentarium te twijfelen. In tabel 1 zijn de regressiecoëfficiënten voor 'Actief leren opgenomen'. De meeste indicatoren hebben een gewicht rond de 0.30. Dit houdt in dat een positieve score op de indicator gemiddeld tot een stijging van het oordeel met drietiende leidt. Indicator 1 levert een bijna dubbele bijdrage. De laatste kolom van tabel 1 geeft de standaarddeviatie van de verdeling van de regressiecoëfficiënten binnen scholen. Voor indicator 1 betekent dit dat 95 procent van de regressiecoëfficiënten tussen de .55 en .79 ligt. Voor standaard 6 is die marge .23 tot .57. De verschillen tussen de scholen betekent dat de indicatoren in wisselende mate tot het oordeel hebben bijgedragen, maar op dat op (vrijwel) alle scholen het verband positief is.



*Steekproeven*

Ten behoeve van de evaluatie van het onderwijsleerproces is een getrapte steekproef getrokken, waarbij eerst scholen (eerste trap) en vervolgens lessen binnen scholen (tweede trap) getrokken zijn.

In de eerste trap heeft het CBS op verzoek van de Inspectie een representatieve steekproef van 120 vestigingen getrokken uit alle circa 1200 (officiële) vestigingen van scholen voor voortgezet onderwijs waar basisvorming wordt verzorgd. Ook de vestigingen vallend onder het Ministerie van Landbouw, Natuurbeheer en Visserij zijn in de populatie opgenomen. Bij het trekken van de steekproef is rekening gehouden met vier factoren: samenstelling qua opleidingsmogelijkheden, zelfstandigheid, grootte en regionale spreiding. Gekozen is voor een omvang van 120 vestigingen omdat daarmee voldoende nauwkeurig uitspraken over het geheel gedaan kunnen worden en verschillen van enige importantie, tussen bijvoorbeeld verschillende opleidingstypen, aangetoond kunnen worden. De steekproef is zelfwegend, zodat bij de analyses geen rekening gehouden hoefde te worden met de indeling in strata. De details zijn opgenomen in de technische rapportage (Peters-Sips & Zwartz, 1999).

In feite zijn twee parallelle steekproeven getrokken. De tweede steekproef fungeerde als reserve steekproef. Bij zeven steekproefelementen is het element uit de eerste steekproef vervangen door een element uit de tweede steekproef: drie vestigingen zijn door hun 'reserve vestiging' vervangen om praktische redenen, en op vier vestigingen kon de evaluatie niet uitgevoerd worden door problemen in de school of omdat er (bij nader inzien) geen (volledige) basisvorming gegeven bleek te worden op de getrokken vestiging.

De representativiteit van de resulterende steekproef is nagegaan voor de stratumindeling, de schoolgrootte en de regionale spreiding. Daarbij bleken geen significante afwijkingen.

In de tweede trap zijn lessen binnen scholen getrokken. Volgens de opzet moesten per school tussen de 30 en 90 lessen bezocht worden. Praktische omstandigheden maakten het moeilijk om deze lessen aselekt te kiezen. Die omstandigheden betroffen met name de beperkte tijd waarin de lessen bezocht moesten worden (één week per school), de beschikbaarheid van (vak)inspecteurs, de samenstelling van de school en het gehanteerde rooster. Om een zekere representativiteit te waarborgen is bij de keuze van de lessen een aantal regels gehanteerd:

- per vak proportioneel met het aantal uren in de adviesurentabel;
- verdeling over de leerjaren overeenkomstig de duur van de basisvorming per vak;
- zo veel mogelijk gespreid over verschillende leraren;
- zo veel mogelijk gespreid over klassen van verschillende samenstelling (naar opleidingstype).

De representativiteit van de gekozen lessen is nagegaan voor de verdeling over de vakken, de verdeling over klassen en de verdeling over leerjaren. Voor vakken is er een significante afwijking van de verdeling zoals die uit de adviesurentabel voortvloeit. De belangrijkste oorzaak is dat lessen in 'kleine' vakken (vakken met een gering aantal lessen) op vrijwel elke school zijn bezocht en daardoor disproportioneel veel aan bod zijn gekomen. De afwijking is echter niet dusdanig dat het nodig is om door selectie of weging de verdeling volgens de adviesurentabel dichter te benaderen.

Met teams van acht à twaalf inspecteurs zijn schoolbezoeken afgelegd van drie à vier dagen bij elk van de 120 vestigingen van scholen voor voortgezet onderwijs. Iedere inspecteur heeft zich daarbij in het bijzonder op één, twee of maximaal drie vakken uit de basisvorming geconcentreerd. Gemiddeld zijn ongeveer 60 lessen per vestiging bezocht (zie Tabel 3). In totaal zijn tijdens de schoolbezoeken 7338 lessen bijgewoond. Daarvan zijn achteraf de gegevens van negen lessen incompleet gebleken. Voor de 'keuzevakken' Dans, Drama en Fries zijn op de 120 bezochte scholen in totaal slechts 51 lessen bijgewoond. Voor deze vakken zijn om een goed beeld te krijgen ook lessen in andere dan de 120 scholen bezocht. Deze lesobserva-

ties zijn in dit artikel buiten beschouwing gebleven. Vandaar dat in Tabel 3 een overzicht gegeven wordt van 7268 lessen, uitgesplitst naar vak en klassensamenstelling.

Tabel 3: Aantal geobserveerde lessen per standaard per vak naar klassensamenstelling.

| Vakken                 | Klassensamenstelling |     |      |       |      |       |       |      | Breed | Totaal |
|------------------------|----------------------|-----|------|-------|------|-------|-------|------|-------|--------|
|                        | (I)VBO               | VBO | MAVO | MAVO/ | HAVO | MAVO/ | HAVO/ | VWO  |       |        |
| Aardrijkskunde         | 38                   | 38  | 79   | 41    | 37   | 58    | 53    | 72   | 26    | 442    |
| Beeldende vorming      | 73                   | 48  | 104  | 54    | 37   | 55    | 49    | 98   | 37    | 555    |
| Biologie               | 52                   | 42  | 70   | 35    | 23   | 44    | 36    | 76   | 27    | 405    |
| Duits                  | 16                   | 73  | 99   | 44    | 62   | 22    | 69    | 46   | 16    | 447    |
| Economie               | 41                   | 37  | 91   | 26    | 32   | 20    | 23    | 30   | 10    | 310    |
| Engels                 | 75                   | 78  | 101  | 39    | 51   | 60    | 80    | 83   | 32    | 599    |
| Frans                  | 0                    | 4   | 73   | 24    | 43   | 59    | 55    | 86   | 27    | 371    |
| Geschiedenis           | 34                   | 44  | 89   | 40    | 41   | 45    | 59    | 81   | 20    | 453    |
| Informatiekunde        | 25                   | 32  | 35   | 30    | 10   | 33    | 13    | 47   | 23    | 248    |
| Lichamelijke opvoeding | 62                   | 63  | 95   | 40    | 33   | 67    | 58    | 79   | 30    | 527    |
| Muziek                 | 38                   | 21  | 58   | 22    | 14   | 42    | 41    | 49   | 19    | 304    |
| Natuur en scheikunde   | 57                   | 53  | 109  | 32    | 53   | 31    | 90    | 42   | 12    | 479    |
| Nederlands             | 88                   | 96  | 125  | 66    | 57   | 77    | 88    | 112  | 33    | 742    |
| Techniek               | 60                   | 33  | 57   | 29    | 17   | 51    | 30    | 64   | 21    | 362    |
| Verzorging             | 45                   | 38  | 59   | 32    | 12   | 38    | 29    | 40   | 17    | 310    |
| Wiskunde               | 81                   | 80  | 109  | 57    | 51   | 69    | 82    | 106  | 34    | 668    |
| Totaal                 | 785                  | 780 | 1351 | 611   | 573  | 771   | 854   | 1111 | 383   | 7268   |

Voor de andere dan de drie genoemde vakken varieert het aantal lessen van bijna 750 voor Nederlands tot iets meer dan 300 voor muziek. Voor klassen van verschillende samenstelling varieert het aantal waargenomen lessen van 1351 voor MAVO-klassen tot 573 voor HAVO-klassen. De afwijking van de verdeling volgens de adviesuren en de klassensamenstelling liggen vooral bij Frans, Duits en techniek in (I)VBO- en VBO-klassen. In IVBO-klassen worden Frans en Duits veel minder gegeven dan in AVO- en VWO-klassen en wordt meer aandacht besteed aan techniek. In VBO-klassen wordt minder Frans gegeven. Voor de talen ligt de verklaring in de mogelijkheid om maar twee moderne vreemde talen te kiezen in MAVO en (I)VBO en het vrijstellingenbeleid in het (I)VBO.

#### *Overeenstemming en betrouwbaarheid*

Betrouwbaarheid omvat een aantal aspecten. In de evaluatie is de beoordelaarsovereenstemming nagegaan, alsmede de betrouwbaarheid van de gegevens die verzameld zijn tijdens de lesbezoeken.

De beoordelaarsovereenstemming is nagegaan door duobezoeken: twee inspecteurs die gemeenschappelijk een les bezoeken en onafhankelijk van elkaar hun bevindingen vastleggen. Voor de indicatoren is het percentage overeenstemming 86 procent. Voor de oordelen is de gewogen overeenstemming, rekening houdend met de ordening van de kwalificaties 92 procent (Heuvelmans & Sanders, 1993). Beide waarden zijn zonder meer acceptabel.

De betrouwbaarheid van de oordelen op de standaarden voor de lessen is bepaald aan de hand van een generaliseerbaarheidstudie. In een generaliseerbaarheidsstudie wordt de score op een

variabele uiteengelegd in één of meer ware scores en één of meer versturende factoren. Als de varianties van de onderscheiden componenten bekend zijn, kan een generaliseerbaarheidscoëfficiënt berekend worden. Voor een model met één versturende factor komt dat overeen met coëfficiënt alfa (Veldhuijzen, Godebeld & Sanders, 1993).

Bij de lesobservaties vallen de scores uiteen in effecten van school, leraar, les, vak, en observator (inspecteur). De eerste vier effecten kunnen in dit verband beschouwd worden als ware scores en de effecten van observator als de versturende factor. De effecten van school, leraar, les en observator zijn (gemodelleerd als) random effecten, terwijl het effect van vak hier beschouwd is als een vast effect. Een complicerende factor is dat lessen genest zijn binnen leraren en leraren genest zijn binnen combinaties van scholen en inspecteurs (observatoren). Er is derhalve sprake van een hiërarchisch model (Goldstein, 1995). De schattingen van de variantiecomponenten zelf staan in Tabel 4.

Tabel 4: Variantiecomponenten en betrouwbaarheid ( $\rho$ ) per standaard.

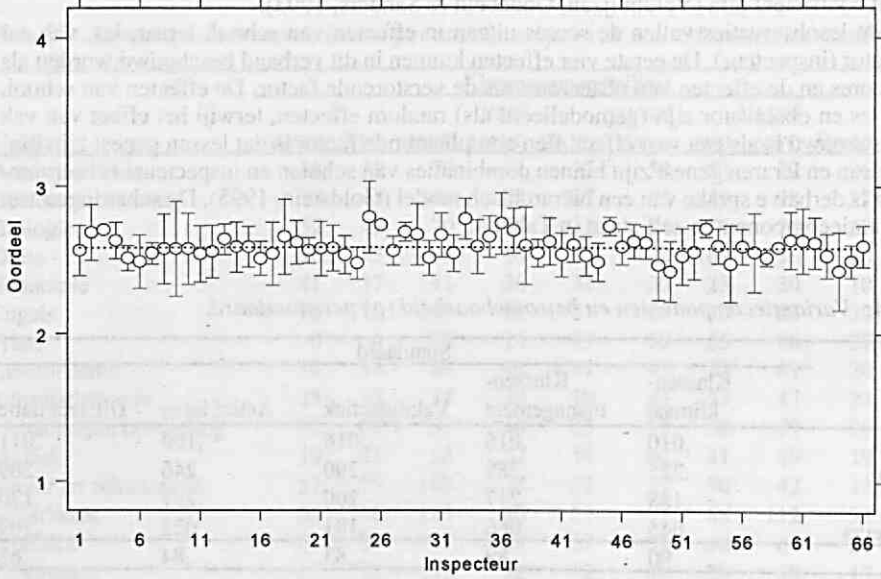
| Factor      | Standaard           |                        |              |              |                |
|-------------|---------------------|------------------------|--------------|--------------|----------------|
|             | Klassen-<br>klimaat | Klassen-<br>management | Vakdidactiek | Actief leren | Differentiatie |
| Scholen     | .010                | .016                   | .016         | .100         | .011           |
| Leraren     | .225                | .289                   | .290         | .246         | .269           |
| Lessen      | .188                | .217                   | .200         | .251         | .170           |
| Inspecteurs | .045                | .066                   | .101         | .022         | .209           |
| $\rho$      | .90                 | .89                    | .83          | .84          | .68            |

In Tabel 4 kan de som van de eerste drie componenten beschouwd worden als ware variantie en de variantie tussen inspecteurs vormt de foutenvariantie. De betrouwbaarheid van het oordeel van één inspecteur die één les bezoekt ( $\rho$ ) is de ratio van de ware en de som van de ware variantie en foutenvariantie. De betrouwbaarheden zijn heel acceptabel, met uitzondering van 'Differentiatie' zijn ze alle groter dan .80. Het is niet onmogelijk dat juist bij 'Differentiatie' interpretatieverschillen tussen inspecteurs een rol spelen. Dit houdt in dat uitspraken over deze standaard met de nodige terughoudendheid gedaan moeten worden. We kunnen derhalve concluderen dat de rangorde van scholen, docenten en lessen voor vier van de vijf standaarden in hoge mate vergelijkbaar zijn over inspecteurs.

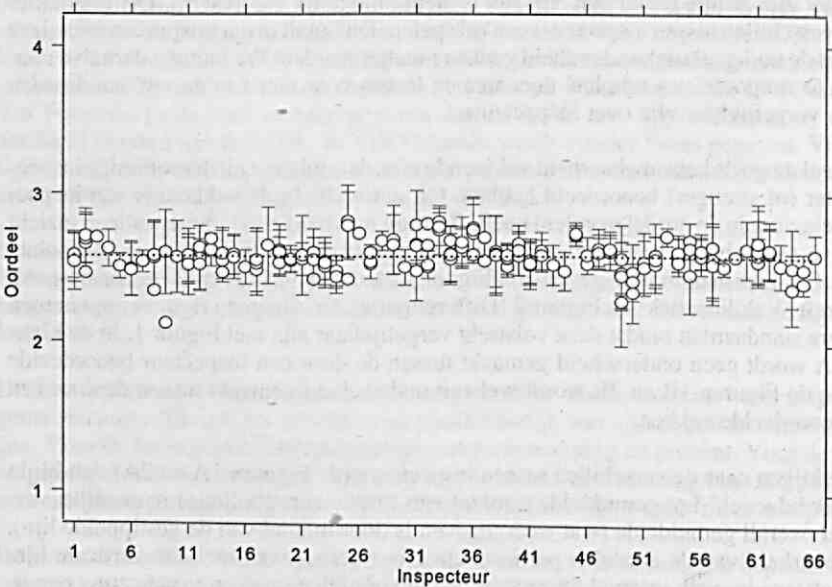
Over het geheel moge de betrouwbaarheid voldoende zijn, dat sluit niet uit dat sommige inspecteurs coulanter (of strenger) beoordeeld hebben. Of, dat wellicht de vakkennis van inspecteurs tot vertekening in de oordelen geleid heeft. Betrouwbaarheid geeft immers alleen inzicht in de relatieve samenhang; betrouwbaarheid is niet geformuleerd in termen van 'absolute samenhang'. Om hier inzicht in te geven zijn Figuur 1 (A en B) en 2 (A en B) geconstrueerd. Figuur 1 betreft 'Vakdidactiek' en Figuur 2 'Differentiatie'. Er zijn geen figuren opgenomen voor de andere standaarden omdat deze volstrekt vergelijkbaar zijn met Figuur 1. In de Figuren 1A en 2A wordt geen onderscheid gemaakt tussen de door een inspecteur beoordeelde vakken, en in de Figuren 1B en 2B wordt wel een onderscheid gemaakt tussen de door een inspecteur beoordeelde vakken.

Wanneer we kijken naar de verschillen tussen inspecteurs (de Figuren 1A en 2A) dan blijkt dat voor 'Vakdidactiek' het gemiddelde oordeel van inspecteurs (bolletje) nauwelijks verschilt van het overall gemiddelde (wat weergegeven is door middel van de gestippelde lijn). De betrouwbaarheid van de oordelen per inspecteur is weergegeven met een verticale lijn. Deze lijn geeft aan in welk interval 85 procent van de oordelen van een inspecteur - condi-

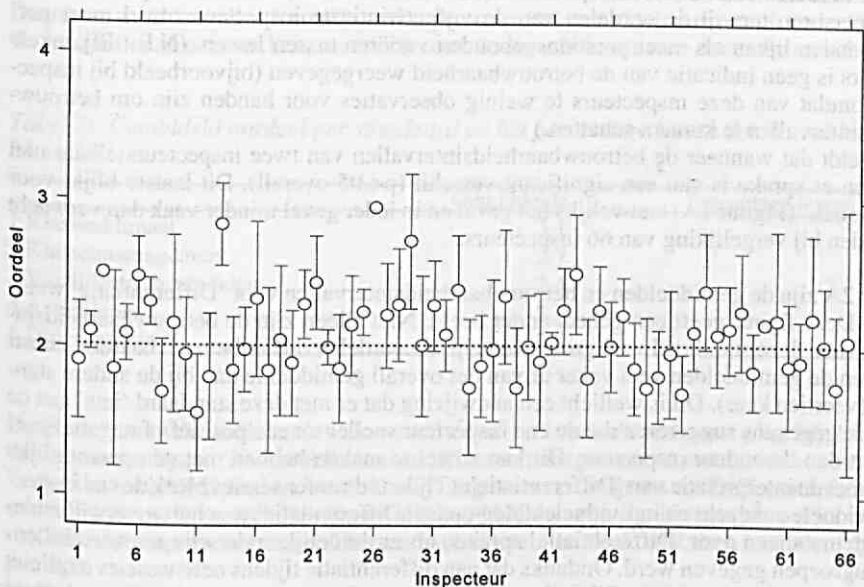
*Figuur 1A: Gemiddeld oordeel per inspecteur (bolletje) en individuele betrouwbaarheidsintervallen (verticale lijn) voor 'Vakdidactiek'.*



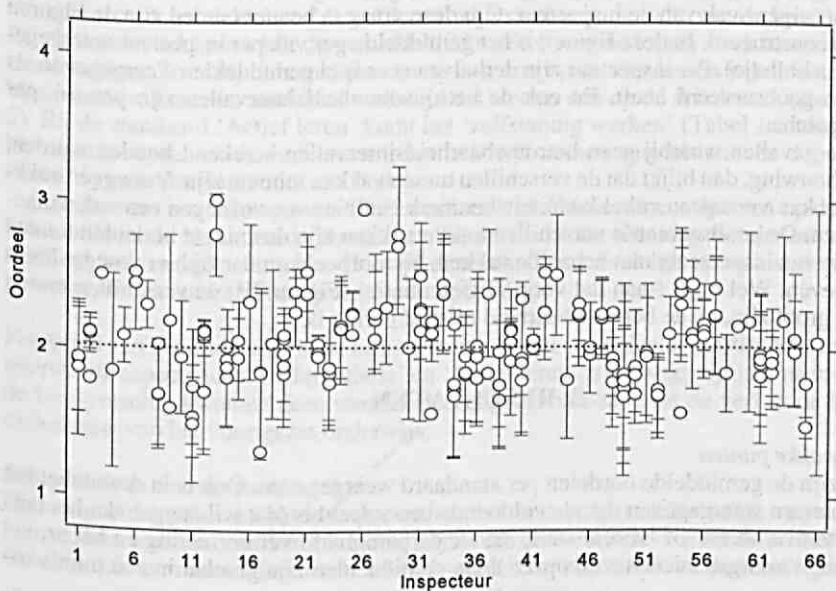
*Figuur 1B: Gemiddeld oordeel per inspecteur per vak (bolletje) en individuele betrouwbaarheidsintervallen (verticale lijn) voor 'Vakdidactiek'.*



Figuur 2A: Gemiddeld oordeel per inspecteur (bolletje) en individuele betrouwbaarheidsintervallen (verticale lijn) voor 'Differentiatie'.



Figuur 2B: Gemiddeld oordeel per inspecteur per vak (bolletje) en individuele betrouwbaarheidsintervallen (verticale lijn) voor 'Differentiatie'.





tioneel op verschillen tussen vakken, lessen, docenten en scholen - verwacht mogen worden. Hoe langer de lijn, hoe groter het betrouwbaarheidsinterval, hoe onbetrouwbaarder de oordelen. De oordelen van de eerste inspecteur vertonen nauwelijks kenmerken van deze beoordeelende persoon, terwijl de oordelen van de vijftiengste inspecteur zowel meer persoonsgebonden lijken als meer persoonsgebonden variëren tussen lessen. (N.B.: Bij enkele inspecteurs is geen indicatie van de betrouwbaarheid weergegeven (bijvoorbeeld bij inspecteur 26) omdat van deze inspecteurs te weinig observaties voor handen zijn om betrouwbaarheidsintervallen te kunnen schatten.)

Ook geldt dat wanneer de betrouwbaarheidsintervallen van twee inspecteurs elkaar niet overlappen er sprake is van een significant verschil ( $p < .05$  overall). Dit laatste blijkt voor 'Vakdidactiek' (Figuur 1A) nauwelijks het geval en in ieder geval minder vaak dan verwacht mag worden bij vergelijking van 66 inspecteurs.

In Figuur 2A zijn de gemiddelden en betrouwbaarheidsintervallen voor 'Differentiatie' weergegeven. Deze figuur geeft een geheel ander beeld. Niet alleen zijn de betrouwbaarheidsintervallen voor deze standaard veel groter - en zijn de oordelen dus onbetrouwbaarder - maar ook wijken de gemiddelden veel vaker af van het overall gemiddelde dan bij de andere standaarden (veertien keer). Dit is wellicht een aanwijzing dat er met deze standaard 'iets' aan de hand is; de gegevens suggereren dat de ene inspecteur sneller tot een positief of negatief oordeel komt dan de andere inspecteur. Dit kan zowel te maken hebben met een persoonlijke 'bias' of met de interpretatie van 'Differentiatie'. (Tijdens de oefensessies bleek de ene inspecteur individuele aandacht en individuele uitleg onder 'Differentiatie' te scharen, terwijl andere inspecteurs alleen over 'Differentiatie' spraken als er duidelijk onderwijs aan verschillende niveaugroepen gegeven werd. Ondanks dat aan differentiatie tijdens oefensessies expliciet aandacht besteed is, is het klaarblijkelijk niet geheel gelukt de inspecteurs voor het onderdeel differentiatie op één lijn te krijgen.)

Een tweede vraag die in dit verband aangeroerd moet worden, betreft de verschillen tussen vakken. De inspecteurs hebben over het algemeen lessen van twee à drie vakken beoordeeld. Het is mogelijk dat het ene vak strenger is beoordeeld dan het andere. Met name kan dit gelden voor het 'eigen' vak van de inspecteur. Om deze vraag te beantwoorden zijn de Figuren 1B en 2B geconstrueerd. In deze figuren is het gemiddelde per vak per inspecteur weergegeven (door een bolletje). Per inspecteur zijn derhalve net zoveel gemiddelden weergegeven als deze vakken geobserveerd heeft. En ook de betrouwbaarheidsintervallen zijn per vak per inspecteur geschat.

Laten we gevallen waarbij geen betrouwbaarheidsintervallen berekend konden worden, buiten beschouwing, dan blijkt dat de verschillen tussen vakken miniem zijn. Voor geen enkele inspecteur kan voor geen enkel kwaliteitskenmerk een bias voor of tegen een vak aangetoond worden. De geobserveerde verschillen tussen vakken zijn dus niet te herleiden tot een preoccupatie van inspecteurs met bepaalde vakken, bijvoorbeeld omdat zij hier vroeger les in hebben gegeven. Wel blijft staan dat voor 'Differentiatie' (Figuur 2B) de verschillen tussen inspecteurs groot zijn, en de betrouwbaarheid relatief gering is.

### 3. RESULTATEN

#### *Sterke en zwakke punten*

In Tabel 5 zijn de gemiddelde oordelen per standaard weergegeven. Ook is in deze tabel het percentage lessen weergegeven dat als voldoende beoordeeld is (dat wil zeggen dat het oordeel 'sterkte overheerst' of 'vooral sterk' is; zie de paragraaf over normering en het artikel 'Beleidsmatige voorgeschiedenis en opzet'). De gemiddelden zijn geschat in een multivari-

aat hiërarchisch model met lessen, leraren en scholen als random componenten en vak als vast effect<sup>2</sup>. Dat wil zeggen: voor alle combinaties van vak en standaard wordt een gemiddelde geschat, en op alle drie de niveaus wordt de covariantiematrix tussen de standaarden geschat. Daardoor wordt zowel informatie verkregen over het gemiddelde niveau (oordeel), als over de verschillen tussen vakken, maar ook over de verschillen tussen docenten en scholen.

Tabel 5: Gemiddeld oordeel per standaard en het percentage lessen dat als voldoende beoordeeld is.

|                          | Gemiddelde | Percentage lessen voldoende |
|--------------------------|------------|-----------------------------|
| 1 Klassenklimaat         | 3.1        | 80                          |
| 2 Klassenmanagement      | 2.7        | 64                          |
| 3 Vakdidactisch handelen | 2.4        | 53                          |
| 4 Actief leren           | 2.4        | 53                          |
| 5 Differentiatie         | 2.0        | 26                          |

De grens tussen de oordelen 'zwakte overheerst' en 'sterkte overheerst', ligt bij 2.5. Gemiddeld genomen liggen 'Klassenklimaat' en 'Klassenmanagement en instructie' boven die grens, 'Vakdidactisch handelen' en 'Actief leren' ongeveer op die grens en 'Differentiatie' duidelijk onder die grens. De verschillen tussen alle gemiddelden zijn significant ( $\chi^2_1 > 156$ ;  $p < .001$ ), met uitzondering van het verschil tussen 'Actief leren' en 'Vakdidactisch handelen' ( $\chi^2_1 = 0.2$ ;  $p < .65$ ).

De verschillen tussen de vijf standaarden blijken nog duidelijker uit de percentages lessen die als voldoende beoordeeld zijn ('sterkte overheerst' en 'vooral sterk'). 'Klassenklimaat' is een (relatief) sterk punt: bij een ruime meerderheid van de lessen gaan leraren en leerlingen positief met elkaar om en zorgt de leraar voor een goed werkklimaat. Het 'Klassenmanagement' komt op de tweede plaats. In bijna twee op de drie lessen is de uitleg van de leerstof duidelijk, wordt de leerstof regelmatig gecontroleerd en zijn de leerlingen taakgericht bezig. 'Vakdidactisch handelen' en 'Actief leren' zijn slechts in de helft van de lessen voldoende. Bij 'Vakdidactiek' verschilt de gemiddelde kwaliteit van de lessen sterk tussen de verschillende vakken. Zo is bij lichamelijke opvoeding drie op de vier lessen vakdidactisch in orde; bij economie, Frans en aardrijkskunde zijn dat slechts één op de drie lessen (zie ook: Figuur 5). Bij de standaard 'Actief leren' komt het 'zelfstandig werken' (Tabel 1: indicator 3) over het algemeen wel uit de verf, maar komt het 'zelf nemen van initiatieven door leerlingen' (Tabel 1: indicator 1) minder aan bod.

Zondermeer een zwak punt is 'Differentiatie': slechts iets meer dan een op de drie lessen is als voldoende beoordeeld. Vooral het 'bij de instructie rekening houden met niveauverschillen' is hier onder de maat. Hoewel deze standaard duidelijk het minst betrouwbaar is gemeten, is deze uitkomst toch een reden voor zorg.

Het beeld van het onderwijs in de basisvorming is niet onverdeeld gunstig. Met name de vernieuwende aspecten als 'Differentiatie' en 'Actief leren' zijn zwak; vijf jaar na invoering van de basisvorming kwamen deze standaarden nog onvoldoende uit de verf in de lessen in de onderbouw van het voortgezet onderwijs.

#### Samenhang tussen de standaarden

De vraag of bepaalde kwaliteitsaspecten van het onderwijs onafhankelijk van andere voorkomen, kan worden beantwoord door analyse van de correlaties tussen de scores op de standaarden. Die correlaties kunnen berekend worden op elk van de onderscheiden niveaus. In

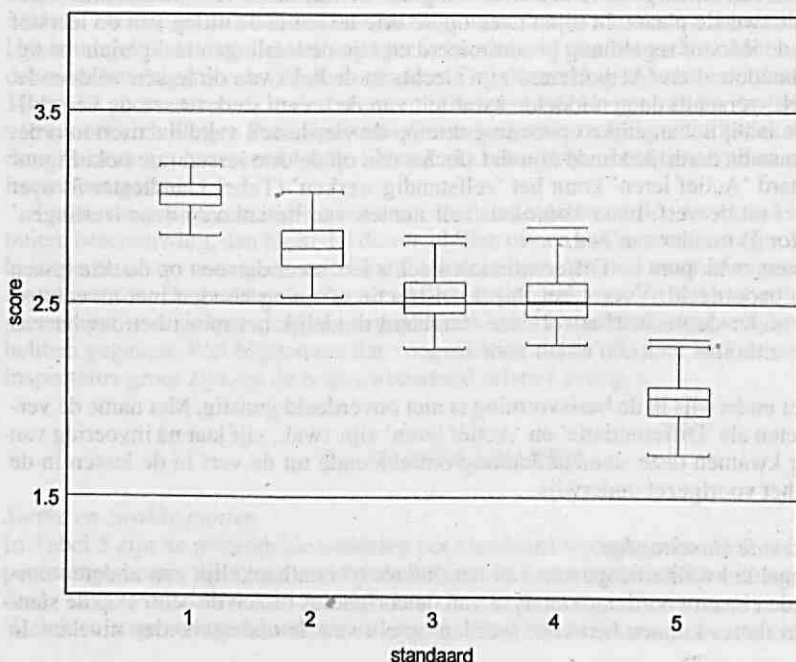
Tabel 6 staan de correlaties tussen de geschatte schoolgemiddelden (onder de diagonaal) en de docentgemiddelden (boven de diagonaal).

Tabel 6: Correlaties tussen standaarden op docent- (boven de diagonaal) en schoolniveau (onder de diagonaal).

| Standaard           | Standaard |     |     |     |     |
|---------------------|-----------|-----|-----|-----|-----|
|                     | 1         | 2   | 3   | 4   | 5   |
| 1 Klassenklimaat    | -         | .60 | .50 | .50 | .33 |
| 2 Klassenmanagement | .62       | -   | .58 | .55 | .37 |
| 3 Vakdidactiek      | .50       | .56 | -   | .54 | .33 |
| 4 Actief leren      | .50       | .54 | .55 | -   | .41 |
| 5 Differentiatie    | .31       | .36 | .33 | .42 | -   |

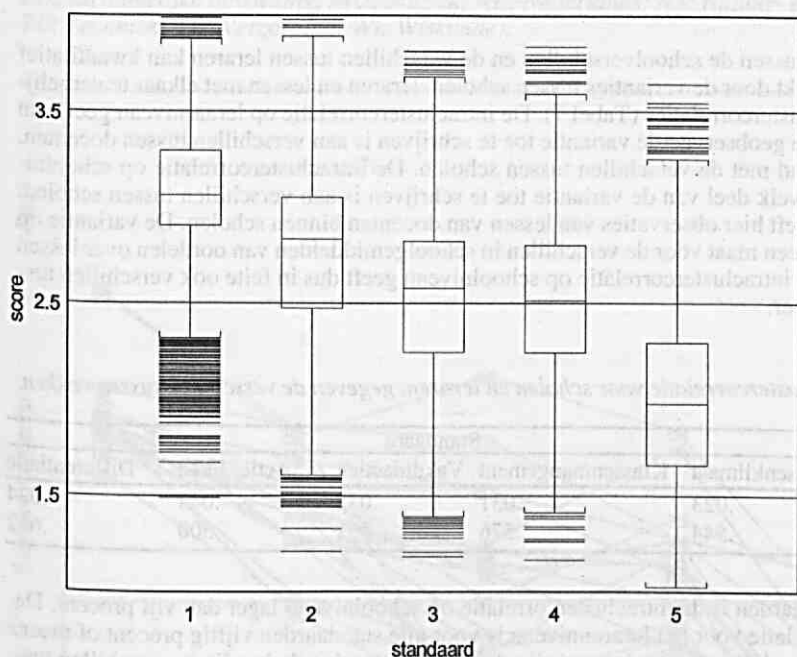
Uit de tabel blijkt dat zowel op school- als docentniveau vier van de vijf standaarden redelijk goed samenhangen: alle correlaties zijn hoger dan 0.50. Dit betekent dat als op een school, of bij een leraar één van de eerste vier standaarden voldoende is, er een redelijke kans is dat één of meer van de andere aspecten ook op orde zijn. Maar anderzijds kan - gezien de hoogte van de correlatiecoëfficiënten - ook geconcludeerd worden dat het hier om vier duidelijk onderscheidbare kenmerken van het onderwijs gaat; bij voorbeeld: een docent die relatief hoog

Figuur 3: 'Box and whisker' - plot van de verdeling van de geschatte schoolgemiddelden (y-as) per standaard (x-as; 1: 'Klassenklimaat'; 2: 'Klassenmanagement en instructie'; 3: 'Vakdidactiek'; 4: 'Actief leren'; 5: 'Differentiatie').





Figuur 4: 'Box and whisker' plot van de verdeling van de leergemiddelden (y-as) per standaard (x-as; 1: 'Klassenklimaat'; 2: 'Klassenmanagement en instructie'; 3: 'Vakdidactiek'; 4: 'Actief leren'; 5: 'Differentiatie').



beoordeeld is op 'Vakdidactiek' is over het algemeen ook relatief goed beoordeeld op 'Klassenklimaat', maar let wel, er zijn ook docenten die (relatief) hoog op de ene, maar (relatief) laag op de andere standaard beoordeeld zijn.

De standaard 'Differentiatie' vormt de uitzondering. Zowel op het niveau van de school als dat van de docent is er een betrekkelijk geringe samenhang tussen deze standaard en de andere kwaliteitskenmerken van onderwijzen.

#### Kwaliteitsverschillen tussen leraren en scholen

Dat er grote verschillen zijn in de kwaliteit van lessen is een gegeven. Interessanter is de vraag naar de verschillen tussen leraren en scholen. Deze verschillen zijn als 'box and whisker'-plots weergegeven in Figuur 3 en Figuur 4 (zie ook Tabel 4). In de figuren omvat de 'box' in vijftig procent van de waarnemingen; de mediaan is weergegeven als een lijn die de 'box' in tweeën deelt. De 'whiskers' worden weergegeven door de rechte haken onder en boven de 'box'. Ze staan op anderhalve keer het interkwartielbereik en omvatten 71 procent van de observaties. Uitschieters zijn als losse lijntjes weergegeven (elke uitschieter is door één lijn gerepresenteerd). De horizontale lijnen in beide figuren geven de grenzen aan van respectievelijk 'vooral zwak' met 'zwakte overheerst', van 'zwakte overheerst' met 'sterkte overheerst' en van 'sterkte overheerst' met 'vooral sterk'.

Figuur 3 laat de verschillen tussen scholen zien. Deze zijn klein - zeker in vergelijking met de verschillen tussen leraren (zie Figuur 4). Bovendien blijkt uit de figuur dat er betrekkelijk weinig uitschieters zijn. Zo liggen alle schoolgemiddelden voor 'Klassenklimaat' en 'Klassenmanagement' in het gebied 'meer sterk dan zwak' en liggen de schoolgemiddelden voor 'Differentiatie' alle in het gebied 'meer zwak dan sterk'.

Zoals uit Figuur 4 blijkt, zijn de verschillen tussen de leraren aanzienlijk. Voor alle standaarden komen zowel leraren voor die gemiddeld 'vooral sterk' zijn als leraren die 'vooral zwak' zijn. Voor 'Klassenklimaat' is de variatie het kleinst, hoewel er veel uitschieters naar beneden zijn.

De verhouding tussen de schoolverschillen en de verschillen tussen leraren kan kwantitatief worden uitgedrukt door de varianties tussen scholen, leraren en lessen met elkaar te vergelijken door intraclustercorrelaties (Tabel 7). De intraclustercorrelatie op leraarniveau geeft aan welk deel van de geobserveerde variantie toe te schrijven is aan verschillen tussen docenten, rekening houdend met de verschillen tussen scholen. De intraclustercorrelatie op schoolniveau geeft aan welk deel van de variantie toe te schrijven is aan verschillen tussen scholen. Let wel, het betreft hier observaties van lessen van docenten binnen scholen. De variantie op schoolniveau is een maat voor de verschillen in schoolgemiddelden van oordelen over lessen en docenten. De intraclustercorrelatie op schoolniveau geeft dus in feite ook verschillen tussen docenten weer.

Tabel 7: Intraclustercorrelatie voor scholen en leraren, gegeven de verschillen tussen vakken.

| Niveau | Standaard      |                   |              |              |                |
|--------|----------------|-------------------|--------------|--------------|----------------|
|        | Klassenklimaat | Klassenmanagement | Vakdidactiek | Actief leren | Differentiatie |
| School | .023           | .031              | .032         | .043         | .024           |
| Leraar | .544           | .576              | .591         | .500         | .612           |

Voor alle standaarden is de intraclustercorrelatie op schoolniveau lager dan vijf procent. De intraclustercorrelatie voor het lerarenniveau is voor alle standaarden vijftig procent of meer. De kwaliteitsverschillen tussen leraren zijn dus veel groter dan de kwaliteitsverschillen tussen scholen. Dit betekent aan de ene kant dat er binnen elke school zeer grote verschillen zijn in de kwaliteit van de docenten en aan de andere kant dat de verschillen tussen scholen (relatief) gering zijn; de marges om de kwaliteit van het onderwijs te verhogen zijn derhalve op leraarniveau veel groter dan op schoolniveau.

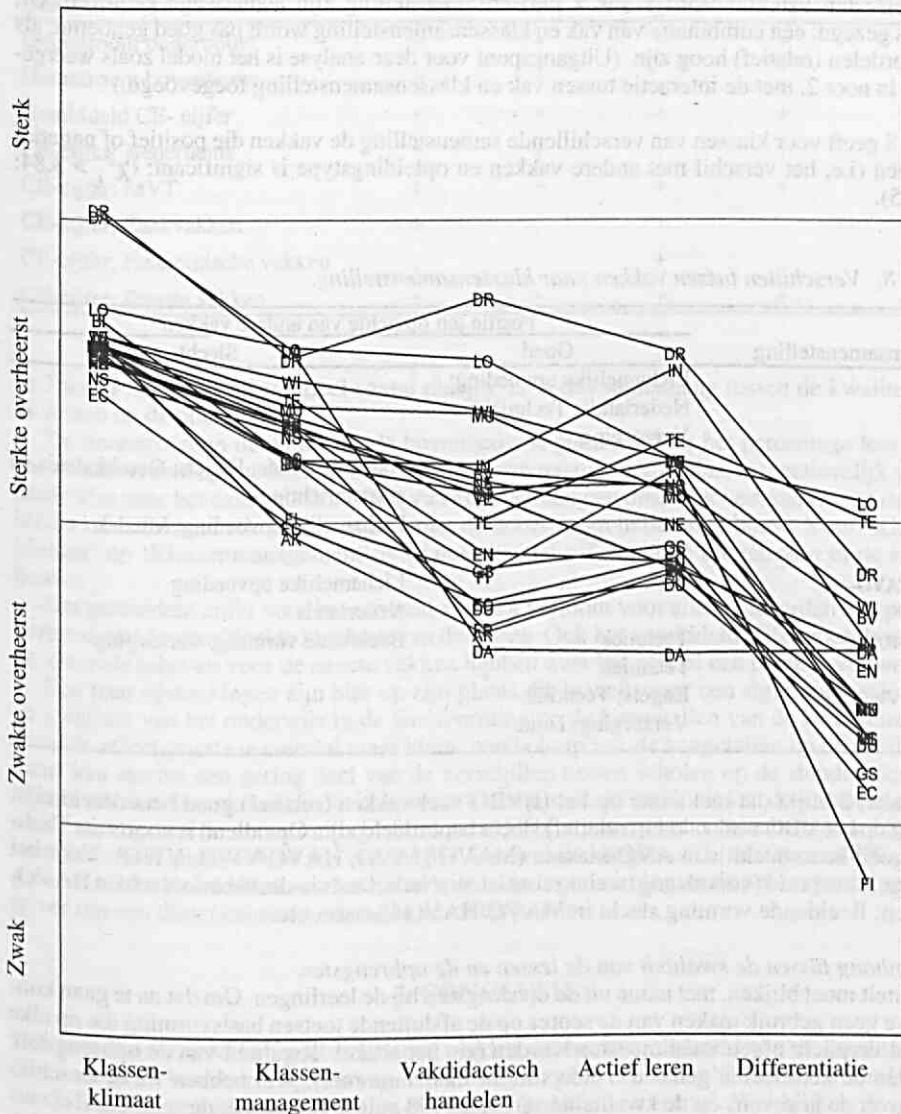
#### *Verschillen tussen vakken*

Wordt in het ene vak beter lesgegeven dan in het andere vak en geldt dat dan alleen voor bepaalde standaarden of voor alle standaarden? Uit Figuur 5 is het antwoord op deze vragen af te leiden. In de figuur is per standaard het gemiddelde oordeel per vak weergegeven. Elke lijn geeft één van de negentien vakken weer die in de basisvorming worden gegeven.

Uit Figuur 5 blijkt duidelijk dat er forse kwaliteitsverschillen zijn in gemiddelden tussen de vakken. Uit de figuur blijkt ook dat de lijnen voor de figuren elkaar betrekkelijk weinig kruisen. Dit houdt in dat er vakken zijn die over alle standaarden beter of slechter zijn dan andere vakken. Voorbeelden van goede vakken zijn lichamelijke opvoeding (LO) en muziek (MU); vakken die slecht uit de bus komen zijn economie (EC) en aardrijkskunde (AK). Deze laatste vakken zijn gemiddeld, met uitzondering van 'Klassenklimaat', op alle standaarden onvoldoende.

Opvallend is verder dat de verschillen tussen de vakken van standaard tot standaard varieert. Bij 'Klassenklimaat' zijn de verschillen betrekkelijk klein. Voor 'Vakdidactiek' en vooral voor 'Differentiatie' zijn de verschillen erg groot. Enkele van deze verschillen zullen voortvloeien uit de aard van het vak, maar voor het overige zijn ze groter dan geaccepteerd kan worden.

Figuur 5: Gemiddelde oordelen per vak (y-as) per standaard (x-as; AK: Aardrijkskunde; BI: Biologie; BV: Beeldende vorming; DA: Dans; DR: Drama; DU: Duits; EC: Economie; EN: Engels; FI: Fries; FR: Frans; GS: Geschiedenis en staatsinrichting; IN: Informatica; LO: Lichamelijke opvoeding; MU: Muziek; NE: Nederlands; NS: Natuur- en scheikunde; TE: Techniek; VE: Verzorging; WI: Wiskunde).



*Verschillen naar klassensamenstelling*

Bij alle lesobservaties is de samenstelling van de klas vastgelegd, zoals de school deze benoemt (VWO, VWO/HAVO, VWO/HAVO/MAVO etc.). De oordelen kunnen ook uitgesplitst worden naar samenstelling van de bezochte klassen. Dan kan nagegaan worden of de kwaliteit van de lessen verschilt met de klassensamenstelling. In de Tabel 8 zijn de resultaten van deze analyse samengevat. Daarbij zijn wij uitgegaan van alle oordelen over het lesbezoek; interactie-effecten van standaard x vak x klassensamenstelling zijn achterwege gebleven. Of, anders gezegd: een combinatie van vak en klassensamenstelling wordt pas goed genoemd, als alle oordelen (relatief) hoog zijn. (Uitgangspunt voor deze analyse is het model zoals weergegeven in noot 2, met de interactie tussen vak en klassensamenstelling toegevoegd.)

Tabel 8 geeft voor klassen van verschillende samenstelling de vakken die positief of negatief opvallen (i.e. het verschil met andere vakken en opleidingstype is significant;  $\chi^2_1 > 3.84$ ;  $p < .05$ ).

Tabel 8: *Verschillen tussen vakken naar klassensamenstelling.*

| Klassensamenstelling | Positie ten opzichte van andere vakken                         |   |
|----------------------|--|---|
|                      | Goed   | Slecht  |
| (1)VBO               | Lichamelijke opvoeding;<br>Nederlands; Techniek;<br>Verzorging |   |
| VBO                  |  | Aardrijkskunde; Engels; Geschiedenis en<br>staatsinrichting;<br>Lichamelijk opvoeding; Muziek |
| MAVO                 | Biologie   |   |
| VBO/MAVO             |  | Lichamelijke opvoeding  |
| HAVO                 |  | Nederlands  |
| MAVO/HAVO            | Techniek   | Beeldende vorming; Verzorging   |
| VWO                  | Techniek   |   |
| HAVO/VWO             | Engels; Techniek   |   |
| Breed                | Verzorging; Duits  |   |

Uit Tabel 8 blijkt dat met name op het (1)VBO veel vakken (relatief) goed beoordeeld zijn, terwijl op het VBO veel vakken (relatief) slecht beoordeeld zijn. Opvallend is voorts dat Techniek goed beoordeeld is in AVO-klassen (MAVO/HAVO, HAVO/VWO, VWO). Voor het overige is het beeld (vooralsnog) weinig consistent; Nederlands is slecht beoordeeld in HAVO-klassen; Beeldende vorming slecht in MAVO/HAVO-klassen, etc.

*Samenhang tussen de kwaliteit van de lessen en de opbrengsten*

Kwaliteit moet blijken, met name uit de opbrengsten bij de leerlingen. Om dat na te gaan konden we geen gebruik maken van de scores op de afsluitende toetsen basisvorming die op elke school verplicht afgenomen moesten worden (zie het artikel 'Kwaliteit van de opbrengsten: Worden de kerndoelen gehaald?' elders in dit themanummer). Wel hebben we de beschikking over de gegevens op de kwaliteitskaart van 1998 van de deelnemende scholen. Het verband tussen beide variabelen is nagegaan met behulp van een univariaat hiërarchisch model, waarin per verklarende variabele (i.e. de kengetallen van de kwaliteitskaart) het effect op een standaard is getoetst. De resultaten van deze analyse naar de relatie tussen de oordelen over de kwaliteit van de lessen en deze kengetallen is in Tabel 9 samengevat.

Tabel 9: Significante relaties tussen de kengetallen van de kwaliteitskaart (1998) en de scores op de standaarden ('+' : positieve verband ( $\alpha = .05$ ); '-' : negatief verband ( $\alpha = .05$ ); '.' : geen verband).

| Onafhankelijke variabele      | Standaard         |                          |                        |                 |                     |
|-------------------------------|-------------------|--------------------------|------------------------|-----------------|---------------------|
|                               | Vak-<br>didactiek | Klassen-<br>kli-<br>maat | Klassen-<br>management | Actief<br>leren | Differen-<br>tiatie |
| Doorstroom onderbouw          | +                 | +                        | +                      | +               | .                   |
| Doorstroom bovenbouw          | +                 | +                        | .                      | .               | .                   |
| Gemiddeld CE- cijfer          | +                 | +                        | +                      | +               | +                   |
| CE-cijfer: Nederlands         | +                 | +                        | .                      | .               | .                   |
| CE-cijfer: MVT                | +                 | +                        | +                      | +               | .                   |
| CE-cijfer: Zaakvakken         | +                 | .                        | +                      | .               | -                   |
| CE-cijfer: Economische vakken | .                 | .                        | +                      | .               | -                   |
| CE-cijfer: Exacte vakken      | +                 | +                        | +                      | +               | .                   |

In Tabel 9 staat een substantieel aantal plusjes: er is dus samenhang tussen de kwaliteit van de lessen en de opbrengsten.

De doorstroom in de onder- en de bovenbouw is gemeten door het percentage leerlingen dat zonder zittenblijven doorstroomt van de eerste naar de derde klas, respectievelijk van de derde klas naar het examen. Er is voor vier standaarden een duidelijk verband tussen de kwaliteit van de lessen in de basisvorming met de doorstroom in de onderbouw. Voor 'Klassenklimaat' en 'Klassenmanagement' is er ook positief effect op de doorstroom in de bovenbouw.

Het gemiddeld cijfer voor het centraal examen vertoont voor alle standaarden een positief verband met de gemiddelde kwaliteit van de lessen. Ook het gemiddeld cijfer voor de moderne vreemde talen en voor de exacte vakken hebben over het geheel een positief verband.

Een paar opmerkingen zijn hier op zijn plaats. Er is weliswaar een significant effect van de kwaliteit van het onderwijs in de basisvorming op de kengetallen van de kwaliteitskaart, maar de effectgrootte is meestal maar klein; met behulp van de kengetallen op de kwaliteitskaart kan slechts een gering deel van de verschillen tussen scholen op de standaarden verklaard worden. Daar komt bij dat de causale keten tussen de kwaliteit van de lessen enerzijds en de meting van de doorstroming of het gemiddeld cijfer op het centraal examen bij (nota bene) een heel ander cohort wel erg lang is. Hoewel de aangetoonde relaties een effect suggereren tussen de kwaliteit van het onderwijs in de basisvorming en de opbrengsten, gaat het te ver om een direct causaal verband te veronderstellen.

#### 4. CONCLUSIE

Het geheel overziend kan geconcludeerd worden dat het overall-beeld van de kwaliteit van de lessen in de basisvorming niet erg rooskleurig is; geen één school komt als 'vooral sterk' uit de bus. Het gemiddelde oordeel is hoogstens 'sterkte overheerst'. Wel blijken de oordelen op de verschillende standaarden redelijk hoog samen te hangen. Dat wil zeggen: scholen die (relatief) goed beoordeeld zijn op de ene standaard, zijn dat ook op een andere standaard. Enigszins gechargeerd zou geconcludeerd kunnen worden dat er goede en slechte scholen zijn, of wellicht beter: sommige scholen zijn in staat betere docenten aan te nemen (dan wel bij of na te scholen) dan andere scholen. Daarbij moet echter wel aantekend worden dat



'goed' hier een zeer relatieve betekenis heeft als zijnde 'sterkte overheerst'. Bovendien moet de score op een standaard, uiteraard, relatief begrepen worden: de relatieve positie van een school (t.o.v. andere scholen) op de ene standaard zegt iets over de relatieve positie van een school op een andere standaard, want ook tussen de standaarden zijn duidelijke verschillen aangetoond.

Er zijn vooral verschillen in kwaliteit tussen leraren. De verschillen in scholen zijn in vergelijking daarmee klein. Ook is gebleken dat de verschillen tussen docenten relatief constant zijn over de vijf standaarden. Dit leidt tot de conclusie dat er 'goede' docenten en 'slechte' docenten zijn. Gezien de grootte van de verschillen tussen docenten enerzijds en de verschillen tussen scholen anderzijds, lijkt het voor de hand te liggen om beleidsmaatregelen met het oog op onderwijsverbetering vooral te formuleren op docentniveau.

Er zijn duidelijke verschillen tussen vakken, groter dan door de aard van de vakken verklaard kan worden. Ook hier moeten we concluderen dat, grosso modo, de rangorde van vakken relatief constant is over de vijf standaarden. Vakken die op de ene standaard (relatief) goed beoordeeld zijn, zijn dat ook (relatief) goed op een andere standaard. Vakken die er relatief positief uitspringen zijn: lichamelijke opvoeding en muziek. In negatief opzicht vallen de drie moderne vreemde talen, economie en aardrijkskunde op.

Er is een zwak, maar consistent verband tussen de kwaliteit van de lessen en de opbrengsten van het onderwijs zoals dat wordt gemeten met de kengetallen van de kwaliteitskaart.

Het ligt voor de hand op grond van het voorafgaande de conclusie te trekken dat de doelen van de basisvorming, in elk geval wat betreft de kwaliteit van de lessen, niet gehaald zijn, en meer somber nauwelijks haalbaar zijn. Echter, zo simpel ligt het niet. De grote verschillen tussen leraren en tussen vakken laten zien dat het ook goed kan. Alleen de kleine verschillen tussen scholen kunnen tot de conclusie leiden dat scholen er niet veel toe doen en dat het niet goed mogelijk is om scholen als geheel beter te laten functioneren, anders dan door voor betere leraren te zorgen. Maar we moeten ons wel bedenken dat de verschillen tussen scholen in feite niets anders zijn dan geaggregeerde oordelen over docenten. Zo bezien laat de kwaliteit van docenten - of het gebrek daaraan - zich zowel op docent- als op schoolniveau voelen.

## NOTEN

1. In feite ligt de zaak iets gecompliceerder dan hier voorgesteld wordt. Door de normering is er een relatie tussen de indicatoren en het onderscheid tussen voldoende en onvoldoende. Uit de aard van de zaak is er daardoor al een positieve relatie tussen de scores op de indicatoren die onderdeel uitmaken van de norm en het oordeel.

2. Stel  $Y_{hijk}$  is het oordeel over de  $h^{de}$  standaard van de  $j^{de}$  les van de  $k^{de}$  docent in school  $k$ . We definiëren vijftien dummy-variabelen, zeg  $Z_{1hijk}, \dots, Z_{15hijk}$ , die de vijftien onderscheiden vakken indiceren (i.e. de dummy  $Z_{1hijk}$  staat alleen 'aan' (is gelijk aan één) als het een oordeel over het eerste vak betreft, in alle andere gevallen staat deze dummy 'uit' (is gelijk aan nul), etc.). Voorts worden vijf dummy-variabelen gedefinieerd, zeg  $X_{1ijk}, \dots, X_{5ijk}$ , die de vijf standaarden indiceren (i.e. de eerste dummy-variabele ( $X_{1ijk}$ ) staat alleen 'aan' als het een oordeel over de eerste standaard betreft, etc.) Met behulp van het product van beide typen dummy-variabelen kan elke combinatie van vak en standaard uniek geïdentificeerd worden (i.e.  $Z_{1hijk} * X_{1ijk}$  indiceert de combinatie van het eerste vak met de eerste standaard). Door deze combinaties op te nemen in het vaste deel van het model kunnen voor al deze combinaties van vak en standaard de gemiddelden geschat worden. Naast deze gemiddelden willen we ook per standaard de variantie op de onderscheiden niveaus (les, docent en school) schatten. Hiervoor worden de dummy-variabelen die de vijf standaarden indiceren opgenomen in het random deel van het model (i.e.  $X_{1ijk}, \dots, X_{5ijk}$ ). Meer formeel kan het model geschreven worden als:

$$\begin{aligned}
 Y_{hijk} = & Z_{1hijk} ( B_1 * X_{1ijk} + B_2 * X_{2ijk} + \dots + B_5 * X_{5ijk} ) + \\
 & Z_{2hijk} ( B_6 * X_{1ijk} + B_7 * X_{2ijk} + \dots + B_{10} * X_{5ijk} ) + \\
 & \dots \\
 & Z_{15ijk} ( B_{71} * X_{1ijk} + B_{72} * X_{2ijk} + \dots + B_{75} * X_{5ijk} ) + \\
 & [ X_{1ijk} * e_{1ijk} + X_{2ijk} * e_{2ijk} + \dots + X_{5ijk} * e_{5ijk} + \\
 & X_{1ijk} * u_{10jk} + X_{2ijk} * u_{20jk} + \dots + X_{5ijk} * u_{50jk} + \\
 & X_{1ijk} * v_{100k} + X_{2ijk} * v_{200k} + \dots + X_{5ijk} * v_{500k} ],
 \end{aligned}$$

$$(h = 1, 2, \dots, 5; i = 1, 2, \dots, I_j; j = 1, 2, \dots, J_k; k = 1, 2, \dots, K).$$

De vijfenzeventig combinaties van standaard en vak leveren evenzo vele gemiddelden (i.e.  $B_1 \dots B_{75}$ ). Op les-, docent- en schoolniveau worden vijf residuele scores onderscheiden; één per standaard. We nemen aan dat deze residuele scores normaal verdeeld zijn met verwachte waarde nul en een variantie van respectievelijk  $S^2_{e_{1ijk}}, \dots, S^2_{e_{5ijk}}, S^2_{u_{10jk}}, \dots, S^2_{u_{50jk}}, S^2_{v_{100k}}, \dots, S^2_{v_{500k}}$ . We nemen aan dat de verschillende residuele scores op de verschillende niveaus ongecorrleerd zijn. Echter aan de covariantiematrices op de drie niveaus worden géén restricties opgelegd; per niveau wordt de volledige [5X5] covariantiematrix geschat.

## LITERATUUR

- Goldstein, H. (1995). *Multilevel statistical models*. (Kendall's library of statistics 3). London: Arnold.
- Heuvelmans, A.P.J.M. & Sanders, P.F. (1993). Beoordelaarsovereenstemming. In: T.J.H.M. Eggen & P.F. Sanders (red.), *Psychometrie in de praktijk*. Arnhem: Cito.
- Janssens, F.J.G. (1997). *Van onderzoek naar evaluatie. De methodologie van de onderwijsinspectie*. Den Haag: Delwel.
- Karstanje, P.N. (1996). *Beleidstheorie basisvorming: een proeve van reconstructie*. Amsterdam: SCO-Kohnstamm Instituut.
- Linden, W.J. van der; Meijer, R.R.Z. & Vos H.J. (1997). *Normeringsmethoden voor inspectie-evaluaties*. Rapport 97-01. Enschede: Universiteit Twente, Toegepaste Onderwijskunde.
- Peters-Sips, M. & Zwarts, M. (1999). *Onderwijsleerproces basisvorming: Evaluatierapport*. Utrecht: Inspectie van het Onderwijs.
- Veldhijzen, N.H., Godebeld, P. & Sanders, P.F. (1993). *Klassieke testtheorie en generaliseerbaarheidstheorie*. In: T.J.H.M. Eggen & P.F. Sanders (red.), *Psychometrie in de praktijk*. Arnhem: Cito.

## Kwaliteit van het vakspecifieke aanbod

Marielle Peters-Sips, *Inspectie van het Onderwijs*

Michel Zwarts, *Inspectie van het Onderwijs*

Huub van den Bergh, *Utrechts Instituut voor Linguïstiek (UU)*

Lisette Schuurmans, *Inspectie van het Onderwijs*

### ABSTRACT

*One of the main objectives of the Education Act of 1992 called 'the Act Basisvorming' was the elimination of the categorisation in the first years of Secondary Education. Within the existing structure of school types a virtually uniform curriculum for all 12 to 15 year olds was formed. For all the pupils, the table indicating the number of lessons per subject comprises at least 15 subjects. Attainment targets have been laid down for all pupils.*

*Between 1997 and 1999 an evaluation of the 'Basisvorming' was carried out. In this article the following evaluation question is answered: 'Is the teaching content in basic secondary education consistent with the subject-specific attainment targets?'*

*Inspectors of schools judged these consistencies for 19 different subjects in 120 Dutch schools. The treatment of 334 attainment targets in methods was studied. Evaluation forms are completed by teachers of all the 120 schools. Inspectors studied the curricula, materials and tests of the schools for all subjects, they interviewed teachers (1800 interviews) and visited more than 7200 lessons.*

*The overall conclusion of the evaluation was that not even one school treated the attainment targets seriously enough for all the subjects. Differences were found between the different subjects and between the different types of schools.*

### 1. INLEIDING

Om het centrale doel van de basisvorming 'een brede vorming voor alle leerlingen' te realiseren was onder meer een modernisering en harmonisering van het onderwijsprogramma in de eerste leerjaren van het voortgezet onderwijs (VO) nodig (zie het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer). Met het formuleren van kerndoelen is het aanbod van het funderend onderwijs geactualiseerd en gemoderniseerd. Voor het VO zijn de kerndoelen opgenomen in het kernprogramma of kerncurriculum van minimaal 15 vakken dat in artikel 11a van de Wet op het Voortgezet Onderwijs (WVO) verplicht is gesteld voor leerlingen in de eerste leerjaren van het voortgezet onderwijs (Wet Basisvorming, Staatsblad 1992, 270). De kerndoelen zijn vastgesteld bij algemene maatregel van bestuur (AMVB).

In het Besluit Kerndoelen en Adviesurentabel Basisvorming 1993-1998 (Staatsblad, 1993, 207) wordt een onderscheid gemaakt tussen: algemene vaardigheidsdoelen, de algemene doelstelling per vak en de kerndoelen per vak. De drie typen doelstellingen heten tezamen (ook) kerndoelen. Wanneer in dit themanummer gesproken wordt van 'de kerndoelen' worden meestal de vakspecifieke kerndoelen bedoeld. Algemene vaardigheidsdoelen beschrijven vakgebied overstijgende kwaliteiten van leerlingen op het gebied van vaardigheden. Ze gelden voor alle vakken gezamenlijk, respectievelijk voor de basisvorming als geheel (zie het artikel 'Kwaliteit van het vakoverstijgende aanbod: het voorkomen van de algemene

vaardigheden' elders in dit themanummer). Per vak is vervolgens een algemene doelstelling geformuleerd. Deze schetst het algemene kader voor het betreffende vak. De algemene doelstelling geeft in veel gevallen aanwijzingen voor de te hanteren (vak)didactische aanpak van het vak. Eveneens per vak ten slotte, zijn specifieke kerndoelen geformuleerd. De vakspecifieke kerndoelen geven een beschrijving van de kwaliteiten van leerlingen op het gebied van kennis, inzicht en vaardigheden. Deze kerndoelen zijn een beredeneerde keuze uit de mogelijkheden die de algemene doelstelling voor het vak biedt.

De kerndoelen (hier is bedoeld: algemene vaardigheidsdoelen, algemene doelstellingen per vak én vakspecifieke kerndoelen) hebben een dubbele gebruiksfunctie. Ze zijn voor de overheid een maat waarmee kwaliteit periodiek bepaald en bewaakt wordt. Kwaliteit in termen die een indicatie geven van de onderwijsinhoud, als selectie uit het culturele erfgoed in samenhang met de eisen vanuit de samenleving (PMB, 1992). Voor de school zijn de kerndoelen te hanteren als aan het eind van de basisvorming te bereiken doelstellingen. Kerndoelen gelden dus voor het eind van de basisvorming; het zijn wettelijk verplichte deugdelijkheidseisen voor het onderwijs. Het kerncurriculum omvat drie leerjaren, maar kan in bepaalde gevallen zo worden ingericht, dat de kerndoelen sneller -in ten minste twee jaar- dan wel later -in ten hoogste vier jaar- worden bereikt (WVO, artikel 11b). Enerzijds moet er voor overheid en samenleving voldoende zicht zijn op de kerninhouden van het funderend onderwijs, anderzijds moet er principieel en voldoende ruimte voor de school zijn voor eigen in- en aanvulling.

Omdat de vakspecifieke kerndoelen aangeven welke inhoudelijke vernieuwing van het onderwijs in de vakken van de basisvorming is beoogd en welke reeds eerder nagestreefde doelen van belang zijn gebleven, heeft de Inspectie bij de evaluatie van de basisvorming veel aandacht aan het aanbod van deze kerndoelen besteed. Voor alle vakken is nagegaan in hoeverre de afzonderlijke kerndoelen deel uitmaken van het leerstofaanbod: frequent gebruikte methoden zijn geanalyseerd op de aanwezigheid van kerninhouden, waarin de kerndoelen uiteengelegd zijn, secties zijn schriftelijk en mondeling ondervraagd en in de lessen is geregistreerd welke domeinen (clusters van kerndoelen) aan de orde zijn gesteld. In het eindrapport over de evaluatie van de basisvorming (Inspectie van het Onderwijs, 1999a) en in het interne technische rapport (Peters-Sips & Zwarts, 1999) is in het algemeen gerapporteerd over het aanbod van de kerndoelen, in de diverse vakrapporten (Inspectie van het Onderwijs, 1999b) is per vak over het aanbod gerapporteerd.

In dit artikel gaan wij nader in op de evaluatie van het aanbod van de vakspecifieke kerndoelen zonder diep in de inhouden van alle negentien vakken te treden, aangezien dat in het bestek van dit themanummer niet mogelijk is. De volgende vragen vormen daarbij het uitgangspunt:

1. Hoe is voor de verschillende vakken het aanbod van de kerndoelen en wat zijn de verschillen tussen de opleidingen?
2. Hoe oordelen vaksecties over de haalbaarheid en het aanbod van de kerndoelen en wat zijn de verschillen tussen de opleidingen?
3. In welke mate komen voor de verschillende vakken de kerndoelen aan bod in de frequent gebruikte methoden?
4. Hoe is voor de verschillende vakken het aanbod van de kerndoelen als daarbij onderscheid wordt gemaakt tussen 'oude' en 'nieuwe' kerndoelen?
5. Is de het aantal lessen dat leerlingen ontvangen in overeenstemming met de adviesuren-tabel?

## 2. METHODE

Voor ieder kerndoel is voor iedere opleidingstype (VBO, MAVO, HAVO en/of VWO) op iedere school beoordeeld of het voldoende aan bod komt. In totaal zijn meer dan 100.000 oor-



delen uitgesproken door inspecteurs. De oordelen betreffen gemiddeld zestien vakken met in totaal 334 kerndoelen en 315 opleidingen op de 120 bezochte scholen met een gemiddelde van bijna 900 oordelen per school. De inspecteurs hebben zich daarbij gebaseerd op het aanbod zoals dat bleek uit: de gebruikte methode(n), aanvullend onderwijsmateriaal, vakwerkplannen, door vaksecties ingevulde vragenlijsten en met vaksecties gevoerde gesprekken. Daarnaast hebben inspecteurs tijdens lesobservaties (zie het artikel 'Kwaliteit van het onderwijsleerproces' elders in dit themanummer) geregistreerd aan welke domeinen van het vak in de les aandacht is geschonken.

Om inspecteurs in korte tijd tot een verantwoord oordeel te laten komen over de kwaliteit van het aanbod op iedere school is een duidelijke behoefte gebleken aan inzicht in de mate waarin met de gebruikte methoden het voorgeschreven aanbod gerealiseerd kan worden. Daarom is -in het jaar voorafgaand aan het schoolbezoek- op de 120 scholen van de steekproef en op de 120 scholen van de reserve steekproef (zie het artikel 'Kwaliteit van het onderwijsleerproces' elders in dit themanummer) geïnventariseerd welke methoden voor de verschillende vakken in de basisvorming werden gebruikt. Vervolgens is aan de betrokken educatieve uitgeverijen gevraagd om de betreffende methoden te (laten) analyseren aan de hand van door de Inspectie opgestelde criteria. Bij deze criteria nemen de kerndoelen een belangrijke plaats in. De kerndoelen waren voor dit doel door vakinspecteurs in samenwerking met vakdidactici uiteengelegd in kerninhouden, waarin de relatie tussen kerndoel en leerstofinhoud werd gelegd. Op basis van de door uitgeverij aangeleverde methodenanalyses zijn door onafhankelijke deskundigen van buiten de inspectie (vakdidactici en onderwijskundigen) vervolgens het basisvormingsgehalte en de didactische kwaliteit van de methoden beoordeeld.

Aan de hand van de gebruikte methode en een schriftelijke opgave over de behandeling en de haalbaarheid van de kerndoelen door de verschillende vaksecties is door de inspectie op iedere bezochte school een gesprek gevoerd met iedere vaksectie over de behandeling van de kerndoelen. Wanneer de methode voor bepaalde vakspecifieke kerndoelen hiaten bevatte, of indien leraren bepaalde delen uit de methode oversloegen, is per kerndoel vastgesteld wat daarvan de consequenties waren. Indien leraren voor ontbrekende of overgeslagen kerndoelen aanvullende materialen gebruikten, diende de vaksectie aannemelijk te maken dat dit aanvullende materiaal de hiaten in de vakspecifieke kerndoelen dekte. De informatie van de vaksectie over de behandeling van de kerndoelen en het methodegebruik is door de inspecteurs gecheckt aan de hand van de het vakwerkplan, (praktisch) werk van leerlingen, gesprekken met leerlingen, observaties in klassen, toetsen en proefwerken e.d.

Inspecteurs hebben per school per opleidingstype voor ieder kerndoel aangegeven of het 'niet', 'matig' dan wel 'voldoende' in het aanbod aan de orde kwam. Als alle leerlingen in een bepaalde opleiding alleen maar met een kerndoel 'in aanraking' zijn gebracht, is gesproken van een 'matig' aanbod van het kerndoel, als er sprake was van zodanige aandacht voor het kerndoel dat de leerlingen daardoor in staat gesteld werden het kerndoel te 'beheersen' is gesproken van een 'voldoende' aanbod. Voor het uitspreken van het oordeel per opleiding per vak is gebruik gemaakt van een vierpuntsschaal lopend van 'zwak', via 'meer zwak dan sterk', naar 'meer sterk dan zwak' en vervolgens naar 'sterk'.

Als alle kerndoelen van een vak op enigerlei wijze ('matig' of 'voldoende') in een opleidingstype aan bod kwamen, is het aanbod van de kerndoelen van dat vak voor dat opleidingstype op die school als 'meer sterk dan zwak' beoordeeld. Als een school meer opleidingstypen had, gold dezelfde eis voor alle opleidingen. Bij het vaststellen van deze norm is uitgegaan van de wettelijke verplichting van de scholen om de kerndoelen te gebruiken als uitgangspunt voor het onderwijs. Aan die eis wordt minimaal voldaan als de leerlingen ten minste bij het aanbod met alle kerndoelen geconfronteerd worden.



### 3. RESULTATEN

#### *Het aanbod van de kerndoelen en de verschillen tussen de opleidingen*

Inspecteurs zijn in iedere bezochte vestiging nagegaan of leraren voor de verschillende vakken erin slagen om in het feitelijk gegeven onderwijs in alle aanwezige opleidingen ieder kerndoelen ('matig' dan wel 'voldoende') aan bod te laten komen. Per vestiging is per opleiding, per vak, een oordeel uitgesproken over de mate waarin het onderwijsaanbod de kerndoelen omvat.

Gemiddeld genomen hebben de scholen voor ongeveer de helft van de vakken een aanbod waarin alle kerndoelen ten minste 'matig' aan bod komen. Er zijn geen scholen waarin dat voor alle vakken het geval is. Geen school voldoet dus aan de verplichting om alle kerndoelen aan alle leerlingen aan te bieden. De in dit opzicht beste 25 procent van de scholen biedt in ongeveer negen vakken alle kerndoelen 'matig' aan. De in dit opzicht slechtste 25 procent van de scholen doet dit in ten hoogste zes vakken of minder.

In Tabel 1 is in de eerste kolom met gegevens het percentage vakken in een school vermeld, waarin volgens het oordeel van de inspecteurs alle kerndoelen ten minste 'matig' aan de orde worden gesteld. In die gevallen heeft de inspecteur het oordeel 'sterkte overheerst zwakte' of het oordeel 'vooral sterk' uitgesproken. Deze twee kwalificaties worden in het vervolg van de tekst samengenomen tot het oordeel 'voldoende', zoals in de kop van de kolom staat aangegeven. Dit percentage is het percentage van alle opleidingen op de school tezamen.

*Tabel 1: Percentage vakken per school met het oordeel 'voldoende' aanbod van kerndoelen en de percentages kerndoelen, die per school in 'voldoende' mate dan wel 'matig' in het aanbod zijn opgenomen.*

| Waardering scholen | Percentage vakken met oordeel 'voldoende' | Percentage kerndoelen 'voldoende' | Percentage kerndoelen 'matig' |
|--------------------|---|-----------------------------------|-------------------------------|
| Laagste percentage | 0   | 38                                | 73                            |
| Percentiel 25      | 40  | 56                                | 84                            |
| Gemiddelde         | 49  | 64                                | 88                            |
| Percentiel 75      | 60  | 70                                | 91                            |
| Hoogste percentage | 88  | 83                                | 96                            |

De tweede kolom van Tabel 1 heeft betrekking op het percentage kerndoelen dat over alle vakken per school 'voldoende' aan bod komt. Gemiddeld genomen komt een leerling met iets minder dan twee van de drie kerndoelen in voldoende mate in aanraking om daar iets van op te steken. In de scholen die tot de beste 25 procent wat dit betreft behoren, worden de leerlingen met iets meer dan tweederde van de kerndoelen in voldoende mate in contact gebracht.

De derde kolom geeft informatie over het percentage kerndoelen dat over alle vakken per school 'ten minste matig' aan bod komt. Bij de school met de laagste waardering komt 73 procent van de kerndoelen 'matig' aan de orde, bij de school met de hoogste waardering bijna 100 procent. Gemiddeld is het percentage 88 procent. Dit houdt in dat gemiddeld genomen een leerling in de loop van de basisvorming met 88 procent van de kerndoelen in aanraking komt. In vijftig procent van de scholen ligt het percentage kerndoelen waarmee de leerlingen in aanraking komen tussen de 84 en 91 procent.

Tabel 2: Percentage van het totaal aantal uitgebrachte oordelen over het aanbod van de kerndoelen per vak dat de kwalificatie 'voldoende' kreeg.

| Vak                    | Totaal | IVBO | VBO | MAVO | HAVO | VWO |
|------------------------|--------|------|-----|------|------|-----|
| Muziek                 | 64     | 53   | 52  | 69   | 71   | 70  |
| Informatiekunde        | 62     | 58   | 59  | 63   | 65   | 63  |
| Wiskunde               | 62     | 53   | 55  | 62   | 67   | 68  |
| Frans                  | 61     | 26   | 30  | 56   | 73   | 76  |
| Engels                 | 58     | 38   | 45  | 60   | 69   | 69  |
| Lichamelijke opvoeding | 58     | 53   | 55  | 58   | 62   | 64  |
| Beeldende vorming      | 56     | 42   | 45  | 58   | 63   | 66  |
| Duits                  | 52     | 31   | 34  | 53   | 66   | 64  |
| Techniek               | 47     | 35   | 43  | 47   | 53   | 54  |
| Verzorging             | 47     | 45   | 43  | 47   | 49   | 50  |
| Geschiedenis           | 45     | 20   | 28  | 45   | 57   | 59  |
| Economie               | 45     | 32   | 35  | 45   | 54   | 52  |
| Natuur-/scheikunde     | 44     | 31   | 33  | 42   | 53   | 54  |
| Nederlands             | 39     | 28   | 32  | 40   | 46   | 46  |
| Biologie               | 38     | 29   | 32  | 35   | 45   | 46  |
| Aardrijkskunde         | 37     | 20   | 22  | 39   | 46   | 49  |
| Totaal                 | 49     | 37   | 40  | 51   | 59   | 59  |

De verschillende vakken verschillen in de mate waarin zij het aanbod van de kerndoelen realiseren. Tabel 2 geeft in de kolom 'totaal' het percentage van het totaal aantal uitgebrachte oordelen dat 'voldoende' was ('sterk' of 'sterkte overheerst'). De vakken zijn in deze tabel in volgorde geplaatst van een hoog percentage oordelen 'voldoende' naar een laag percentage. Van alle uitgebrachte oordelen over de vakken heen luidt de helft 'voldoende'. De oordelen 'vooral sterk' en 'vooral zwak' zijn beide in 10 procent van de gevallen geveld, terwijl de oordelen 'sterkte overheerst zwakte' en 'zwakte overheerst sterkte' beide in 40 procent van de gevallen gegeven zijn.

De verschillen tussen de vakken zijn aanzienlijk. De beoordeling van de mate waarin de kerndoelen 'voldoende' aan de orde worden gesteld, is voor de vakken Muziek, Informatiekunde, Wiskunde, Frans, Engels, Lichamelijke opvoeding en Beeldende vorming significant beter dan voor de vakken Techniek, Verzorging, Geschiedenis, Economie, Natuur-/scheikunde, Nederlands, Biologie en Aardrijkskunde. Ook de vakken Duits, Techniek en Verzorging worden significant beter beoordeeld in dit opzicht dan de vakken Biologie en Aardrijkskunde<sup>1</sup>.

Bij dit overzicht kan worden opgemerkt dat in het algemeen door de inspecteurs volgens het normeringsvoorschrift is beoordeeld. Afwijkingen van het normeringsvoorschrift zijn vooral te vinden bij Wiskunde en Verzorging, waar ruim één vijfde respectievelijk ruim een kwart van de oordelen positiever zijn uitgevallen dan het normeringsvoorschrift aangaf.

Tussen de verschillende opleidingen bestaan eveneens verschillen in de mate waarin de kerndoelen in het aanbod voor de leerlingen zijn opgenomen. In het VWO en het HAVO zijn significant meer kerndoelen in het aanbod opgenomen, dan in het MAVO, VBO en IVBO. Het MAVO verschilt volgens het oordeel van de inspecteurs verder significant van het VBO en het IVBO. In het VWO en HAVO zijn bijna 60 procent van de oordelen over de mate waarin de kerndoelen in het aanbod aan de leerlingen zijn verwerkt 'voldoende', in het MAVO ruim 50 procent, in het VBO 40 procent en in het IVBO nog geen 40 procent.

De percentages oordelen 'voldoende' per opleiding lopen voor de verschillende vakken uiteen. De verschillen zijn aanzienlijk bij Aardrijkskunde, Geschiedenis en de Moderne Vreemde Talen, terwijl ze gering zijn bij Informatiekunde en Verzorging.

*Haalbaarheid en aanbod van kerndoelen volgens vaksecties*

Voorafgaand aan het schoolbezoek is aan alle secties gevraagd schriftelijk per kerndoel en per vaardigheidsdoel voor iedere opleiding die de school verzorgt, aan te geven of de sectie het betreffende doel haalbaar acht en of het in het aanbod is opgenomen (zie Tabel 3). (Omwille van de vergelijkbaarheid zijn in Tabel 3 de vakken in dezelfde volgorde geplaatst als in Tabel 2.)

*Tabel 3: Gemiddeld percentage kerndoelen per vak, dat ten minste matig behandeld wordt volgens de vakinspecteurs, en volgens de secties behandeld wordt respectievelijk haalbaar wordt geacht, vergeleken met het percentage kerndoelen dat gemiddeld in de methoden voor het vak aan de orde wordt gesteld.*

| Vak                    | Tenminste matig | Behandeld | Haalbaar | Aanbod in methoden |
|------------------------|-----------------|-----------|----------|--------------------|
| Muziek                 | 92              | 84        | 84       | 100                |
| Informatiekunde        | 91              | 90        | 84       | 100                |
| Wiskunde               | 95              | 92        | 81       | 93                 |
| Frans                  | 93              | 88        | 84       | 90                 |
| Engels                 | 93              | 86        | 86       | 79                 |
| Lichamelijke opvoeding | 93              | 84        | 84       | —                  |
| Beeldende vorming      | 82              | 78        | 75       | 100                |
| Duits                  | 91              | 86        | 82       | 93                 |
| Techniek               | 88              | 84        | 84       | 83                 |
| Verzorging             | 83              | 78        | 83       | 94                 |
| Geschiedenis           | 86              | 79        | 76       | 98                 |
| Economie               | 86              | 84        | 85       | 99                 |
| Natuur-/scheikunde     | 88              | 83        | 83       | 91                 |
| Nederlands             | 84              | 76        | 81       | 54                 |
| Biologie               | 85              | 82        | 87       | 95                 |
| Aardrijkskunde         | 85              | 80        | 83       | 88                 |
| Totaal                 | 88              | 83        | 83       | 90                 |

Geconcludeerd kan worden dat er ook op basis van de gegevens die de sectie verstrekten verschillen tussen de vakken geconstateerd worden. De verschillen tussen de vakken zijn daarbij echter -evenals bij de uitspraken 'ten minste matig' van de inspecteurs (weergegeven in de eerste kolom van Tabel 3)- minder groot dan de verschillen bij de oordelen 'voldoende' (Tabel 2). De antwoorden van de vaksecties liggen dichterbij de gemiddelde percentages 'ten minste matig' dan bij de gemiddelde percentages 'voldoende' van de inspecteurs.

Nagegaan is wat de relatie is tussen de inspectie-uitspraken 'ten minste matig' en de informatie die de secties leverden over het behandeld dan wel haalbaar zijn van de kerndoelen. De correlaties zijn hoog: respectievelijk 0.86 en 0.72.

*Tabel 4: Percentage kerndoelen dat door de secties aangeboden wordt dan wel realiseerbaar wordt geacht.*

| Opleiding | Behandeld | Haalbaar |
|-----------|-----------|----------|
| IVBO      | 75        | 67       |
| VBO       | 77        | 74       |
| MAVO      | 82        | 82       |
| HAVO      | 83        | 85       |
| VWO       | 83        | 86       |

Over de vakken heen zijn er verschillen tussen de opleidingen (Tabel 4). De kerndoelen worden in het algemeen meer haalbaar geacht naarmate de 'moeilijkheidsgraad' van de opleiding (achtereenvolgens IVBO, VBO, MAVO, HAVO en VWO) toeneemt, waarbij de onderlinge verschillen tussen IVBO, VBO en MAVO groter zijn dan tussen MAVO, HAVO en VWO. In het IVBO zou 67 procent van de kerndoelen haalbaar zijn, in het VWO 86 procent. Van de vakken wijken Beeldende vorming, Lichamelijke opvoeding, Techniek en Verzorging positief af van het beeld dat in het (I)VBO veel meer kerndoelen niet haalbaar zouden zijn dan in het AVO en VWO. De verschillen tussen de opleidingen zijn bij deze vakken namelijk klein.

Ook ten aanzien van de mate waarin de kerndoelen behandeld worden zijn er verschillen tussen vakken en opleidingen, maar de verschillen zijn minder groot dan bij de haalbaarheid. In het IVBO zou circa 75 procent van de kerndoelen worden aangeboden, in het VWO 83 procent. Bij Beeldende vorming, Informatiekunde en Verzorging zijn de verschillen tussen de opleidingen gering, bij Wiskunde, Aardrijkskunde en Geschiedenis zijn ze wel duidelijk aanwezig. Opmerkelijk is het hoge percentage behandelde kerndoelen bij Informatiekunde in het IVBO. Verdere details staan in de technische rapportage (Peters-Sips & Zwarts, 1999).

#### *Het aanbod van de kerndoelen in de methoden*

Tabel 3 geeft in de laatste kolom het percentage kerndoelen dat gemiddeld in de methoden voor het vak aan de orde wordt gesteld. Omdat veel methoden afzonderlijke delen hebben voor de verschillende opleidingstypen is zowel gemiddeld over de verschillende delen van één methode als over de verschillende methoden voor één vak. Bij de meeste vakken komt tachtig procent of meer van de kerndoelen aan de orde. Duidelijk onder deze grens scoren Nederlands, biologie en aardrijkskunde. Gekeken is ook wat het verband is tussen aan de orde komen in de methode en voorkomen in het aanbod. Geconcludeerd kan worden dat er over het geheel genomen een matig positief verband is. De correlatie verschilt echter per vak. Het hoogste positieve verband is bij Aardrijkskunde en Biologie gevonden, waar de correlatie circa .40 is. Bij Nederlands, Frans, Natuur-/scheikunde en Economie is geen verband gevonden. Verdere details staan in Peters-Sips & Zwarts (1999).

#### *Verschillen in beoordeling tussen 'oude' en 'nieuwe' doelen*

Voor alle vakken zijn de kerndoelen in te delen in doelen die ook al vóór de basisvorming deel uitmaakten van het curriculum en doelen die nieuw zijn in het curriculum voor de eerste leerjaren. Vakinspecteurs maakten tezamen met andere vakdeskundigen een indeling. Voor de oude en nieuwe doelen is nagegaan hoe de oordelen van de inspecteurs uitvielen en hoe de secties oordeelden over de haalbaarheid en de behandeling van deze doelen. In Tabel 5 zijn de resultaten van deze splitsing in oude en nieuwe kerndoelen opgenomen. (Indien geen oude dan wel nieuwe kerndoelen aangetroffen zijn, is dit in de tabel aangegeven met een liggend streepje.)

Of het nu gaat om oordelen van inspecteurs over de mate waarin de doelen 'voldoende' in het aanbod zijn opgenomen of om meningen van secties over de haalbaarheid en het aanbod van de doelen, in alle gevallen blijken de waarderungen positiever uit te vallen als het om doelen gaat die in het algemeen reeds vóór de basisvorming werden nagestreefd dan wanneer het om doelen gaat die met de basisvorming geïntroduceerd zijn. Met name ten aanzien van de beoordeling van de mate waarin de doelen 'voldoende' in het aanbod zijn opgenomen zijn de verschillen tussen 'oude' en 'nieuwe' doelen groot. De tendens is echter in alle gevallen hetzelfde. Bijvoorbeeld bij Aardrijkskunde en Nederlands zijn de verschillen tussen 'oud' en 'nieuw' aanzienlijk.

#### *Het aantal lessen dat leerlingen ontvangen in vergelijking met de adviesurentabel*

Voor het beoordelen van het onderwijsaanbod van een school is het nodig inzicht te hebben in de gehanteerde uren tabellen, in de feitelijke duur van het leerjaar en in de werkelijke leer-

Tabel 5: Gemiddeld percentage van de oordelen 'voldoende' over de oude en nieuwe doelen per vak en het gemiddelde percentage doelen dat haalbaar dan wel behandeld wordt geacht.

| Vak                    | Oordeel 'oud' | Oordeel 'nieuw' | Haalbaar 'oud' | Haalbaar 'nieuw' | Behandeld 'oud' | Behandeld 'nieuw' |
|------------------------|---------------|-----------------|----------------|------------------|-----------------|-------------------|
| Muziek                 | 67            | —               | 84             | —                | 84              | —                 |
| Informatiekunde        | 67            | 50              | 87             | 81               | 93              | 87                |
| Wiskunde               | 88            | 68              | 83             | 75               | 96              | 81                |
| Frans                  | 65            | 46              | 88             | 76               | 92              | 81                |
| Engels                 | 67            | 46              | 88             | 81               | 90              | 81                |
| Lichamelijke opvoeding | 81            | 39              | 89             | 64               | 84              | 83                |
| Duits                  | 66            | 49              | 86             | 74               | 90              | 80                |
| Techniek               | —             | 52              | —              | 84               | —               | 84                |
| Verzorging             | —             | 64              | —              | 83               | —               | 78                |
| Geschiedenis           | 65            | 61              | 77             | 72               | 79              | 78                |
| Economie               | —             | 64              | —              | 85               | —               | 84                |
| Natuur-/scheikunde     | 79            | 68              | 84             | 79               | 86              | 76                |
| Nederlands             | 72            | 40              | 89             | 75               | 88              | 66                |
| Biologie               | 76            | 55              | 89             | 81               | 86              | 68                |
| Aardrijkskunde         | 75            | 27              | 88             | 70               | 86              | 65                |
| Totaal                 | 72            | 51              | 86             | 77               | 88              | 78                |

Tabel 6: Aantal uren dat per vak in de adviesentabel staat, respectievelijk gemiddeld op de lessentabellen van de verschillende opleidingen en gemiddeld over alle opleidingen voorkomt op de meerderheid van de lessentabellen.

| Vak <sup>1</sup>       | Advies | IVBO             | VBO              | MAVO             | HAVO | VWO  | Totaal |
|------------------------|--------|------------------|------------------|------------------|------|------|--------|
| Nederlands             | 10,0   | 9,0              | 9,0              | 10,0             | 10,0 | 9,5  | 9,6    |
| Engels                 | 7,0    | 6,8              | 7,5              | 8,0              | 8,0  | 7,9  | 7,2    |
| Frans                  | 6,0    | —                | 3,8              | 8,0              | 8,0  | 8,0  | 8,0    |
| Duits (2) <sup>3</sup> | 6,0    | 4,0              | 4,0              | 6,0              | 6,0  | 5,5  | 5,5    |
| Wiskunde               | 10,0   | 8,0              | 9,0              | 10,0             | 10,0 | 10,0 | 9,6    |
| Natuur-/scheikunde (2) | 5,0    | 4,0              | 4,0              | 5,4              | 6,0  | 6,3  | 5,5    |
| Biologie (2)           | 3,0    | 3,6              | 3,0              | 6,0 <sup>2</sup> | 4,0  | 4,0  | 3,8    |
| Techniek (2)           | 4,5    | 5,4              | 4,0              | 4,0              | 4,0  | 4,0  | 4,2    |
| Geschiedenis           | 5,0    | 4,0 <sup>2</sup> | 4,0 <sup>2</sup> | 6,0              | 6,0  | 6,0  | 5,9    |
| Aardrijkskunde         | 3,5    | 3,0 <sup>2</sup> | 3,0 <sup>2</sup> | 5,5              | 6,0  | 5,5  | 5,6    |
| Economie (1)           | 2,0    | 2,0              | 2,0              | 5,0 <sup>2</sup> | 2,0  | 2,0  | 2,0    |
| Verzorging (1)         | 2,5    | 4,0 <sup>2</sup> | 2,0              | 2,0              | 2,0  | 2,0  | 2,0    |
| Informatiekunde (1)    | 0,5    | 1,0              | 1,0              | 1,0              | 1,0  | 1,0  | 1,0    |
| Lichamelijke opvoeding | 9,0    | 8,0              | 8,0              | 8,0              | 7,3  | 7,0  | 7,6    |
| Kunstvakken            | 7,0    | 9,0 <sup>2</sup> | 8,0 <sup>2</sup> | 8,1              | 9,0  | 8,5  | 8,8    |

<sup>1</sup> Tenzij anders vermeld, zijn per vak de uren vermeld bij een duur van drie jaar.

<sup>2</sup> Het urenantal geldt voor een andere duur van de basisvorming voor die opleiding.

<sup>3</sup> Het merendeel van de scholen heeft voor het vak de tussen haakjes aangegeven basisvormingduur.

tijd. Bij de evaluatie van de basisvorming zijn hierover in de bezochte scholen gegevens verzameld door de lesroosters voor de leerlingen in de verschillende opleidingen op te vragen en de duur van de periode van basisvorming in de school. Tijdens de lesobservaties is de leer-tijd in kaart gebracht. Voor het berekenen van de aantallen lessen (lessen met een andere



duur dan 50 minuten zijn omgerekend naar lessen van 50 minuten) is steeds uitgegaan van de opleidingsmogelijkheden die de school in het derde leerjaar biedt. Voor al die opleidingen is per school nagaan hoeveel uren leerlingen in de categorale, dakpan- of breder samengestelde derde, tweede en eerste klassen voor de verschillende vakken onderwijs hebben gehad. De meeste scholen sluiten de basisvorming na drie jaar af. In Tabel 6 staan de gemiddelde uren aantallen (in lessen van 50 minuten) per lesweek per opleidingstype afgezet tegen de adviesurentabel.

De scholen die de basisvorming na drie jaar afsluiten hebben in de periode van de basisvorming meestal drie jaar onderwijs verzorgd in de vakken Nederlands, Engels, Frans (vrijwel niet in het IVBO en VBO), Wiskunde, Geschiedenis en staatsinrichting (in het VBO meestal twee jaar), Aardrijkskunde (in het IVBO en VBO meestal twee jaar), Lichamelijke opvoeding en de Kunstvakken. Deze scholen hebben dan in de meeste gevallen twee jaar onderwijs verzorgd in Duits, Natuur-/scheikunde, Biologie (in het MAVO drie jaar) en Techniek, terwijl bij de scholen die de basisvorming na drie jaar afsluiten de vakken Economie (in het MAVO twee jaar), Verzorging (in het IVBO in twee jaar) en Informatiekunde meestal maar één jaar op het rooster hebben staan.

In veel scholen wordt de adviesurentabel niet exact gevolgd. Gemiddeld genomen staan er gedurende de periode van drie jaar basisvorming meer jaaruren op de rooster voor Aardrijkskunde, de Kunstvakken en Frans (circa twee jaaruren meer), Geschiedenis en staatsinrichting en Biologie (circa één jaaruur meer) en Natuur-/scheikunde en Informatiekunde (circa een half jaaruur meer). Daarentegen staan er minder jaaruren op het rooster voor Lichamelijke opvoeding (circa anderhalf jaaruur minder), Nederlands, Duits, Wiskunde en Verzorging (circa een half jaaruur minder).

*Tabel 7: Aantal (en percentage) vestigingen dat genoemde vakken ten minste in één leerjaar op het rooster voor een bepaalde opleiding combineert.*

| Combinatie                               | IVBO   | VBO    | MAVO  | HAVO  | VWO   |
|--|--------|--------|-------|-------|-------|
| Wiskunde en informatiekunde              | 2 (5)  | 1 (2)  | 3 (4) | 3 (4) | 5 (7) |
| Natuur-/scheikunde en informatiekunde    | 0 (0)  | 2 (4)  | 6 (7) | 3 (4) | 6 (9) |
| Natuur-/scheikunde en biologie           | 4 (11) | 7 (14) | 5 (6) | 5 (7) | 3 (4) |
| Techniek en informatiekunde              | 0 (0)  | 4 (8)  | 4 (5) | 2 (3) | 4 (6) |
| Aardrijkskunde, geschiedenis en economie | 8 (21) | 4 (8)  | 3 (4) | 3 (4) | 2 (3) |

In Tabel 6 is geen rekening gehouden met het voorkomen van vakkencombinaties op de lesentabel. Het aantal scholen dat in het onderzoeksjaar 1997/1998 vakkencombinaties op het lesrooster heeft staan is klein (Tabel 7). Een belangrijke reden voor het achterblijven van het gecombineerd aanbieden van vakken is dat hiervoor vrijwel geen leermiddelen op de markt worden aangeboden.

Aan de vakken van de basisvorming worden in totaal in de eerste drie leerjaren circa 84 in plaats van de geadviseerde 75 lessen van 50 minuten besteed (Tabel 8). Het surplus aan uren voor de vakken van de basisvorming blijkt uit alle lesroosters met uitzondering van de roosters voor het VBO, waar 70 in plaats van 75 uren aan de vakken van de basisvorming worden besteed in de eerste drie leerjaren, en het IVBO, waar in drie jaar één uur minder aan basisvormingvakken wordt besteed. Gemiddeld genomen worden de resterende 12,6 uren besteed aan Studielessen (twee uren), Mentoruur (één uur), Klassieke talen (ruim één uur), Oriëntatie op studie en beroep (bijna een half uur) en overige (meestal beroepsvoorbereidende) vakken (ruim acht en een half uur). Van de zogenaamde 'vrije ruimte' van 21 uren wor-

den er door de scholen negen besteed aan vakken van de basisvorming. Het totaal aantal van 96 lesuren in de eerste drie leerjaren wordt gemiddeld volgens alle lesroosters gerealiseerd, met uitzondering van het VBO. Gemiddeld genomen komen de leerlingen in het VBO een half jaaruur van de 96 uren te kort. In scholengemeenschappen met VBO, in HAVO/VWO en het Gymnasium krijgen de leerlingen in de eerste drie leerjaren gemiddeld tezamen 2 uren meer dan de voorgeschreven 96 uren.

Tabel 8: Aantal lesuren (van 50 minuten) besteed aan vakken en overige lessen.

| Vakken                         | Leerjaar 1 | Leerjaar 2 | Leerjaar 3 | Totaal |
|--------------------------------|------------|------------|------------|--------|
| Vakken basisvorming            | 29,1       | 29,5       | 25,2       | 83,8   |
| Studielessen                   | 1,0        | 0,4        | 0,3        | 1,7    |
| Mentoruren                     | 0,5        | 0,4        | 0,2        | 1,0    |
| Oriëntatie op studie en beroep | 0,1        | 0,3        | 0,1        | 0,5    |
| Grieks en Latijn               | 0,1        | 0,6        | 0,8        | 1,4    |
| Overige vakken                 | 1,4        | 1,2        | 5,4        | 8,0    |
| Totaal                         | 32,2       | 32,3       | 31,9       | 96,4   |

Tabel 9: Lesvervangende activiteiten in dagen op jaarbasis.

| Activiteiten         | Leerjaar 1 | Leerjaar 2 | Leerjaar 3 | Totaal |
|----------------------|------------|------------|------------|--------|
| Introductie          | 2,3        | 0,6        | 0,5        | 1,1    |
| Projecten            | 1,2        | 1,2        | 1,1        | 1,2    |
| Sportdagen           | 1,3        | 1,1        | 1,1        | 1,2    |
| Proefwerken          | 3,8        | 4,1        | 4,5        | 4,1    |
| Excursies            | 0,9        | 1,1        | 1,6        | 1,2    |
| Overige activiteiten | 0,7        | 0,8        | 1,1        | 0,9    |
| Totaal               | 10,0       | 8,8        | 9,9        | 9,6    |

Het aantal schoolweken bedraagt op jaarbasis minder dan de voorgeschreven 40 weken. Er gaan gemiddeld twee weken van het jaarrooster op aan zogenaamde lesvervangende activiteiten: introductieperiode, projecten, sportdagen, proefwerkweken, excursies en overige lesvervangende activiteiten (Tabel 9). Daarnaast gaan gemiddeld twee weken van het jaarrooster verloren door niet-lesvervangende activiteiten zoals rapportvergaderingen en vrije dagen aan het begin en einde van het schooljaar (Tabel 10). In totaal worden gemiddeld ongeveer 21 dagen per jaar anders ingevuld dan het weekrooster voorziet. Het aantal lesweken op jaarbasis bedraagt op de gemiddelde school 38 in plaats van het wettelijk voorgeschreven aantal van 40.

Een ander aspect van de tijdsbesteding is de werkelijke leertijd. Tijdens de lesbezoeken van de inspecteurs is voor iedere les genoteerd hoeveel procent van de bedoelde lestijd werd besteed aan organisatie, aan klassikale instructie, aan individueel of groepswork en hoeveel tijd niet effectief werd besteed. Gemiddeld over alle bijgewoonde lessen voor alle vakken heen, blijkt ongeveer 6 procent van de lestijd besteed te worden aan organisatie, 50 procent aan klassikale instructie, 37 procent aan individueel of groepswork, terwijl 7 procent niet effectief werd besteed.

Tabel 10: Niet-lesvervangende activiteiten in dagen op jaarbasis.

| Activiteiten            | Leerjaar 1 | Leerjaar 2 | Leerjaar 3 | Totaal |
|-------------------------|------------|------------|------------|--------|
| Rapportvergaderingen    | 2,4        | 2,4        | 2,5        | 2,4    |
| Begin leerjaar          | 1,0        | 1,3        | 1,3        | 1,2    |
| Einde leerjaar          | 4,5        | 4,4        | 4,6        | 4,5    |
| Overige geplande uitval | 1,0        | 1,1        | 0,8        | 1,0    |
| Onverklaarde uitval     | 0,7        | 0,7        | 0,6        | 0,7    |
| Totaal                  | 9,7        | 9,8        | 9,7        | 9,7    |

## 5 SAMENVATTING EN CONCLUSIES

Het voorafgaande kan voor wat betreft het aanbod van de vakspecifieke kerndoelen als volgt worden samengevat:

### *Inspecteuroordelen over het aanbod van kerndoelen*

- Geen enkele school voldoet aan de verplichting om alle kerndoelen aan te bieden. Er zijn wat dit betreft grote verschillen tussen de scholen.
- Een leerling krijgt gemiddeld bijna negentig procent van de kerndoelen op een minimaal niveau aangeboden. Er zijn ook scholen waar dat gemiddeld maar vijfenzeventig procent is.
- Gemiddeld genomen krijgt een leerling ruim zestig procent van de kerndoelen op een voldoende wijze aangeboden. Er zijn ook scholen waar dat maar veertig procent of minder is.
- De verschillen tussen de vakken zijn groot: bij Muziek, Informatiekunde, Wiskunde en Frans heeft de inspectie in meer dan 60 procent van de gevallen het oordeel 'voldoende' geveld over het aanbod, bij Aardrijkskunde, Biologie en Nederlands gebeurde dat in minder dan 40 procent van de gevallen.
- De verschillen tussen de opleidingen zijn groot: in het IVBO en het VBO heeft de inspectie in circa 40 procent van de gevallen een oordeel 'voldoende' over het aanbod uitgesproken, in het HAVO en VWO in circa 60 procent van de gevallen.

### *Sectieoordelen over het aanbod van kerndoelen*

- Gemiddeld genomen zijn leraren van mening 83 procent van de kerndoelen te behandelen. Tevens achten zij gemiddeld 83 procent van de kerndoelen haalbaar.
- Het percentage kerndoelen dat door leraren haalbaar wordt geacht ligt in het HAVO en VWO beduidend hoger dan in het IVBO en VBO. Ook worden in het VWO, HAVO en MAVO meer kerndoelen behandeld.

### *Aanbod van kerndoelen in de methoden*

- Tussen het aanbod van de kerndoelen in de methoden en het aan de orde stellen van de kerndoelen in het onderwijs bestaat een matig positief verband. Vakken verschillen in dit opzicht van elkaar.

### *Het aanbod van 'oude' en 'nieuwe kerndoelen*

- De oordelen van inspecteurs over het aanbod en de meningen van secties over de haalbaarheid en het aanbod van de doelen blijken in alle gevallen positiever uit te vallen als het om de reeds vóór de basisvorming bekende doelen gaat dan om doelen die met de basisvorming geïntroduceerd zijn. Gemiddeld over de vakken genomen zijn volgens de inspecteurs de 'oude' kerndoelen door circa 70 procent van de scholen in het aanbod opgenomen en de 'nieuwe' door circa 50 procent.

### *Tijd besteedt aan de basisvorming*

- Er wordt in het MAVO, HAVO en VWO meer tijd aan de vakken van de basisvorming besteed dan geadviseerd en in het VBO en IVBO minder.
- Vernieuwing van het aanbod door het gecombineerd aanbieden van vakken vindt maar sporadisch plaats.
- Vrijwel alle scholen bieden het wettelijk voorgeschreven weekrooster van 32 uren van 50 minuten aan.
- Gemiddeld bieden scholen meer uren aan in de basisvormingvakken dan wordt geadviseerd.
- Ongeveer twee weken per jaar krijgen leerlingen lesvervangende activiteiten aangeboden.
- Ongeveer twee weken per jaar vervallen de lessen voor de leerlingen, anders dan door ziekte van de leraren.

Uit de hiervoor beschreven resultaten moet worden geconcludeerd dat het feitelijk onderwijsaanbod in het schooljaar 1997/1998 strikt genomen niet voldeed aan de kenmerken van goed onderwijs in de basisvorming. Geen enkele school bood in alle vakken alle kerndoelen aan en daarmee voldeed geen enkele school aan de wettelijke verplichting om alle kerndoelen aan te bieden.

Op basis van deze evaluatieresultaten moet tevens worden geconcludeerd dat de kerndoelen in een groot aantal vakken tot een overladen programma leidden. *"De kerndoelen zijn soms nogal abstract of zo gedetailleerd dat een voldoende behandeling meer tijd vraagt dan beschikbaar is. Op schoolniveau is zeker sprake van overladenheid, getuige het feit dat geen enkele steekproefschool alle kerndoelen kon behandelen. Ook het feit dat de scholen meer tijd voor de basisvorming hebben ingeruimd dan geadviseerd, vormt een indicatie voor de geconstateerde overladenheid."* (Inspectie van het Onderwijs, 1999a).

In de onderzoeksresultaten betreffende het onderwijsaanbod zijn opmerkelijke verschillen aangetroffen tussen ('oude' en 'nieuwe') vakken, tussen opleidingstypen en tussen 'oude' en 'nieuwe' kerndoelen. Het lijkt zinvol gedetailleerder naar deze verschillen te kijken om meer grip te krijgen op het fenomeen 'overladenheid'. In het evaluatieonderzoek dat de inspectie uitvoerde was dit fenomeen niet expliciet in de vraagstelling opgenomen en derhalve niet expliciet onderzocht. Nader onderzoek lijkt gewenst.

Niet onvermeld moet blijven, dat bij de invoering van de basisvorming de kerndoelen niet direct in alle methoden zijn verwerkt, daarvoor was enige tijd nodig. Ook duurde het enige tijd voordat de scholen alle methoden door nieuwe hadden vervangen. In 1998 is de tweede generatie kerndoelen vastgesteld. Ook voor deze generatie zal gelden dat het enige tijd duurt voordat geschikte methoden door alle leerlingen worden gebruikt. Het kostenaspect moet in dit geval zeker niet verwaarloosd worden.

### NOOT

<sup>1</sup> De data zijn in een hiërarchisch model geanalyseerd en binnen dit model zijn de verschillen tussen de vakken als geheel en per vak getoetst.

## LITERATUUR

- Inspectie van het Onderwijs (1999a). *Werk aan de basis. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (1999b). *Werk aan de basis. Vakrapport Aardrijkskunde (c.q. Beeldende vakken/Biologie/Dans/Drama/Duits/Economie/Engels/Frans/Fries/Geschiedenis en staatsinrichting/Informatiekunde/Lichamelijke opvoeding/Muziek/Natuur-en scheikunde/Nederlands/Techniek/Verzorging/Wiskunde)*. Utrecht: Inspectie van het Onderwijs
- Peters-Sips, M. & Zwarts, M. (1999). *Onderwijsleerproces basisvorming. Evaluatierapport*. Utrecht: Inspectie van het Onderwijs.
- Procesmanagement Basisvorming (1992). *Ontwerp kerndoelen voor de basisvorming in basisonderwijs en voortgezet onderwijs. Inforeeks basisvorming 2*. PMB.

## SAMENVATTING EN CONCLUSIES

De inspectie van het Onderwijs heeft in 1994 een onderzoek naar de basisvorming in het basisonderwijs en het voortgezet onderwijs uitgevoerd. Het onderzoek is gericht op de vraag of de basisvorming in deze onderwijsvormen voldoende aandacht krijgt en of de kwaliteit van de basisvorming voldoende is. Het onderzoek is uitgevoerd op basis van een literatuurstudie, een enquête onder leerkrachten en een analyse van de leerplannen van de basisonderwijs en het voortgezet onderwijs. De conclusies van het onderzoek zijn dat de basisvorming in deze onderwijsvormen niet voldoende aandacht krijgt en dat de kwaliteit van de basisvorming niet voldoende is. Dit betekent dat de basisvorming in deze onderwijsvormen moet worden verbeterd. De inspectie van het Onderwijs heeft een aantal aanbevelingen gedaan om de basisvorming in deze onderwijsvormen te verbeteren. Deze aanbevelingen zijn gericht op de verbetering van de aandacht voor de basisvorming en de verbetering van de kwaliteit van de basisvorming. De inspectie van het Onderwijs zal de uitvoering van deze aanbevelingen volgen.



# Kwaliteit van het vakoverstijgende aanbod: het voorkomen van de algemene vaardigheden

Michel Zwarts, *Inspectie van het onderwijs*

Huub van den Bergh, *Utrechts Instituut voor Linguïstiek (UU)*

Marielle Peters, *Inspectie van het onderwijs*

## ABSTRACT

*One of the objectives of the educational reform called 'basisvorming' was the introduction of general, subject independent skills in the curriculum.*

*This study focusses on how the six formulated general skills are dealt with. The difference between school methods, subjects and school types is analysed.*

*For the six formulated general skills it is studied how they are treated in methods, if they are treated in different subjects and if they are treated the same in different types of schools.*

*It is shown that the treatment of general skills differs largely between subject matters. This conclusion is justified for both the methods used and the actual teaching in class.*

*If the notion of general skills is taken seriously they should be taught in each subject. However, even for a mild requirement such as 'each general skill has to be treated in two subjects' the vast majority of schools fail.*

*All in all, it has to be concluded that the general skills are treated rather poorly in Dutch secondary education.*

## 1. INLEIDING

Eén van de doelstellingen van de basisvorming is de modernisering en gedeeltelijke harmonisering van het onderwijsprogramma (zie het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer). Hiertoe zijn de zogenaamde kerndoelen geformuleerd. De kerndoelen hebben betrekking op het *verwerven* van kennis, inzicht en vaardigheden. De meeste zijn vakspecifiek (zie het artikel 'Kwaliteit van het vakspecifieke aanbod' elders in dit themanummer). Een aantal vaardigheidsdoelen is echter algemeen, vakoverstijgend, geformuleerd. Deze algemene vaardigheidsdoelen vormen de gemeenschappelijke kern in handlingsdoelen, die in de vakken verder aangevuld worden met specifieke vaardigheden (Procesmanagement Basisvorming, 1992). Met de tweede generatie kerndoelen, die vanaf 1998 van kracht zijn, zijn de algemene vaardigheidsdoelen in algemene onderwijsdoelen vertaald. Het belang ervan is sinds de invoering van de basisvorming alleen maar toegenomen door de nadruk die ook in de tweede fase ligt op actief en interactief leren.

Met de algemene vaardigheidsdoelen wordt de gemeenschappelijkheid van het curriculum benadrukt. De algemene vaardigheidsdoelen kunnen kort samengevat worden als:

1. de leerlingen kunnen eenvoudig onderzoek doen;
2. de leerlingen kunnen eigen standpunten verwoorden;
3. de leerlingen kunnen samenwerken aan opdrachten;
4. de leerlingen kunnen een relatie leggen met de beroepspraktijk;

5. de leerlingen leren eigen mogelijkheden en interesses ontdekken voor de keuze van een vervolgopleiding;
6. de leerlingen leren criteria hanteren bij het zelf beoordelen van werkstukken.

Twee van de zes algemene vaardigheden (4 en 5) zijn vaardigheden die betrekking hebben op de Oriëntatie op studie en beroep (OSB), de vier andere betreffen vooral vaardigheden die van belang zijn om zich zelfstandig kennis eigen te maken en greep te krijgen op het eigen leerproces: leren leren.

Omdat met de introductie van de algemene vaardigheden een specifieke inhoudelijke vernieuwing van het onderwijs beoogd is, heeft de inspectie hier bij de evaluatie van de basisvorming ruim aandacht aan besteed. Voor alle vakken werd nagegaan in hoeverre de algemene vaardigheden onderdeel uitmaken van het leerstofaanbod; veelgebruikte methoden werden geanalyseerd op de aanwezigheid van de algemene vaardigheden en het voorkomen van de algemene vaardigheden in de lessen is geregistreerd. In het eindrapport over de evaluatie van de basisvorming (Inspectie van het Onderwijs, 1999) is in het algemeen gerapporteerd over het aanbod van de algemene vaardigheden. In dit artikel gaan wij dieper in op het voorkomen van de algemene vaardigheden. Uitgangspunt daarbij zijn de volgende vragen:

1. Hoe is het aanbod van de algemene vaardigheden en wat zijn de verschillen tussen vakken en opleidingen?
2. In hoeverre worden de algemene vaardigheden vakoverstijgend aangeboden?
3. In welke mate komen de algemene vaardigheden aan bod in de frequent gebruikte methoden?
4. In welke mate komen de verschillende algemene vaardigheden in de lessen voor en in hoeverre zijn er verschillen tussen vakken, klassen en lestypen?

## 2. METHODE

Of een algemene vaardigheid op een school voldoende aan bod komt, is voor elk vak van elk opleidingstype van iedere school bepaald. In totaal gaat het om 5100 oordelen over de algemene vaardigheden in zestien vakken van 120 scholen (met een variabel aantal opleidingen per school). De inspecteurs hebben zich daarbij gebaseerd op het aanbod zoals dat bleek uit: de gebruikte methode(n), aanvullende onderwijsleermateriaal, vakwerkplannen, door vaksecties ingevulde vragenlijsten en gesprekken met de secties. Daarnaast hebben inspecteurs tijdens lesobservaties (zie het artikel 'Kwaliteit van het onderwijsleerproces' elders in dit themanummer) aangegeven of er in een les al dan niet aandacht voor een algemene vaardigheid geweest was.

Om inspecteurs in korte tijd tot een verantwoord oordeel te laten komen over de kwaliteit van het aanbod op een school is een duidelijke behoefte gebleken aan inzicht in de relatie tussen het onderwijsaanbod en de gebruikte methoden. Daarom is aan uitgeverijen gevraagd om de betreffende methoden te (laten) analyseren aan de hand van door de inspectie opgestelde criteria. Bij deze criteria is een belangrijke plaats ingeruimd voor de algemene vaardigheidsdoelen. Op basis van de door uitgeverijen aangeleverde methodenanalyses zijn de methoden vervolgens door onafhankelijke deskundigen (vakdidactici en onderwijskundigen) beoordeeld op hun basisvormingsgehalte en didactische kwaliteit. Wanneer in een methode bepaalde doelen onderbelicht zijn (er zijn bijvoorbeeld weinig opdrachten waarvoor de leerlingen onderzoek moeten doen (1), hun eigen mening moeten verwoorden (2), samen moeten werken (3), etc.) dan kunnen inspecteurs leraren gericht bevragen over het gebruik van aanvullende materialen. Voorts kunnen inspecteurs met behulp van deze informatie de consequenties voor het leerstofaanbod overzien wanneer leraren of vaksecties afwijken van de methode: onderdelen overslaan, hulpmiddelen niet gebruiken en dergelijke.

De inspecteurs hebben per algemene vaardigheid per combinatie van vak en opleidingstype aangegeven of deze 'voldoende', 'matig' of 'niet' in het aanbod voorkwamen dan wel 'niet van toepassing' zijn. Als alle leerlingen in een bepaalde opleiding alleen maar met een kerndoel 'in aanraking' zijn gebracht, is gesproken van een 'matig' aanbod van het kerndoel, als er sprake was van zodanige aandacht voor het kerndoel dat de leerlingen daardoor in staat gesteld werden het kerndoel te 'beheersen' is gesproken van een 'voldoende' aanbod. Bij de lesbezoeken hebben de inspecteurs aangegeven welke algemene vaardigheden zij in de les hebben waargenomen.

Voor de analyse is uitgegaan van het voldoende of niet-voldoende zijn van het aanbod in de algemene vaardigheden en het wel of niet voorkomen ervan in de klas. In beide gevallen gaat het om dichotome oordelen, en kan dus gebruik gemaakt worden van logitanalyse (waarbij de verklarende variabelen variëren al naar gelang de specifieke vraag/analyse. Vandaar dat de specifieke modellen weergegeven zijn in de betreffende tabellen). In zo'n logitanalyse wordt getoetst of een bepaald model voldoende past bij de data. Bij geneste modellen, waarvan hier sprake is, kan bovendien getoetst worden of toevoeging van factoren (verklarende variabele of combinatie van verklarende variabelen) aan het model tot een verbetering van de passing leidt.

### 3. RESULTATEN

#### *De algemene vaardigheden in het aanbod*

De inspecteurs hebben tijdens de evaluatie van de basisvorming het aanbod van de algemene vaardigheden per vak beoordeeld. In Tabel 1 is aangegeven welk percentage van de opleidingen de algemene vaardigheden voldoende aanbiedt.

Tabel 1: Percentage opleidingen dat de algemene vaardigheden voldoende aanbiedt.

| Algemene vaardigheid                          | Percentage |
|---|------------|
| 3 Samenwerken aan opdrachten                  | 54         |
| 2 Standpunten verwoorden                      | 43         |
| 1 Eenvoudig onderzoek doen                    | 36         |
| 4 Relaties met de beroepspraktijk             | 23         |
| 5 Eigen mogelijkheden en interesses ontdekken | 18         |
| 6 Criteria hanteren                           | 15         |
| Aantal combinaties van opleidingen en vak     | 5100       |

Over het geheel genomen is het oordeel van de inspecteurs niet positief. Slechts één vaardigheid ('Samenwerken aan opdrachten') wordt op meer dan de helft van de opleidingen voldoende aangeboden. 'Standpunten verwoorden' en 'Onderzoek doen' komen nog enigermate aan bod. Voor de twee OSB-vaardigheden en 'Criteria hanteren' is zelfs dat niet het geval: bij minder dan één op de vier vaardigheden is het aanbod voldoende.

Uitgangspunt voor de verdere analyse is een model waarin alleen de algemene vaardigheden als factor zijn opgenomen (Model 1). Voor het tweede model is de interactie van de vaardigheden met de (zestien) vakken toegevoegd. Het derde model omvatte ook de (vijf) opleidingen en het vierde model de interactie van opleidingen met de vaardigheden. De uitkomsten van de analyse zijn in Tabel 2 weergegeven.

Tabel 2: Toetsing van de verschillen tussen vakken en opleidingen.

| Model:<br>Logit (Freq) = Constante + | Passing  |     |             | Passingsverschil |    |             |
|--------------------------------------|----------|-----|-------------|------------------|----|-------------|
|                                      | $\chi^2$ | df  | $p(\chi^2)$ | $\chi^2$         | df | $p(\chi^2)$ |
| 1 Algemene vaardigheid (AV)          | 3959.3   | 474 | <.0001      |                  |    |             |
| 2 + Vak * AV                         | 695.2    | 384 | <.0001      | 3264.1           | 90 | <.0001      |
| 3 + Opleiding                        | 298.9    | 380 | .9992       | 396.4            | 4  | <.001       |
| 4 + Opleiding * AV                   | 256.1    | 360 | .9999       | 427              | 20 | .002        |

In Tabel 2 bevatten de eerste drie kolommen de toetsingsresultaten voor de modellen zelf. De laatste drie kolommen geven aan in hoeverre de toevoeging van een factor tot een verbetering leidt van de passing van het model aan de data.

Uit Tabel 2 blijkt duidelijk dat de toevoeging van de interactie van vak met algemene vaardigheden tot een aanzienlijke verbetering van de passing van het model leidt en dat vervolgens toevoegen van de opleiding de passing vrijwel perfect maakt. De toevoeging van de interactie van opleiding met de algemene vaardigheden leidt slechts tot een marginale passingsverbetering. Voor de interpretatie kunnen we dan ook uitgaan van Model 3. Er zijn dan twee tabellen nodig voor de beschrijving van de data: de percentages voldoende voor de algemene vaardigheden per vak (Tabel 3) en de percentages voldoende per opleiding over de vaardigheden heen (Tabel 4).

Tabel 3: Mate waarin de algemene vaardigheden voldoende aan bod komen in de vakken (percentages en gemiddeld aantal).

| Vak                    | Algemene vaardigheid |    |    |    |    |    | Aantal<br>vaardigheden |
|------------------------|----------------------|----|----|----|----|----|------------------------|
|                        | 3                    | 2  | 1  | 4  | 5  | 6  |                        |
| Verzorging             | 74                   | 67 | 55 | 22 | 18 | 26 | 2.6                    |
| Muziek                 | 70                   | 63 | 38 | 20 | 13 | 38 | 2.4                    |
| Natuur- en scheikunde  | 62                   | 37 | 62 | 39 | 36 | 4  | 2.4                    |
| Nederlands             | 61                   | 74 | 34 | 18 | 12 | 36 | 2.3                    |
| Geschiedenis           | 58                   | 66 | 47 | 8  | 11 | 12 | 2.0                    |
| Lichamelijke opvoeding | 90                   | 37 | 16 | 15 | 22 | 11 | 1.9                    |
| Beeldende vorming      | 21                   | 33 | 52 | 15 | 19 | 47 | 1.9                    |
| Biologie               | 40                   | 36 | 65 | 21 | 15 | 8  | 1.8                    |
| Aardrijkskunde         | 56                   | 37 | 45 | 20 | 17 | 7  | 1.8                    |
| Techniek               | 47                   | 21 | 38 | 20 | 17 | 20 | 1.6                    |
| Economie               | 37                   | 43 | 26 | 29 | 16 | 4  | 1.5                    |
| Wiskunde               | 46                   | 35 | 21 | 28 | 18 | 3  | 1.5                    |
| Duits                  | 51                   | 33 | 12 | 18 | 17 | 3  | 1.3                    |
| Frans                  | 51                   | 20 | 20 | 20 | 16 | 3  | 1.3                    |
| Engels                 | 49                   | 34 | 10 | 17 | 12 | 2  | 1.2                    |
| Informatiekunde        | 28                   | 18 | 11 | 41 | 15 | 1  | 1.1                    |
| Gemiddeld              | 54                   | 43 | 36 | 23 | 18 | 15 | 1.9                    |

Het percentage voldoende oordelen per vak per algemene vaardigheid is in Tabel 3 weergegeven. In de laatste kolom staat het gemiddeld aantal algemene vaardigheden dat binnen een vak 'voldoende' aan de orde komt. De vakken zijn op dit aantal geordend. Dat aantal varieert

sterk tussen de vakken. Het hoogste scoort Verzorging met ruim 2.6 algemene vaardigheden. Nog altijd minder dan de helft van de zes die mogelijk zijn. Ook de Kunstvakken, de Mens- en maatschappijvakken en de Exacte vakken scoren relatief hoog. De Talen, Wiskunde en Informatiekunde scoren laag: gemiddelde iets meer dan één algemene vaardigheid per vak.

De mate waarin de algemene vaardigheden in de vakken aan de orde komen varieert zeer sterk, van bijna 75 procent voor 'Samenwerken' (3) bij Verzorging tot een minimale één procent voor 'Criteria hanteren' (6) bij Informatiekunde. Interessant zijn vooral de vakken waar- bij het aanbod van de algemene vaardigheden aanzienlijk afwijken van het algemene patroon zoals dat uit de laatste rij blijkt.

Het ligt voor de hand dat bij Natuur- en scheikunde (net als bij Biologie) de algemene vaardigheid 'Onderzoek doen' (1) vaak aan bod komt. Dat ligt minder voor de hand bij 'Relatie met de beroepspraktijk' (4) en 'Eigen mogelijkheden en interesses ontdekken' (5). Daar staat weer tegenover dat 'Criteria hanteren' (6) bij Natuur- en scheikunde nauwelijks voorkomt.

Bij Geschiedenis en Nederlands komen 'Standpunten verwoorden' (2) relatief vaak voor. Voor Nederlands is dat meer vanzelfsprekend dan voor Geschiedenis. Ook voor de hand liggend zijn de hoge scores voor 'Samenwerken aan opdrachten' (3) bij Lichamelijke opvoeding en 'Criteria hanteren' (6) bij Beeldende vorming.

Bij Informatiekunde is de aandacht voor de 'Relatie met de beroepspraktijk' (4) relatief hoog. Dit is des te opvallender omdat er over dit vak verder nauwelijks iets positiefs gezegd kan worden.

Uit de analyse die in Tabel 2 is gepresenteerd is gebleken dat er tussen de opleidingen een verschil is in het aanbod van de algemene vaardigheden en dat dit verschil voor alle algemene vaardigheden hetzelfde is. Tabel 4 bevat de percentages voldoende per opleiding. Er is een duidelijk verloop van de mate waarin de algemene vaardigheden voldoende aan bod komen: van één op de drie voor HAVO en VWO tot één op de vier bij (I)VBO. De knip ligt tussen MAVO en (I)VBO. We kunnen de conclusie trekken dat het aanbod in de algemene vaardigheden in het beroepsonderwijs van minder kwaliteit is dan in het algemeen vormend onderwijs. Echter, al met al blijken de algemene vaardigheden op alle opleidingstypen onvoldoende aandacht te krijgen.

Tabel 4: Mate waarin de algemene vaardigheden voldoende aan bod komen in de opleidingen.

| Opleiding | Percentage |
|-----------|------------|
| VWO       | 35         |
| HAVO      | 35         |
| MAVO      | 32         |
| VBO       | 25         |
| IVBO      | 23         |
| Gemiddeld | 32         |

#### Vakoverstijgend aanbod

De algemene vaardigheden zijn bedoeld als vaardigheden die gemeenschappelijk zijn in de vakken. Een leerling moet dus dezelfde vaardigheid in verschillende vakken in voldoende mate tegenkomen. Om dat na te gaan, hebben we bij elke combinatie van opleiding en school bepaald in hoeverre een vaardigheid 'opleidingsbreed' voorkomt.



Tabel 5: Verdeling over het aantal vakken waarin de algemene vaardigheden in een opleiding voldoende aan bod komen (percentages).

| Algemene vaardigheid                | Aantal vakken |    |    |     |      |       |
|-------------------------------------|---------------|----|----|-----|------|-------|
|                                     | 0             | 1  | 2  | 3-5 | 6-10 | 11-16 |
| 3 Samenwerken aan opdrachten        | 0             | 1  | 1  | 17  | 58   | 23    |
| 2 Standpunt verwoorden              | 1             | 1  | 3  | 31  | 56   | 8     |
| 1 Eenvoudig onderzoek doen          | 2             | 2  | 8  | 39  | 46   | 3     |
| 4 Relatie beroepspraktijk           | 5             | 12 | 16 | 51  | 15   | 0     |
| 5 Eigen mogelijkheden en interesses | 13            | 18 | 17 | 42  | 9    | 0     |
| 6 Hanteren criteria                 | 9             | 26 | 25 | 37  | 2    | 0     |

De algemene vaardigheden vormen het gemeenschappelijke element in het curriculum. Een minimale eis is dan dat een algemene vaardigheid in minstens twee vakken voldoende aan bod moet komen, wil er van het 'algemene' in de algemene vaardigheden ook maar iets terecht kunnen komen. Duidelijk is dat er voor elke vaardigheid opleidingen zijn waarvoor aan deze minimumeis niet wordt voldaan. Voor de eerste vier vaardigheden is het percentage opleidingen waarbij deze algemene vaardigheden bij minder dan twee vakken aan bod komt nog laag (bij voorbeeld voor de algemene vaardigheid 'Samenwerken aan opdrachten' (3) voldoet slechts één procent van de opleidingen niet aan deze eis). Maar voor de vaardigheid 'Eigen mogelijkheden en interesses' (5) en 'Criteria hanteren' (6) voldoet ruim 30 procent van de opleidingen niet aan deze uiterst terughoudende norm.

Als een strengere eis wordt gesteld aan het gemeenschappelijk voorkomen van de algemene vaardigheden, bijvoorbeeld het voldoende voorkomen in ten minste zes van de zestien in basisvormingsvakken die in de analyse zijn meegenomen, dan is het resultaat beduidend slechter. Voor de vaardigheid 'Samenwerken aan opdrachten' (3) voldoet ruim tachtig procent van de opleidingen aan het criterium. Ook 'Standpunten verwoorden' (2) scoort met bijna vijfenzeventig procent vrij hoog. Bij de andere vaardigheden wordt het vakoverstijgende karakter van de algemene vaardigheden volgens dit criterium op maar weinig opleidingen in voldoende mate gerealiseerd.

Bij de beoordeling van de algemene vaardigheden als geheel op een school is de inspectie uitgegaan van de zo-even genoemde minimale norm van een voldoende aanbod bij ten minste twee vakken per algemene vaardigheid (Inspectie van het Onderwijs, 1999). Het oordeel is voldoende als voor elke opleiding in een school alle algemene vaardigheden in twee of meer vakken aan bod komen. Bij deze norm scoort 43 procent van de scholen voldoende. Als de strengere norm van zes vakken per opleiding wordt gehanteerd, is er geen enkele school waarvoor het aanbod van de algemene vaardigheden voldoende is.

#### Ondersteuning in de methoden

In het Nederlandse onderwijs volgen de meeste leraren trouw de methode (Inspectie van het Onderwijs, 1999). Als de algemene vaardigheden niet of niet voldoende in de methode voorkomen, moeten leraren daar zelf vorm aan geven. Voor de meeste algemene vaardigheden is dat geen sinecure. Om te zien in welke mate de algemene vaardigheden het aanbieden van de algemene vaardigheden ondersteunen zijn de methoden van de meest voorkomende vakken geanalyseerd. Deze analyse is, zoals eerder opgemerkt, in samenwerking met de Groep Educatieve Uitgeverijen (GEU) uitgevoerd door vakdeskundigen onder supervisie van de Inspectie (zie ook het artikel 'Kwaliteit van het vakspecifieke aanbod' elder in dit themanummer).

Tabel 6 bevat de resultaten van deze beoordeling. Aangegeven is het percentage algemene vaardigheden dat in de methode voorkomt, gemiddeld over de beoordeelde methodes. Daarbij is rekening gehouden met de mate waarin de methodes in feite gebruikt worden.

Tabel 6: *Voldoende aanwezigheid van de algemene vaardigheden in de methoden.*

|                        | Percentage |
|------------------------|------------|
| Muziek                 | 94         |
| Verzorging             | 94         |
| Economie               | 47         |
| Natuur- en scheikunde  | 46         |
| Beeldende vorming      | 43         |
| Aardrijkskunde         | 36         |
| Techniek               | 21         |
| Geschiedenis           | 21         |
| Lichamelijke opvoeding | 17         |
| Nederlands             | 16         |
| Biologie               | 13         |
| Informatiekunde        | 13         |
| Duits                  | 9          |
| Engels                 | 0          |
| Wiskunde               | 0          |
| Frans                  | 0          |

Bij slechts twee vakken, Muziek en Verzorging, zijn de algemene vaardigheden in alle methodes in ruime mate aanwezig. Bij vier vakken, de Moderne Vreemde Talen en Wiskunde, komen bij minder dan tien procent van de methodes de algemene vaardigheden voldoende aan bod. De volgorde waarin de vakken zijn opgenomen in deze tabel komt in hoge mate overeen met de mate waarin de algemene vaardigheden volgens de inspecteurs voldoende aan bod komen. Ten dele vloeit dat voort uit het feit dat deze informatie gebruikt is bij de beoordeling van het aanbod. De conclusie blijft echter gerechtvaardigd dat er een samenhang is tussen het voorkomen in de methoden en het realiseren van de algemene vaardigheden.

#### *Algemene vaardigheden in de lessen*

Het aanbod van de algemene vaardigheden moet eigenlijk blijken in de onderwijssituatie. Vandaar dat de inspecteurs tijdens de lesbezoeken het voorkomen van de algemene vaardigheden in de les hebben geregistreerd. In 7218 van de 7268 lessen was hierover informatie aanwezig. Gemiddeld genomen komt in één op de zes lessen één van de algemene vaardigheden aan bod. Voor een voldoende aanbod in de hele periode van de basisvorming is het natuurlijk niet nodig dat er in elke les aandacht aan wordt besteed, maar één algemene vaardigheid in één op de zes lessen is wel erg mager.

De meeste aandacht is voor 'Samenwerken aan opdrachten' (3; ruim eenderde van de lessen) en het minst voor algemene vaardigheden die met OSB samenhangen (een op de twintig). In Tabel 7 staan de precieze percentages per algemene vaardigheid. De volgorde van waarin er aandacht wordt besteed aan de algemene vaardigheden komt redelijk overeen met de kwaliteit van het aanbod zoals dat in Tabel 1 is weergegeven.

Het algemene beeld van Tabel 7 kan gedifferentieerd worden door bestudering van verschillen tussen vakken, samenstelling van de klas en leerjaren. Ook het voorkomen in de verschillende typen lessen kan van belang zijn. Immers, in een practicumles zal eerder onderzoek gedaan worden dan in een klassikale les. Er zijn drie typen lessen onderscheiden: practicumlessen, lessen waarin de leerlingen hoofdzakelijk zelfwerkzaam waren, en klassikale lessen.

Tabel 7: Mate waarin de algemene vaardigheden in de lessen aan bod komen.

| Algemene vaardigheid                    | Percentage |
|---|------------|
| 3 Samenwerken aan opdrachten            | 37         |
| 2 Standpunten verwoorden                | 22         |
| 1 Eenvoudig onderzoek doen              | 16         |
| 5 Mogelijkheden en interesses ontdekken | 6          |
| 6 Criteria hanteren                     | 5          |
| 4 Relaties met de beroepspraktijk       | 3          |
| Aantal lessen                           | 7218       |

Op dezelfde wijze als voor de verschillen tussen vakken en opleidingen (Tabel 2) kan door een logitanalyse nagegaan worden welk model het best past bij de data. De uitkomsten van deze analyse staat in Tabel 8.

Tabel 8: Toetsing van modellen op de algemene vaardigheden in de lessen.

| Model:                      | Passing  |      |             | Passingsverschil |    |             |
|-----------------------------|----------|------|-------------|------------------|----|-------------|
|                             | $\chi^2$ | df   | $p(\chi^2)$ | $\chi^2$         | df | $p(\chi^2)$ |
| Logit (Freq) = Constante +  |          |      |             |                  |    |             |
| 1 Algemene vaardigheid (AV) | 8772.2   | 3600 |             |                  |    |             |
| 2 + Vak * AV                | 5002.3   | 3510 | <.0001      | 3769.92          | 90 | <.0001      |
| 3 + Lestype * AV            | 4042.5   | 3498 | <.0001      | 959.85           | 12 | <.0001      |
| 4 + Klas * AV               | 3810.4   | 3468 | <.0001      | 232.06           | 30 | <.0001      |
| 5 + Leerjaar * AV           | 3773.2   | 3456 | .0001       | 37.20            | 12 | .0002       |

In Tabel 8 staan vijf modellen. Uitgaande van een model waarbij alleen de algemene vaardigheden zijn opgenomen (1), is steeds een variabele en de interactie van die variabele met de algemene vaardigheden toegevoegd. In de eerste drie kolommen staan de  $\chi^2$ -waarden en de bijbehorende vrijheidsgraden, waarmee getoetst kan worden of de modellen een adequate beschrijving geven van de data. Geen van de modellen past in voldoende mate. De overschrijdingskans bij het laatste model is nog steeds te klein ( $p = .0001$ ). Dat houdt echter niet in dat alle factoren even belangrijk zijn, zoals de laatste drie kolommen laten zien. Die kolommen geven aan in welke mate toevoeging van een factor tot verbetering van de passing leidt. Voor vak, lestype en klas (sensamenstelling) is die verbetering substantieel; voor leerjaar is de verbetering aanzienlijk kleiner. Voor de interpretatie gaan we daarom uit van Model 4. We hebben dan drie tabellen nodig om de data te beschrijven: Tabel 9 waarin een onderscheid gemaakt wordt tussen vakken, Tabel 10 waarin de relatie tussen de algemene vaardigheden en de samenstelling van de klassen aan bod komen, en Tabel 11 waarin de combinatie tussen de algemene vaardigheden en het lestype weergegeven is.

Zoals Tabel 9 laat zien, zijn er grote verschillen tussen de vakken. De laatste kolom geeft het gemiddeld aantal vaardigheden dat in een les aan de orde komt. Voor vijf vakken is dat één of meer, voor drie vakken is dat een half of minder. Wellicht ten overvloede zei opgemerkt dat het weliswaar niet nodig is dat alle algemene vaardigheden in alle lessen op enigerlei wijze aan bod komen, maar over het geheel genomen moet het aanbod in de lessen toch substantieel zijn. Merk op dat de volgorde van de vakken redelijk overeen komt met de volgorde in het aanbod (Tabel 3).

Tabel 9: Mate waarin de algemene vaardigheden in de lessen aan de orde komen (percentages) en gemiddeld aantal vaardigheden per les.

| Vak                    | Algemene vaardigheid |    |    |    |    |    | Aantal<br>vaardigheden |
|------------------------|----------------------|----|----|----|----|----|------------------------|
|                        | 3                    | 2  | 1  | 5  | 6  | 4  |                        |
| Techniek               | 40                   | 16 | 41 | 35 | 18 | 16 | 1.7                    |
| Verzorging             | 56                   | 48 | 27 | 7  | 11 | 4  | 1.5                    |
| Natuur- en scheikunde  | 45                   | 30 | 51 | 11 | 1  | 4  | 1.4                    |
| Muziek                 | 54                   | 35 | 14 | 7  | 21 | 3  | 1.3                    |
| Lichamelijke opvoeding | 75                   | 12 | 3  | 11 | 3  | 1  | 1.0                    |
| Geschiedenis           | 32                   | 41 | 11 | 8  | 1  | 1  | .9                     |
| Biologie               | 32                   | 12 | 32 | 4  | 2  | 2  | .8                     |
| Beeldende vorming      | 8                    | 19 | 26 | 9  | 17 | 3  | .8                     |
| Wiskunde               | 36                   | 29 | 3  | 3  | 1  | 2  | .7                     |
| Aardrijkskunde         | 36                   | 18 | 10 | 2  | 2  | 2  | .7                     |
| Nederlands             | 29                   | 22 | 5  | 2  | 8  | 2  | .7                     |
| Economie               | 22                   | 31 | 5  | 2  | 0  | 4  | .6                     |
| Engels                 | 41                   | 13 | 1  | 1  | 1  | 1  | .6                     |
| Informatiekunde        | 19                   | 15 | 5  | 7  | 0  | 3  | .5                     |
| Duits                  | 36                   | 7  | 1  | 1  | 1  | 1  | .5                     |
| Frans                  | 31                   | 6  | 1  | 1  | 1  | 1  | .4                     |
| Gemiddeld              | 37                   | 22 | 14 | 6  | 5  | 3  | .9                     |

Als we naar de uitschieters in de tabel kijken dan valt Techniek het meest op. Voor dit vak is er een redelijk uniforme verdeling over de vaardigheden: alle algemene vaardigheden komen in redelijke mate voor. Met name de vaardigheden die bij de andere vakken nauwelijks aan bod komen, doen het goed bij techniek. Verzorging, een ander vak dat met de basisvorming is ingevoerd, scoort gemiddeld ook vrij hoog. De verdeling over de algemene vaardigheden komt echter meer overeen met dat van de andere vakken.

Bij de vakken die laag scoren staan alle Moderne Vreemde Talen. Ook Informatiekunde en Economie staan onderin de tabel. Ook dit zijn vakken die met de basisvorming verplicht in de eerste leerjaren zijn ingevoerd. Kennelijk is het bij deze vakken niet gelukt om een redelijk basisvormingsgehalte te realiseren.

In Tabel 10 staan de gegevens over het voorkomen van de algemene vaardigheden in klassen van verschillende samenstelling. Zoals op grond van de kwaliteit van het aanbod verwacht mocht worden (zie Tabel 4) daalt de aandacht van HAVO/VWO-klassen via MAVO naar VBO- en IVBO-klassen. Breed samengestelde klassen (MAVO/HAVO/VWO al dan niet met VBO) nemen een middenpositie in. Als we kijken naar de afzonderlijke vaardigheden, dan vertonen ze ongeveer dezelfde verdeling over de klassen. Een uitzondering vormen de OSB-standaardtaarden: in IVBO- en in mindere mate in VBO-klassen komen 'Relatie met de beroepspraktijk' (4) en 'Eigen mogelijkheden en interesses ontdekken' (5) relatief vaker voor dan in de klassen met overwegend AVO-leerlingen.

In Tabel 11 staat voor de verschillende typen lessen het percentage waarin de algemene vaardigheden aan de orde komen. Zoals te verwachten, komen bij practicumlessen de algemene vaardigheden veel meer aan bod dan bij klassikale lessen. Het meest uitgesproken geldt dat voor 'Onderzoek doen' (1): in meer dan de helft van de practicumlessen tegenover in één op de zestien klassikale lessen. Ook 'Eigen mogelijkheden en interesses ontdekken' (5) en 'Criteria hanteren' (6) komen bij practicumlessen substantieel vaker voor.

Tabel 10: Mate waarin de algemene vaardigheden in klassen van verschillende samenstelling aan de orde komen (percentage) en het gemiddeld voorkomen per les.

| Klassensamenstelling | Algemene vaardigheid |    |    |    |   |   | Aantal vaardigheden |
|----------------------|----------------------|----|----|----|---|---|---------------------|
|                      | 3                    | 2  | 1  | 5  | 6 | 4 |                     |
| VWO                  | 46                   | 25 | 17 | 7  | 5 | 1 | 1.0                 |
| HAVO; HAVO/VWO       | 42                   | 25 | 16 | 7  | 6 | 2 | 1.0                 |
| MAVO; MAVO/HAVO      | 37                   | 21 | 13 | 6  | 5 | 3 | .9                  |
| (VBO)MAVO/HAVO/VWO   | 34                   | 21 | 11 | 7  | 5 | 3 | .8                  |
| IVBO                 | 27                   | 20 | 12 | 10 | 4 | 4 | .8                  |
| VBO; VBO/MAVO        | 30                   | 18 | 13 | 6  | 4 | 3 | .7                  |
| Gemiddeld            | 37                   | 22 | 14 | 6  | 5 | 3 | .9                  |

Lessen waarin zelfwerkzaamheid van de leerlingen centraal staan, nemen een middenpositie in tussen practicumlessen en klassikale lessen. De uitzondering vormt 'Samenwerken' (3), dat in een les waarin zelfwerkzaamheid centraal staat even vaak voorkomt als in een practicumles.

Tabel 11: Mate van voorkomen van de algemene vaardigheden in verschillende typen lessen (percentage) en gemiddelde voorkomen per les.

| Type les         | Algemene vaardigheid |    |    |    |    |   | Aantal vaardigheden |
|------------------|----------------------|----|----|----|----|---|---------------------|
|                  | 3                    | 2  | 1  | 5  | 6  | 4 |                     |
| Practicum        | 45                   | 20 | 52 | 14 | 11 | 4 | 1.5                 |
| Zelfwerkzaamheid | 46                   | 22 | 16 | 7  | 6  | 3 | 1.0                 |
| Klassikaal       | 31                   | 22 | 6  | 5  | 4  | 2 | .7                  |
| Gemiddeld        | 37                   | 22 | 14 | 6  | 5  | 3 | .9                  |

#### 4. CONCLUSIES

Het aanbod in de algemene vaardigheden is over het geheel genomen onvoldoende. 'Eenvoudig onderzoek doen', 'Standpunten verwoorden' en 'Samenwerken aan opdrachten' komen nog enigermate in het aanbod voor. Voor 'Relatie leggen met de beroepspraktijk', 'Eigen mogelijkheden en interesses ontdekken' en 'Criteria hanteren' is dat nauwelijks het geval.

Per vak zijn er afwijkingen in dit patroon. Verzorging, Muziek en Natuur- en scheikunde zijn vakken waarin de algemene vaardigheden relatief veel in het aanbod voorkomen. In het aanbod van de Moderne Vreemde Talen en Informatiekunde komen algemene vaardigheden slechts sporadisch voor.

De kwaliteit van het aanbod tussen de opleidingen loopt uiteen. Met name in het (I)VBO is het aanbod duidelijk minder van kwaliteit in vergelijking met de AVO-opleidingen.

Algemene vaardigheden zijn vaardigheden die gemeenschappelijk zijn in de vakken en dus minstens in twee, maar eigenlijk in meerdere vakken voor horen te komen. Dit is lang niet altijd het geval. Alleen 'Samenwerken aan opdrachten' komt bij tachtig procent van de opleidingen bij eenderde of meer van de vakken voor. Bij het 'Hanteren van criteria' is dat bij slechts twee procent van de opleidingen het geval.



Als een minimalistisch criterium wordt aangelegd: in alle opleidingen van een school moeten de algemene vaardigheden in twee op meer vakken aan bod komen, scores slechts 43 procent van de scholen voldoende. Worden strengere criteria aangelegd, dan daalt dit percentage aanzienlijk.

Het aanbod wordt voor een aanzienlijk deel bepaald door de methodes. Slechts voor twee vakken komen alle algemene vaardigheden in de meest gangbare methoden voldoende voor. Voor de Moderne Vreemde Talen en Wiskunde komen de algemene vaardigheden nauwelijks in de methoden voor. Voor een deel kan dit een verklaring zijn voor het matige aanbod van de algemene vaardigheden.

Het aanbod moet gerealiseerd worden in de lessen. Gemiddeld komt in één op de zes lessen een algemene vaardigheid aan bod. Een niet al te hoge score. Het voorkomen in de verschillende vakken en in klassen van verschillende samenstelling komt in grote lijnen overeen met voorkomen in het aanbod in het algemeen.

Een belangrijk aspect van de basisvorming, de modernisering van het aanbod door de algemene vaardigheidsdoelen, is niet goed van de grond gekomen. In het aanbod komen de vaardigheden onvoldoende over meerdere vakken aan de orde. In de methoden zijn ze voor de meeste vakken onvoldoende verdisconteerd en in de lessen zijn ze maar mondjesmaat zichtbaar.

## LITERATUUR

- Agresti, A. (1990). *Categorical Data Analysis*. New York: Wiley.
- Inspectie van het Onderwijs (1999). *Werk aan de basis. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het onderwijs.
- Procesmanagement Basisvorming (1992). *Ontwerp kerndoelen voor de basisvorming in basisonderwijs en voortgezet onderwijs. Inforeeks basisvorming 2*. PMB.

# Kwaliteit van de opbrengsten: Worden de kerndoelen gehaald?

Frans Kleintjes, *Cito*

Michel Zwarts, *Inspectie van het Onderwijs*

Marielle Peters-Sips, *Inspectie van het Onderwijs*

## ABSTRACT

*The evaluation study reported in this article focusses on the results of the pupils in the 'Basisvorming' (basic secondary education). Aim of the study was to assess the achievement of pupils after the implementation of a common core curriculum. The results of 64.500 pupils are used. At least 200 pupils were involved in the testing of each of the 50 domains for 17 subjects. Teachers and other experts judged the results. They found the results for the languages 'good', for the humanities 'sufficient', but for the 'new' subjects 'insufficient'. They also found the results for the 'lower' types of secondary education better than the results for the 'higher' types.*

## 1. INLEIDING

'Kwaliteit moet blijken.' Het onderzoek naar de opbrengsten maakt daarom deel uit van de evaluatie van de basisvorming (zie het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer). In de opzet van de evaluatie zijn de opbrengsten op twee manieren meegenomen: longitudinaal zijn de resultaten op enkele onderdelen van het leerplan onderzocht, éénmalig de resultaten van vrijwel alle onderdelen van het leerplan.

De cohortonderzoeken worden al vanaf het schooljaar 1989/1990 uitgevoerd (zie het artikel 'Kwaliteit van de opbrengsten: Zijn de prestaties bij Tekstbegrip Nederlands en Wiskunde vooruitgegaan?' elders in dit themanummer en Van der Werf, Lubbers & Kuypers, 1999). In deze cohortstudies worden de ontwikkelingen in de tijd onderzocht.

In de periode voorjaar 1997 - voorjaar 1999 is een afzonderlijk onderzoek gedaan naar de mate waarin de 'domeinen', waarin de kerndoelen zijn ondergebracht, bereikt worden. Over dit laatste onderzoek wordt in dit artikel gerapporteerd.

De evaluatievraag die in dit artikel wordt beantwoord, luidt: 'Beheersen voldoende leerlingen aan het eind van de basisvorming de kerndoelen?' Met 'beheersen' wordt hier bedoeld wat van een leerling verwacht mag worden die onderwijs heeft gehad op basis van de kerndoelen en onderwijs wil vervolgen in één van de opleidingstypen in het VO.

Formeel is de duur van de basisvorming drie jaar, maar het bevoegd gezag kan met het oog op het belang van leerlingen de basisvorming zodanig inrichten dat leerlingen de kerndoelen in ten minste twee en ten hoogste vier leerjaren kunnen bereiken. De basisvorming mag per vak worden afgesloten als de leerstof die nodig is voor de beheersing van de kerndoelen aan de orde is geweest en een toets basisvorming is afgenomen. Uit het moment waarop de toetsen basisvorming worden afgenomen (Inspectie van het Onderwijs, 1997b, Kuhlemeier 1997) is duidelijk dat voor de meeste vakken de basisvorming eind leerjaar twee of leerjaar drie wordt afgesloten, maar dat er grote verschillen zijn per vak en per opleiding (zie ook het artikel: 'Kwaliteit van het vakspecifieke aanbod' elders in dit themanummer).

De kerndoelen die golden voor de periode van de evaluatie (1993 - 1998) zijn vastgelegd in het Besluit Kerndoelen en Adviesurentabel Basisvorming 1993 - 1998. De kerndoelen bestaan uit vakspecifieke doelen voor de vakken die in de periode van de basisvorming gegeven worden en algemene vaardigheidsdoelen die niet in elk vak afzonderlijk maar binnen het gehele aanbod van vakken tezamen moeten worden gerealiseerd (zie de artikelen: 'Kwaliteit van het vakspecifieke aanbod' en 'Kwaliteit van het vakoverstijgende aanbod: het voorkomen van de algemene vaardigheden' elders in dit themanummer).

Om de evaluatievraag te beantwoorden was het nodig om de prestaties van de leerlingen te meten en met behulp van een adequate procedure te bepalen of deze prestaties voldoende zijn.

Voor het meten van de prestaties van de leerlingen is gebruik gemaakt van gegevens van de door het Cito ontwikkelde (proef)toetsen basisvorming<sup>1</sup>. Voor de beoordelingsprocedure is gebruik gemaakt van de methodiek zoals die voor peilingsonderzoeken van het Cito is ontwikkeld.

In de volgende paragrafen werken we de opzet voor deze twee aspecten afzonderlijk uit.

## 2. OPZET VOOR HET METEN VAN DE PRESTATIES VAN DE LEERLINGEN

### *Keuze van de opgaven*

Voor het meten van de beheersing van de kerndoelen is gebruik gemaakt van de opgaven die in het voorjaar van 1997 en 1998 zijn beproefd. Deze proeftoetsen omvatten samen vrijwel alle domeinen van de kerndoelen. De proeftoetsen zijn gebruikt in plaats van de definitieve versies van de afsluitende toetsen omdat de proeftoetsen onder gecontroleerde condities zijn afgenomen en een groter aantal opgaven bevatten.

De praktijktoetsen voor de Expressievakken en de toetsen over omgevingsonderwijs bij Aardrijkskunde en Geschiedenis zijn niet in de beoordeling van de opbrengsten meegenomen omdat het scoringsvoorschrift alleen een vergelijking binnen scholen en niet tussen scholen mogelijk maakt.

De algemene vaardigheden maken als zodanig geen onderdeel uit van de afsluitende toetsen basisvorming<sup>2</sup>. Aspecten van de algemene vaardigheden zijn wel in de afsluitende toetsen zijn opgenomen, onder meer in de praktijk- en practicumtoetsen.

Als eenheid van beoordeling werd gekozen voor de domeinen van de kerndoelen. Uitspraken over de beheersing van de afzonderlijke kerndoelen zijn voor de beoordeling van de opbrengsten te gedetailleerd, terwijl het niveau van het vak te grofmazig is. In totaal waren van 50 domeinen en 17 vakken opgaven uit de proeftoetsen beschikbaar (Tabel 1).

### *Verzameling van de gegevens*

De proefafnames werden gehouden bij een steekproef van klassen die de basisvorming voor het vak in kwestie afsluiten. Ieder najaar stuurt het Cito een vragenlijst naar de scholen met het verzoek per klas aan te geven voor welke vakken de basisvorming in het voorjaar van het betreffende schooljaar wordt afgesloten. Deze klassen komen in aanmerking om mee te doen met de proefafname voor het vak dat ze afsluiten, ze vormen het steekproefkader waaruit de benodigde klassen voor een bepaald vak worden getrokken.

Van elke klas geeft de school ook het opleidingstype aan. In dit artikel zijn drie 'BAVO-categorieën' onderscheiden: VBO +, MAVO+ en HAVO+. De + geeft aan dat de categorie zowel categoriale als heterogene opleidingstypen bevat. De betekenis is als volgt:

VBO+: alle combinaties met (I)VBO: IVBO, VBO, IVBO/VBO, (I)VBO/MAVO, (I)VBO/MAVO/HAVO en (I)VBO/MAVO/HAVO/VWO;

MAVO+: MAVO en MAVO/HAVO;

HAVO+: MAVO/HAVO/VWO, HAVO/VWO, HAVO en VWO.

In het vervolg worden deze categorieën aangeduid met (I)VBO, MAVO en HAVO/VWO.

De groepering van de opleidingstypen in de BAVO-categorieën berust op een indeling op grond van de scores op de Eindtoets Basisonderwijs. De scores van de leerlingen op de Eindtoets Basisonderwijs ondersteunen een indeling in deze drie BAVO-categorieën (Uiterwijk, 1998).

De steekproefomvang voor een bepaald vakonderdeel werd enerzijds bepaald door het aantal te toetsen items en anderzijds door nauwkeurigheidseisen die men moet stellen in verband met het beoogde gebruik voor evaluatiedoeleinden. Zowel de itemkenmerken als de populatiekenmerken moeten voldoende nauwkeurig geschat worden. Per BAVO-categorie werden ongeveer 200 leerlingantwoorden per item verzameld. Per item zijn er dus 600 leerlingantwoorden. Tevens is er voor gezorgd dat minimaal 500 leerlingen per BAVO-categorie per vakonderdeel aan de proefafname deelnamen om populatiekenmerken voldoende nauwkeurig te kunnen schatten. Gemiddeld deden er per vakonderdeel 1850 leerlingen mee aan de proefafnames.

De proefafnames voor Aardrijkskunde, Wiskunde, Natuur-en scheikunde, Nederlands en Muziek waren met ongeveer 4000 leerlingen het grootst. In totaal zijn de prestaties van ongeveer 64.500 leerlingen uit ongeveer 3000 klassen gebruikt voor de opbrengstmeting. De omvang van de proefafnames van de spreektoetsen Moderne vreemde talen was met ongeveer 200 leerlingen het kleinst. Een gedetailleerd overzicht staat in Tabel 1.

Om de productieve taalvaardigheden, Spreken en Schrijven, te beoordelen is het nodig om over een selectie leerlingproducten te beschikken. Voor de schrijftaken worden standaard de leerlingproducten naar het Cito gestuurd. Voor de spreektaken is dit niet het geval. Het was dus nodig om voor de spreektaken een aanvullende gegevensverzameling op te zetten. Deze gegevens werden verzameld op de scholen die betrokken zijn bij de evaluatie van het onderwijsleerproces (zie het artikel: 'Kwaliteit van het onderwijsleerproces' elders in dit themanummer).

### *Analyse*

Voor de gekozen beoordelingsprocedure is het nodig dat de opgaven van de domeinen geschaald worden en dat de gemiddelde vaardigheid van de leerlingen op dezelfde schaal weergegeven kan worden. Als analysemethode is gekozen voor het itemresponsemodel, het One Parameter Logistic Model (OPLM; Verhelst, Glas & Verstralen, 1995). Dit model is een uitbreiding van het Raschmodel, waarbij voor elke opgave een discriminatie-index gespecificeerd kan worden. Met het model kan getoetst worden of alle opgaven voor de betrokken populatie op één schaal liggen, dat wil zeggen één vaardigheid meten. Is dit niet het geval, dan kan geprobeerd worden door wijziging van de discriminatie-index, weglaten van opgaven of door splitsing van de populatie tot één schaal te komen.

Een analyse met OPLM levert voor elke opgave een schatting op van zijn positie op de schaal (de moeilijkheid). Voor de leerlingen levert de analyse een schatting op van het gemiddelde en de variantie van de verdeling van vaardigheid per BAVO-categorie.

Over het algemeen was het mogelijk om de items uit de proefafnames te schalen, waarbij het hier en daar nodig was om een opgave te verwijderen. Voor een beperkt aantal toetsen was het nodig om onderscheid te maken in de deelpopulaties (I)VBO, MAVO en HAVO/VWO. Dit betrof het onderdeel theorie voor Informatiekunde, de onderscheiden domeinen van de Beeldende vakken en Techniek.

Tabel 1: *Overzicht van de domeinen (zijnde de rapportage-eenheden) per vak.*

| Vak                   | Rapportage-eenheid                      | Aantal leerlingen | Aantal opgaven       |
|-----------------------|---|-------------------|----------------------|
| Nederlands            | Kijken en luisteren                     | 1289              | 22                   |
|                       | Lezen                                   | 3709              | 37                   |
|                       | Schrijven                               | 1878              | product per leerling |
| Engels                | Lezen                                   | 2111              | 26                   |
|                       | Luisteren                               | 1568              | 22                   |
|                       | Spreken                                 | 253               | product per leerling |
| Frans                 | Schrijven                               | 1464              | product per leerling |
|                       | Lezen                                   | 1255              | 27                   |
|                       | Luisteren                               | 1883              | 24                   |
|                       | Spreken                                 | 186               | product per leerling |
| Duits                 | Schrijven                               | 901               | product per leerling |
|                       | Lezen                                   | 1450              | 30                   |
|                       | Luisteren                               | 1733              | 22                   |
|                       | Spreken                                 | 206               | product per leerling |
| Wiskunde              | Schrijven                               | 1438              | product per leerling |
|                       | B Rekenen, meten en schatten            | 4388              | 21                   |
|                       | C Algebra/verbanden/grafieken/funcities | 4388              | 20                   |
|                       | D Meetkunde                             | 4388              | 20                   |
|                       | E Statistiek en kans                    | 4388              | 21                   |
| Natuur- en scheikunde | B + H Stoffen en materialen             | 3725              | 14                   |
|                       | C + D Electriciteit/energie/verbranding | 3725              | 19                   |
|                       | E + F Licht en geluid                   | 3725              | 17                   |
|                       | G + A Krachten en veiligheid            | 3725              | 16                   |
|                       | Practicum                               | 1342              | 21                   |
|                       | B De mens                               | 2814              | 23                   |
| Biologie              | C Planten en dieren                     | 2814              | 23                   |
|                       | A Informatievaardigheden                | 2171              | 23                   |
|                       | Practicum                               | 1122              | 17                   |
|                       | Theorie                                 | 2104              | 19                   |
| Techniek              | Praktijk                                | 2104              | 17                   |
|                       | C Stedelijk en landelijk                | 4058              | 24                   |
| Aardrijkskunde        | E Multiculturele samenleving            | 4058              | 15                   |
|                       | F Natuur en milieu                      | 4058              | 27                   |
|                       | G Internationale ontwikkelingen         | 4058              | 26                   |
|                       | Geschiedenis                            | 1304              | 16                   |
| Geschiedenis          | E Staatsinrichting                      | 1304              | 22                   |
|                       | B Consumptie                            | 3275              | 25                   |
|                       | C Arbeid en productie                   | 3275              | 20                   |
|                       | D Overheid en bestuur                   | 3275              | 16                   |
|                       | F Internationale ontwikkelingen         | 3275              | 25                   |
| Informatiekunde       | Praktijk                                | 1805              | 21                   |
|                       | Theorie                                 | 523               | 22                   |
| Verzorging            | A Leren over zorg                       | 1808              | 27                   |
|                       | B Leren zorgen                          | 1808              | 25                   |
|                       | Praktijk                                | 2526              | 17                   |
| Muziek                | Kerndoel 10: Betekenis                  | 3815              | 19                   |
|                       | Kerndoelen 8 en 9: Muzikale aspecten    | 3815              | 28                   |
| Tekenen               | Beschouwing                             | 1179              | 24                   |
| Handenarbeid          | Beschouwing                             | 1039              | 24                   |
| Textiel               | Beschouwing                             | 522               | 20                   |



De proeftoetsen die niet uit afzonderlijke items bestaan maar uit een of meer opdrachten die op verschillende wijzen beoordeeld zijn, zoals de spreek- en schrijftaken, konden niet met OPLM geanalyseerd worden. Voor deze opgaven is het resultaat van de beoordeling van de taak gebruikt als schatting van de moeilijkheid.

#### *Samenstellen van de schalen*

De rapportage-eenheden voor de beoordeling van de opbrengsten zijn de domeinen van de kerndoelen. De items in de proefafnames komen daar niet zonder meer mee overeen. Meestal omvatten ze meerdere of alle domeinen van een vak (onderdeel). Het was dus nodig om de opgaven van de proefafnames opnieuw in te delen, zodanig dat zo veel mogelijk werd aangesloten bij de indeling in domeinen.

In een enkel geval (Muziek) waren de onderdelen van een domein apart getoetst. Soms ook spoorde de indeling van de proefafname niet helemaal met de indeling in domeinen. Dit was met name het geval bij de practicum- en praktijktoetsen. Als de proeftoetsen een afwijkende indeling kende, werd deze overgenomen.

Voor het samenstellen van verzamelingen opgaven bij een domein van de kerndoelen werd uitgegaan van de indeling die het Cito bij de constructie van de opgaven heeft gehanteerd. Voor wiskunde was geen indeling voorhanden. Medewerkers van het IOWO hebben alsnog de opgaven ingedeeld bij de kerndoelen. Dit leidde voor alle opgaven tot een eenduidige toewijzing aan de domeinen.

Bij de samenstelling van de verzameling opgaven werd gestreefd naar een gelijkmatig verdeling van opgaven over het bereik van de schaal. Daardoor werd bereikt dat de (lokale) betrouwbaarheid op alle punten van de schaal zo hoog mogelijk is (Kleintjes, 1999). Als tweede criterium gold de representativiteit van de opgaven voor het domein. Als twee of meer opgaven met een min of meer gelijke moeilijkheid beschikbaar waren, werd dit selectie criterium gehanteerd. Verder werd per schaal gestreefd naar 20 tot 25 opgaven. Dit aantal is voor de beoordelaars te overzien en levert een voldoende informatiewaarde voor de schaal op.

Indien te weinig opgaven voor een domein beschikbaar waren, werden in overleg met de toetsconstructeurs domeinen samengevoegd. Dit was bijvoorbeeld het geval bij Natuur- en scheikunde en bij Economie. Indien door samenvoegen van domeinen geen hoger aantal opgaven bereikt kon worden, was genoeg genomen met minder dan 20 opgaven indien alle domeinen voldoende gerepresenteerd waren.

### **3. OPZET VAN DE BEOORDELING**

#### *Keuze van de procedure*

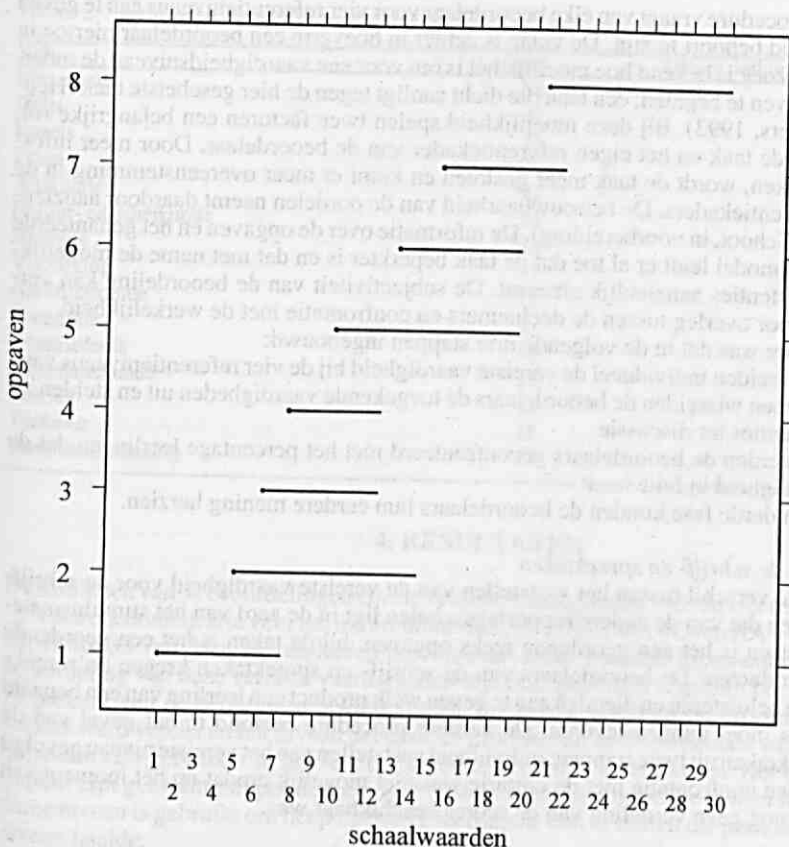
Voor het beoordelen van opbrengsten is een aantal procedures mogelijk. Een overzicht wordt gegeven door Van der Linden, Meijer en Vos (1997). Gekozen werd voor de procedure die werd ontwikkeld in het PPOON-onderzoek (Van der Schoot, 1996). Dit is een variant op de Angoff-procedure (Verstralen, 1993) waarbij in plaats van de parameters van de klassieke testtheorie gebruik wordt gemaakt van de moeilijkheid van de opgaven en de vaardigheid van de leerlingen zoals die geschat worden in het gebruikte itemresponsemodel. Daardoor is het mogelijk om de beoordelaars meer en betere informatie te verschaffen voor het uitvoeren van hun taak.

Bij de gekozen procedure krijgen de beoordelaars de opgaven geordend voorgelegd, samen met psychometrische informatie over de opgaven, die grafisch worden weergegeven. Op de grafiek is voor elke opgave door een lijn weergegeven bij welke vaardigheid vijftig tot tachtig procent van de leerlingen de opgaven goed maakt (Figuur 1). Daarmee wordt de vaardig-

heidschaal in drie sectoren verdeeld: minder dan vijftig procent kans op goed (links van de lijn), tussen de vijftig en tachtig procent kans op goed en meer dan tachtig procent kans op goed (rechts van de lijn). Deze gebieden zijn benoemd met 'onvoldoende beheersing', 'redelijke tot goede beheersing' en 'goede beheersing'. Zo wordt de derde opgave door leerlingen met een vaardigheid van 6 of lager onvoldoende beheerst en beheersen leerlingen met een vaardigheid tussen 6 en 15 de opgave redelijk tot goed. Bij een vaardigheid van 16 of hoger wordt de opgave goed beheerst. Omgekeerd kan, door vast te stellen in welke mate de opgaven beheerst moeten worden, vastgesteld worden welke vaardigheid voor deze beheersing vereist is.

Het was de taak van de beoordelaars om, bij een aantal referentieniveaus te bepalen welke opgaven in welke mate beheerst moesten worden. Daarmee stelden ze de vereiste vaardigheid bij deze niveaus vast.

Figuur 1: Grafisch weergegeven voorbeeld van acht opgaven geordend op een vaardigheidschaal die loopt van 1 tot en met 30.



De keuze voor de vier referentieniveaus vloeit voort uit de gekozen uitwerking van de beheersingsdefinitie: het vereiste niveau van leerlingen die volgens de kerndoelen onderwijs hebben gehad en door willen stromen naar de hogere leerjaren van het VO. Het gaat er in de basis-

vorming op de eerste plaats om dat de leerling een minimumniveau behaalt. Daarnaast is het echter ook nodig, dat de leerling zich kwalificeert voor de hogere leerjaren van een van de opleidingstypen in het VO. Dit houdt in dat in ieder geval rekening gehouden moest worden met de opleidingsmogelijkheden na de basisvorming. Het gemiddelde niveau voor de onderscheiden opleidingen sloot hierbij aan. Deze vier referentieniveaus werden voor de beoordelaars derhalve als volgt omschreven:

1. Het minimumniveau: het niveau dat door 90 procent van de leerlingen gehaald zou moeten worden.
2. Het gemiddelde (I)VBO-niveau: het niveau dat door vijftig procent van de leerlingen in een (I)VBO-opleiding gehaald zou moeten worden.
3. Het gemiddelde MAVO-niveau: het niveau dat door vijftig procent van de leerlingen in een MAVO-opleiding gehaald zou moeten worden.
4. Het gemiddelde HAVO/VWO-niveau: het niveau dat door vijftig procent van de leerlingen in een HAVO- of VWO-opleiding gehaald zou moeten worden.

Het minimumniveau is geoperationaliseerd op 90 procent van de populatie als geheel, hetgeen ongeveer overeenkomt met 75 procent van de (I)VBO-populatie<sup>4</sup>.

De beschreven procedure vraagt van elke beoordelaar voor vier referentieniveaus aan te geven wat de vaardigheid behoort te zijn. De vraag is echter in hoeverre een beoordelaar hiertoe in staat is. Uit onderzoek is bekend hoe moeilijk het is om voor een vaardigheidsniveau de moeilijkheid van opgaven te bepalen; een taak die dicht aanligt tegen de hier geschetste taak (Heuvelmans & Sanders, 1993). Bij deze moeilijkheid spelen twee factoren een belangrijke rol: de openheid van de taak en het eigen referentiekader van de beoordelaar. Door meer informatie te verstrekken, wordt de taak meer gesloten en komt er meer overeenstemming in de gehanteerde referentiekaders. De betrouwbaarheid van de oordelen neemt daardoor aanzienlijk toe (Van der Schoot, in voorbereiding). De informatie over de opgaven en het gehanteerde psychometrische model leidt er al toe dat de taak beperkter is en dat met name de mogelijkheid van inconsistenties aanzienlijk afneemt. De subjectiviteit van de beoordeling kan verbeterd worden door overleg tussen de deelnemers en confrontatie met de werkelijkheid.

In de procedure was dat in de volgende drie stappen ingebouwd:

1. Beoordelaars stelden individueel de vereiste vaardigheid bij de vier referentieniveaus vast.
2. In kleine groepen wisselden de beoordelaars de toegekende vaardigheden uit en stelden zij grote discrepanties ter discussie.
3. Vervolgens werden de beoordelaars geconfronteerd met het percentage leerlingen dat de vereiste vaardigheid in feite haalt.

Na de tweede en derde fase konden de beoordelaars hun eerdere mening herzien.

#### *Procedure voor de schrijf- en spreektaken*

Het voornaamste verschil tussen het vaststellen van de vereiste vaardigheid voor de schrijf- en spreektaken en dat van de andere rapportageschalen ligt in de aard van het stimulusmateriaal. Bij de toetsen is het een geordende reeks opgaven, bij de taken is het een geordende reeks leerlingproducten. De beoordelaars van de schrijf- en spreektaken kregen zo'n reeks voorgelegd of te beluisteren en dienden aan te geven welk product een leerling van een bepaald niveau minstens moet kunnen leveren. De verdere procedure bestond in het geval van de schrijf- en spreektaken uit twee stappen: individueel vaststellen van het vereiste niveau gevolgd door overleg. Een confrontatie met de empirie was niet mogelijk omdat op het moment van de beoordeling nog geen verdeling van de scores beschikbaar was.

#### *Selectie van beoordelaars*

De beoordeling van het vereiste niveau behoort plaats te vinden door een representatieve steekproef van ter zake kundigen. Er zijn drie groepen ter zake kundigen onderscheiden:

- leraren die lesgeven aan leerlingen in de basisvorming;

- opleiders van leraren;
- overige vakdidactici die betrokken zijn bij de ontwikkelingen in de basisvorming, zoals medewerkers van de SLO en de LPC.

Omdat de beoordeling per vak plaatsvond, waren de groepen gestratificeerd naar vak.

Voor de leraren werd een aselecte steekproef van scholen benaderd. Via de directie waren de secties benaderd met het verzoek of één van de leden van de sectie bereid was om aan de beoordeling deel te nemen.

Voor de opleiders en de overige vakdidactici werden degenen benaderd die al eerder bij de evaluatie van de basisvorming betrokken waren geweest. Gestreefd werd naar een verdeling van vijftien leraren en tien andere deskundigen per vak, met uitzondering van de Moderne vreemde talen waarvoor een groter aantal werd geworven, omdat meer taken te verrichten waren. Tabel 2 bevat de aantallen leraren en andere deskundigen die aan de beoordeling hebben meegewerkt. Uit de tabel blijkt dat bij de meeste vakken voldoende leraren bereid gevonden waren om deel te nemen. Voor de deskundigen was dat duidelijk minder het geval.

Tabel 2: Aantallen beoordelaars per vak.

| Vak                   | Leraren | Andere deskundigen |
|-----------------------|---------|--------------------|
| Nederlands            | 16      | 6                  |
| Duits                 | 25      | 7                  |
| Engels                | 24      | 5                  |
| Frans                 | 23      | 6                  |
| Wiskunde              | 22      | 11                 |
| Natuur- en scheikunde | 16      | 10                 |
| Biologie              | 16      | 7                  |
| Verzorging            | 17      | 7                  |
| Aardrijkskunde        | 15      | 10                 |
| Economie              | 13      | 8                  |
| Geschiedenis          | 18      | 11                 |
| Informatiekunde       | 12      | 3                  |
| Muziek                | 14      | 4                  |
| Techniek              | 18      | 6                  |
| Beeldende vorming     | 12      | 6                  |

#### 4. RESULTATEN

De resultaten van de beoordelingen zijn de vereiste vaardigheden voor de vier referentieniveaus.

Voor elk domein zijn voor het minimumniveau, (I)VBO-, MAVO- en HAVO/VWO-niveau elk 18 tot 33 van deze vaardigheden beschikbaar. Deze waarden lopen uiteraard uiteen. Uit de verdeling van deze vereiste vaardigheden zijn drie ijkpunten afgeleid: het niveau dat door 75 procent van de beoordelaars werd vereist, het niveau dat door 50 procent van de beoordelaars werd vereist en het niveau dat door 25 procent van de beoordelaars werd vereist. Deze ijkpunten zijn gebruikt om het feitelijk bereikte niveau te kwalificeren. Het eerste en laatste ijkpunt zijn gebruikt om per domein vast te stellen of het voldoende beheerst werd. Het mediane niveau is gebruikt om het percentage leerlingen vast te stellen dat per vak het minimumniveau haalde.

##### *Beheersing van de domeinen*

Voor de beheersing van de domeinen (rapportage-eenheden) zijn drie kwalificaties gebruikt. Als het feitelijk niveau lager is dan het niveau dat driekwart van de deskundigen minstens

eist, dan presteren de leerlingen beneden het gewenste niveau. De leerlingen beheersen de kerndoelen die bij dat domein horen dan onvoldoende. We spreken dan van 'onder niveau'. Als het feitelijke niveau hoger ligt dan het niveau dat driekwart van de deskundigen hoogstens eist, dan presteren de leerlingen boven het geëiste niveau. De leerlingen beheersen de kerndoelen die bij dat domein behoren in ruime mate. We spreken dan van 'boven niveau'. In alle andere gevallen spreken we van 'op niveau'.

Het belangrijkste punt is uiteraard het onderscheid tussen enerzijds 'onder' en 'op niveau' en anderzijds 'boven niveau', omdat daarvan ons oordeel over de opbrengsten afhangt.

Tabel 3: Percentage van de domeinen dat 'onder', 'op' of 'boven' niveau wordt beheerst.

| Kwalificatie | VBO | MAVO | HAVO/VWO |
|--------------|-----|------|----------|
| Onder niveau | 25  | 24   | 35       |
| Op niveau    | 31  | 57   | 51       |
| Boven niveau | 44  | 19   | 14       |

Tabel 3 geeft voor de afzonderlijk opleidingen het percentage van de domeinen dat 'onder', 'op' of 'boven' niveau wordt beheerst.

Een gemiddelde HAVO/VWO-leerling beheerst twee van de drie domeinen 'op' of 'boven' niveau. Voor een gemiddelde VBO- of MAVO-leerling is de situatie wat gunstiger: driekwart van de domeinen worden 'op' of 'boven' niveau beheerst. Verder valt op dat een gemiddelde VBO-leerling een substantieel aantal domeinen 'boven' niveau beheerst, terwijl dit aantal voor een gemiddelde MAVO- of HAVO/VWO-leerling veel lager ligt. De gemiddelde MAVO- of HAVO/VWO-leerling beheerst de domeinen vooral 'op' niveau. Daarbij moet er uiteraard rekening mee worden gehouden dat het gaat om verschillende vereisten die afzonderlijk voor de onderscheiden opleidingen waren vastgesteld.

#### Beheersing per cluster van vakken

Tabel 4 bevat de resultaten per cluster van vakken. Het aantal domeinen per cluster verschilt nogal. Bij het cluster Talen zijn de meeste domeinen beoordeeld, dat komt omdat de verschillende onderdelen van de schrijftaken uitgebreid aan de orde kwamen. Bij het cluster Beeldend was bij Tekenen, Handvaardigheid en Textiele werkvormen de beoordeling beperkt tot de schriftelijke beschouwing.

Tabel 4: Aantal domeinen per cluster en percentage domeinen dat 'op' of 'boven' niveau wordt beheerst.

| Vakkencluster        | Aantal domeinen | Percentage |     |      |          |
|----------------------|-----------------|------------|-----|------|----------|
|                      |                 | Minimum    | VBO | MAVO | HAVO/VWO |
| Talen                | 24              | 79         | 95  | 96   | 79       |
| Exact                | 15              | 60         | 67  | 67   | 53       |
| Mens en maatschappij | 15              | 60         | 67  | 47   | 40       |
| Beeldend             | 5               | 40         | 40  | 80   | 100      |

Het percentage domeinen bij de Talen dat 'op' of 'boven' niveau wordt beheerst, is relatief hoog; bij VBO en MAVO bijna honderd procent. Voor HAVO/VWO en voor het minimum-niveau is dat met bijna tachtig procent fors lager, maar het is hoger dan voor de vakkenclusters Exact en Mens en maatschappij in alle opleidingen. In de tabel valt verder op dat bij de Mens-



en maatschappijvakken de beheersing van de domeinen bij MAVO en HAVO/VWO laag is. Voor de leerlingen van het VBO en op het minimumniveau geldt dat voor de Beeldende vakken.

#### *Percentage leerlingen dat het minimumniveau haalt*

Om voor de vakken tot een overall oordeel te komen is het wenselijk dat de opbrengsten voor het vak in één getal worden uitgedrukt. Gekozen is voor het percentage leerlingen dat het minimumniveau haalt, gemiddeld over de domeinen van het desbetreffende vak. Het minimumniveau is gedefinieerd als het niveau dat door 50 procent van de beoordelaars is aangegeven. De voornaamste reden voor deze keuze ligt in het feit dat dit overeenkomt met de 'minimum'-leerstandaard die de Onderwijsraad voorstelt (Onderwijsraad, 1999).

De verdeling van de vaardigheden voor de domeinen is niet bekend. Wel is de verdeling apart voor (I)VBO, MAVO en HAVO/VWO gegeven. Het percentage leerlingen dat het minimumniveau beheerst is daarom apart berekend voor elke deelpopulatie. Vervolgens is het gewogen gemiddelde berekend, met als gewichten de relatieve omvang van de deelpopulaties. Tabel 5 bevat de uitkomsten.

*Tabel 5: Percentage leerlingen met prestaties boven het minimumniveau*

| Vak                   | Percentage |
|-----------------------|------------|
| Engels                | 99         |
| Nederlands            | 98         |
| Duits                 | 98         |
| Biologie              | 96         |
| Muziek                | 94         |
| Natuur- en scheikunde | 90         |
| Economie              | 89         |
| Frans                 | 89         |
| Geschiedenis          | 88         |
| Wiskunde              | 87         |
| Aardrijkskunde        | 84         |
| Techniek              | 82         |
| Informatiekunde       | 81         |
| Beeldende vorming     | 72         |
| Verzorging            | 63         |

Het minimumniveau is gedefinieerd als het niveau dat door 90 procent van de leerlingen gehaald moet worden. Dat komt ongeveer overeen met het niveau dat 75 procent van de leerlingen in het (I)VBO moet halen.

Bij zes vakken is het percentage leerlingen dat boven het minimumniveau presteert hoger dan 90 procent. Bij vier vakken ligt het percentage dicht tegen de 90 procent aan. Vijf vakken wijken meer dan vijf procent af van de gestelde grens. Daarbij zijn de vier vakken die met de basisvorming zijn ingevoerd: Techniek, Informatiekunde, Beeldende vorming en Verzorging. Bij Verzorging kan een rol spelen dat de toetsen zijn gebaseerd op de nieuwe kerndoelen, terwijl het onderwijs nog op de oude kerndoelen is gebaseerd.

In Tabel 5 zijn de resultaten per vak gepresenteerd. Uit de tabel mag niet worden afgeleid dat de prestaties van een groot deel van de leerlingen op alle vakken boven het minimumniveau zouden liggen. Het is aannemelijker dat slechts weinig leerlingen alle kerndoelen voor de

onderscheiden vakken op het minimumniveau beheersen. De hier gepresenteerde uitkomsten zijn ontleend aan een reeks univariate verdelingen en niet aan één multivariate verdeling (van den Bergh & Kuhlemeier, 1990).

## 5. CONCLUSIE

De gestelde vraag 'Beheersen voldoende leerlingen aan het eind van de basisvorming de kern-doelen' betreft enerzijds de vraag of voldoende leerlingen het minimumniveau halen en anderzijds de vraag of leerlingen het onderwijs in één van de opleidingstypen van het VO kunnen vervolgen.

### *Wat het minimumniveau betreft*

Bij de Talen voldoen de prestaties aan de door de beoordelaars gestelde eisen: 90 procent van de leerlingen of meer presteert boven het vereiste minimumniveau. Voor de Mens- en maatschappijvakken en de Exacte vakken voldoet net iets minder dan het vereiste aantal leerlingen aan de gestelde eisen. Voor de 'nieuwe' vakken halen niet voldoende leerlingen het minimumniveau. Dit houdt echter niet in dat individuele leerlingen voor alle vakken de kerndoelen op minimumniveau beheersen.

### *Wat de opleidingstypen betreft*

In het VBO en MAVO presteren de leerlingen op drie van de vier domeinen 'op' of 'boven' het niveau dat door de beoordelaars is vastgesteld. In het HAVO/VWO is dat voor twee van de drie domeinen het geval. Voor dat opleidingstype zijn met name de Exacte vakken en de Mens- en maatschappijvakken problematisch.

## NOTEN

1. Helaas was het niet mogelijk om gebruik te maken van de resultaten van de afsluitende toetsen basisvorming. Door de keuzemogelijkheden die de scholen bij de afname hebben, lieten de leerlingresultaten een vergelijking tussen de scholen niet toe.
2. In het kader van het VOCL-onderzoek zijn specifieke toetsen voor het meten van de algemene vaardigheden geconstrueerd. Over de uitkomsten hiervan is afzonderlijk gerapporteerd (van den Bergh e.a., 2000).
3. In deze procedure worden per school maar één of enkele toetsen afgenomen. De overlap met de 120 scholen die betrokken zijn in de evaluatie is daardoor te klein om een relatie te leggen tussen de proeftoetsen en de gegevens van de in de evaluatie betrokken scholen.
4. De (I)VBO populatie bestond blijkens de integrale leerlingtelling in het schooljaar 1998/1999 uit 34.1 procent van alle leerlingen. Strikt genomen is 25 procent van deze (deel)populatie 8.5 procent van de gehele populatie. Om voor de hand liggende redenen is echter steeds gesproken over 25 procent van de (I)VBO-populatie, overeenkomend met 10 procent van de totale populatie.

## LITERATUUR

- Bergh, H. van den, Bosker, R., Brandsma, H., Meijer J. & Elshout-Mohr, M. (2000). *Algemene vaardigheden. Veranderingen in de tijd en relaties met het onderwijsaanbod*. Artikel aangeboden voor publicatie.
- Bergh, H. & Kuhlemeier H. (1990). De haalbaarheid van de eindtermen Nederlands in de basisvorming. *Pedagogische Studiën*, 67, 1-15.
- Cito (1993, 1996) *Toetsmodellen Afsluitingstoetsen Basisvorming*. Arnhem: Cito.
- Commissie toetsing basisvorming (1996). *De toets der kritiek*. Zoetermeer: Ministerie van OCenW.

- Heuvelmans, T. & Sanders, P.F. (1993). Beoordelaarsovereenstemming. Hoofdstuk 12 in: Eggen, T.J.H.M. en Sanders, P.F. (red.). *Psychometrie in de praktijk*. Arnhem: Cito.
- Inspectie van het onderwijs (1997a). *Plan voor de evaluatie van de basisvorming*. Utrecht: Inspectie van het onderwijs.
- Inspectie van het onderwijs (1997b). *Gebruik van de toetsen basisvorming in het schooljaar 1996/'97*. Utrecht: Inspectie van het onderwijs.
- Inspectie van het onderwijs (1999). *Werk aan de basis. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het onderwijs.
- Kleintjes, F.G.M. (1998). *Bepaling van de meetnauwkeurigheid in evaluatieonderzoek. Toepassing van een IRT-model bij de evaluatie van de basisvorming*. Paper gepresenteerd op de Onderwijs Research Dagen 1998 te Enschede.
- Kuhlemeier, J.B., Kremers, E.J.J. & Kleintjes F.G.M. (1997). *De eerste generatie afsluitingstoetsen basisvorming: gebruik, betrouwbaarheid en maakbaarheid. Onderzoeksrapporten algemeen voortgezet onderwijs*. Arnhem: Cito.
- Linden, W.J. van der, Meijer, R.R. & Vos, H.J. (1997). *Normeringsmethoden voor inspectie-evaluaties*. Rapport 97-01. Enschede: Universiteit Twente, Vakgroep OMD.
- Onderwijsraad (1999). *Zeker Weten. Leerstandaarden als basis voor toegankelijkheid*. Den Haag: Onderwijsraad.
- Schoot F. van der (1996). *Setting standards for national attainment targets on PPON ability scales for Dutch language*. Paper gepresenteerd op de European Conference on Educational Research te Sevilla (Spanje).
- Schoot, F. van der. *Standaarden voor kerndoelen basisonderwijs. De ontwikkeling van standaarden basisonderwijs op basis van de resultaten van peilingsonderzoek*. Proefschrift in voorbereiding.
- Uiterwijk, J.B. (1998). *Interpretatie van het leerlingrapport. Toelatings- en doorstroomgegevens van leerlingen die in 1997 aan de Eindtoets Basisonderwijs deelnamen*. Arnhem: Cito.
- Verhelst, N.D., Glas, C.A.W. & Verstralen, H.H.F.M. (1995). *One Parameter Logistic Model OPLM, manual and computer program*. Arnhem: Cito.
- Verstralen, H.H.F.M. (1993). Schalen, normen en cijfers. Hoofdstuk 13 in: Eggen, T.J.H.M. en Sanders, P.F. (red.). *Psychometrie in de praktijk*. Arnhem: Cito.
- Werf, M.P.C. van der, M.J. Lubbers en H. Kuyper (1999). *Onderwijsresultaten van VOCL '89 en VOCL '93 leerlingen*. GION: Groningen.

# Kwaliteit van de opbrengsten: Zijn de prestaties tekstbegrip Nederlands en Wiskunde vooruitgegaan?

Huub van den Bergh, *Utrechts Instituut voor Linguïstiek*  
Roel Bosker, *Universiteit Twente*  
Hennie Brandsma, *Universiteit Twente*

## ABSTRACT

*In this study the achievements of students of about fifteen years of age on a test for reading achievement and a test for mathematics of three cohorts were analyzed, in order to assess differences in achievement due to the implementation of a common core curriculum in secondary education ('basisvorming'). This is a large-scale reform of the first three years of secondary education, not only with respect to the content of curricula (for all subject matter domains), but also with respect to pedagogical and didactical issues.*

*The first cohort in this study consists of students who enrolled in secondary education in '89, well before the common core curriculum as a reform of secondary education was implemented. The second cohort is constituted by students who started secondary education in the year in which the innovation actually was implemented (1993). The last cohort consists of students who started secondary education when the innovation was five years on its way, and their reading and mathematics achievements were measured when they reached the third class of 'basisvorming'.*

*It is shown that there are only very small differences in mean scores for both reading or mathematics between the cohorts. Differences between subpopulations (boys versus girls, ethnic minority versus privileged students, etc.) appeared to be rather consistent over cohorts. Despite of both sampling problems or measurement problems, it is concluded that the high expectations of the effects of the common core curriculum were not met (yet).*

## 1. INLEIDING

Met de invoering van de basisvorming beoogde de overheid een algehele verhoging van het onderwijsniveau te bewerkstelligen, de verschillen tussen leerlingen te verkleinen, de inhoud van het onderwijs beter controleerbaar te maken voor het parlement etc. (Ministerie van Onderwijs en Wetenschappen, 1989). Het is dan ook niet overdreven om te stellen dat bij de invoering van de basisvorming de verwachtingen omtrent de effecten hiervan hooggespannen waren. Het logische gevolg is dat bij de evaluatie van de basisvorming aandacht besteed zou moeten worden aan de prestaties van de leerlingen. Eén van de vragen van het Platform Evaluatie Basisvorming (1985) is dan ook: worden de prestatie-effecten die met de basisvorming zijn beoogd op langer(e) termijn gerealiseerd?

Het prestatieniveau heeft (o.a.) gestalte gekregen in een vergelijking van de prestaties van leerlingen in de periode 1987-1998 voor de vakken Nederlands en Wiskunde. Hierbij gaat het

met name om veranderingen in het gemiddelde prestatieniveau bij beide vakken, maar ook om veranderingen in de verschillen tussen leerlingen. Hierbij werd verwacht dat het gemiddelde prestatieniveau zou stijgen, en dat de verschillen tussen leerlingen zouden afnemen (Ministerie van Onderwijs en Wetenschappen, 1989)<sup>1</sup>.

Natuurlijk is de basisvorming niet de enige verandering in het onderwijs in de onderhavige periode. Te denken valt o.a. aan de schaalvergroting die in het voortgezet onderwijs in ruwweg dezelfde periode plaatsgevonden heeft, en (eventuele) effecten van campagnes als 'kies exact', en 'een slimme meid is op de toekomst voorbereid', etc.

We kunnen eventuele veranderingen in de prestaties dan ook niet zonder meer toeschrijven aan de introductie van de basisvorming. Door bij voorbeeld het ontstaan van brede scholengemeenschappen is het nu veel makkelijker om van opleidingsniveau te veranderen dan het vroeger was. Vandaar dat prestatieveranderingen ook bestudeerd moeten worden in relatie tot diverse leerlingkenmerken.

Bij de invoering van de basisvorming in 1992 kon natuurlijk niet met terugwerkende kracht een onderzoek naar de vaardigheden van de leerlingen uitgevoerd worden. Vandaar dat besloten is om voor de evaluatie van eventuele prestatieveranderingen aan te sluiten bij het onderzoek van NWO uitgevoerde VOCL-onderzoek (Voortgezet Onderwijs Cohort Leerlingen), waarin leerlingen gedurende verschillende jaren gevolgd worden. Door de prestaties van leerlingen in het derde leerjaar uit verschillende cohorten te vergelijken kan inzicht verkregen worden in de mogelijke gevolgen van de invoering van de basisvorming.

## 2. METHODE

### *Steekproeven*

Voor de vergelijking van de leerlingprestaties is gebruik gemaakt van de gegevens uit het VOCL-onderzoek (Van der Werf, Kuyper & Lubbers, 1999). Dit is een cohortstudie die tot nu toe tweemaal is uitgevoerd, waarbij een grote groep scholen (en leerlingen) gedurende een langere periode gevolgd zijn. De eerste cohortstudie is gestart in het schooljaar '88/'89, voordat de basisvorming ingevoerd is, en de tweede cohortstudie is gestart in het schooljaar '93/'94 ten tijde van de wettelijke aanvang van de basisvorming. Als de basisvorming het verwachte positieve effect zou hebben op de prestaties van de leerlingen, dan zouden de leerlingen die in schooljaar '93/'94 begonnen zijn met het voortgezet onderwijs aan het einde van de basisvorming hogere prestaties behaald moeten hebben dan de leerlingen die voor de invoering van de basisvorming gestart zijn.

Nu is de invoering van een zo'n grootschalige vernieuwing als de basisvorming natuurlijk geen eenvoudige zaak; docenten moeten zich heroriënteren, hun didactiek en lesmateriaal aanpassen etc. Er kan dan ook niet verwacht worden dat de eerste lichting basisvormingsleerlingen (i.e. leerlingen die in 1993 met het VO begonnen zijn) veel beter presteren dan eerdere leerlingen. Pas in latere jaren, wanneer het onderwijs de tijd heeft gehad om zich aan de veranderde doelstellingen, didactiek en werkvormen aan te passen mogen positieve effecten verwacht worden. Om deze reden zijn de gegevens uit beide cohortstudies aangevuld met een steekproef van leerlingen die in het schooljaar '95/'96 met het voortgezet onderwijs begonnen zijn, en in '98 in de derde klas zijn aangeland (Inspectie van het Onderwijs, 1999). Hoewel de vaardigheden van de leerlingen in deze aanvullende steekproef alleen aan het einde van de basisvorming gemeten zijn (i.e. derde klas), en er dus strikt genomen géén sprake is van een cohortstudie, zullen we de gegevens uit deze steekproef om redenen van uniformiteit toch aanduiden met de term cohort.



In de onderstaande tabel wordt een overzicht gegeven van de gegevens in de drie cohorten. Hierbij zijn niet alleen de totaal aantallen weergegeven, maar is tevens een opsplitsing gemaakt naar verschillende deelpopulaties.

*Tabel 1: Aantallen leerlingen per cohort per opleidingstype voor Nederlands en Wiskunde, alsmede een onderscheid naar verschillende deelpopulaties, alsmede het gemiddelde opleidingsniveau en het 'advies aan het einde van de basisschool' per combinatie van cohort en opleidingstype (standaarddeviaties tussen haakjes)*

|                        | Nederlands    |                |                |                |                |                |                |               |               | Wiskunde       |                |                |                |                |                |                |               |               |
|------------------------|---------------|----------------|----------------|----------------|----------------|----------------|----------------|---------------|---------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|---------------|---------------|
|                        | (I)VBO        |                |                | MAVO           |                |                | HAVO/VWO       |               |               | (I)VBO         |                |                | MAVO           |                |                | HAVO/VWO       |               |               |
|                        | 89-3          | 93-3           | 95-3           | 89-3           | 93-3           | 95-3           | 89-3           | 93-3          | 95-3          | 89-3           | 93-3           | 95-3           | 89-3           | 93-3           | 95-3           | 89-3           | 93-3          | 95-3          |
| N-scholen              | 194           | 175            | 31             | 205            | 173            | 53             | 72             | 74            | 35            | 195            | 162            | 30             | 203            | 164            | 54             | 71             | 69            | 35            |
| N-leerlingen           | 3419          | 2950           | 1074           | 4728           | 4226           | 1406           | 2045           | 1727          | 659           | 3738           | 4724           | 2033           | 2796           | 3765           | 1539           | 1043           | 1392          | 660           |
| % Jongens              | 59.05         | 57.93          | 60.61          | 44.86          | 45.84          | 48.65          | 48.26          | 45.98         | 44.46         | 60.43          | 45.47          | 48.35          | 58.63          | 46.43          | 46.65          | 61.69          | 47.49         | 42.88         |
| % Cult Acht            | 7.55          | 6.88           | 4.19           | 5.52           | 3.12           | 5.41           | 5.43           | 2.03          | 4.40          | 7.78           | 5.50           | 5.31           | 7.40           | 2.95           | 2.27           | 3.93           | 5.17          | 3.33          |
| % Afstroom             | 12.78         | 19.97          | 18.06          | 8.52           | 11.90          | 16.22          | 0.49           | 5.91          | 6.22          | 13.35          | 11.60          | 8.80           | 14.66          | 12.01          | 13.71          | 0.96           | 6.54          | 6.21          |
| % Opstroom             | 2.37          | 3.19           | 2.51           | 5.18           | 7.03           | 2.92           | 26.31          | 26.23         | 19.97         | 2.78           | 1.97           | 1.28           | 8.76           | 7.17           | 4.94           | 4.89           | 29.09         | 18.79         |
| % 1-ouder              | 6.87          | 11.90          | 11.55          | 5.12           | 7.74           | 11.81          | 4.99           | 8.40          | 7.89          | 6.89           | 7.22           | 6.05           | 8.91           | 7.46           | 9.62           | 9.78           | 10.06         | 6.67          |
| Opleiding range: 1 - 4 | 2.13<br>(.81) | 2.39<br>(.78)  | 2.31<br>(.53)  | 2.71<br>(.88)  | 2.93<br>(.78)  | 2.73<br>(.71)  | 3.23<br>(.80)  | 3.36<br>(.74) | 3.17<br>(.78) | 2.13<br>(.81)  | 2.40<br>(.79)  | 2.33<br>(.52)  | 2.68<br>(.88)  | 2.89<br>(.78)  | 2.70<br>(.70)  | 3.24<br>(.80)  | 3.36<br>(.74) | 3.18<br>(.74) |
| Advies range: 1 - 9    | 2.92<br>(.98) | 3.11<br>(1.18) | 3.00<br>(1.17) | 5.52<br>(1.16) | 5.58<br>(1.25) | 5.95<br>(1.29) | 7.78<br>(1.04) | 7.81<br>(.98) | 8.04<br>(.88) | 2.81<br>(1.01) | 3.11<br>(1.20) | 2.93<br>(1.16) | 5.42<br>(1.17) | 5.48<br>(1.25) | 5.84<br>(1.28) | 7.76<br>(1.03) | 7.80<br>(.98) | 7.83<br>(.90) |

Uit Tabel 1 blijkt dat de steekproeven qua aantal sterk verschillen. Zo zijn voor het vak Nederlands in cohort '89 de prestaties van 3419 leerlingen van 194 scholen met een (I)VBO-opleiding gemeten, en cohort '93 zijn de prestaties van 2950 leerlingen van 175 scholen met een (I)VBO-opleiding geëvalueerd. Daarmee zijn de steekproeven in deze beide cohorten in termen van deelnemende opleidingen ruim vijf keer zo groot als de steekproef in cohort '95. Ook voor het MAVO is het aantal opleidingen in cohort '95 zo'n vier keer kleiner dan in de beide andere cohorten. Alleen voor HAVO/VWO-opleidingen is het aantal opleidingen in het laatste cohort 'slechts' tweemaal zo klein.

Hoewel het aantal opleidingen voor (I)VBO en MAVO in cohort '95 ongeveer vijf keer zo klein is als in de beide andere cohorten, is het aantal leerlingen twee tot drie keer zo klein. Dit is wellicht een gevolg van de schaalvergroting in het onderwijs. Hetgeen alleen maar aangeeft dat er gedurende de onderzochte periode meer veranderingen in het onderwijs plaatsgevonden hebben dan alleen de invoering van de basisvorming. Dit kan de interpretatie van de resultaten bemoeilijken.

Naast verschillen in de grootte van de steekproeven zijn er ook verschillen tussen de cohorten voor andere variabelen. Zo varieert voor bij voorbeeld Nederlands het percentage jongens van 57.9 tot 60.6 op het (I)VBO, van 44.9 tot 48.7 op het MAVO en van 44.5 tot 48.3 procent op het HAVO/VWO, etc. (zie Tabel 1). Ook met betrekking tot het hoogste door (één van) de ouders genoten opleidingsniveau en het advies dat aan de leerlingen aan het einde van de basisschool gegeven wordt kunnen verschillen tussen de drie cohorten aangegeven worden. Zo is het gemiddelde opleidingsniveau van de ouders van leerlingen in cohort '93 duidelijk hoger dan dat in de beide andere cohorten (zie Tabel 1). En, het advies aan het einde van de basisschool is voor leerlingen op MAVO- en HAVO/VWO-scholen duidelijk hoger in cohort '95, terwijl voor leerlingen op (I)VBO-scholen het gemiddelde advies in cohort '93 hoger is.

Hoewel de bovenstaande voorbeelden over de voor de verschillen tussen de onderscheiden cohorten toegelicht zijn voor 'Nederlands', gelden zij in gelijke mate voor 'Wiskunde'.

### Instrumenten

Om de prestaties van de drie cohorten te kunnen vergelijken is in elk cohort een tekstbegrip-toets Nederlands en een toets Wiskunde afgenomen. Deze toetsen zijn niet identiek in alle drie de cohorten. Omdat de derdeklastoetsen in het eerste cohort (i.e. '89-3) niet volledig dekkend bleken voor de (latere) kerndoelen - i.c. het aangeboden curriculum - zijn vanaf cohort '93-3 aangepaste versies gebruikt. Een directe vergelijking van de geobserveerde scores van de derdeklassers is derhalve niet zonder meer mogelijk. Maar, omdat er in beide versies van de toetsen voldoende zogenaamde anker items (i.e. identieke items) opgenomen zijn, kunnen de scores op beide toetsen met behulp van een itemrespons- model op één schaal uitgedrukt worden. Hierdoor wordt een vergelijking van de drie cohorten, of een vergelijking in de tijd, mogelijk. Voor de equivalering van de toetsscores is gebruik gemaakt van het één parameter logistisch model (zie: Verhelst, Glas & Verstralen, 1993). Tevens bleek het door gebruik van itemrespons-modellen mogelijk de scores op de wiskundetoetsen die in het VBO en die in het AVO afgenomen zijn op één schaal te plaatsen, waardoor ook een vergelijking tussen opleidingstypen kan plaatsvinden. In Tabel 2 zijn enkele psychometrische gegevens voor de toetsen voor beide vakken vermeld, voor de geobserveerde gegevens in het derde cohort '95-3.

Tabel 2: Enkele psychometrische gegevens van de toetsen voor beide vakken.

|                              | Nederlands   | Wiskunde |       |
|------------------------------|--------------|----------|-------|
|                              | ((I)VBO/AVO) | (I)VBO   | AVO   |
| Gemiddeld aantal items goed  | 23.76        | 56.93    | 42.67 |
| Sd                           | 5.67         | 17.32    | 15.07 |
| Range: in theorie            | 0-40         | 0-101    | 0-77  |
| Range: geobserveerd          | 0-37         | 9-97     | 2-76  |
| Betrouwbaarheid ( $\alpha$ ) | .76          | .77      | .84   |
| N per toets                  | 3381         | 2190     | 1104  |

### Analyse

In feite willen we, gegeven de onderzoeksvragen, een onderscheid maken tussen de gemiddelden per vak per cohort, alsmede een schatting krijgen van de verschillen tussen leerlingen. Om recht te doen aan de hiërarchische structuur van de populatie is bij de analyse gebruik gemaakt van verschillende multi niveau modellen (zie Appendix), waarbij voor elk cohort een onderscheid gemaakt wordt tussen vaste parameters, de variantie tussen leerlingen (binnen scholen), en de variantie tussen scholen.

## 3. RESULTATEN

In Tabel 3 zijn per opleidingstype de verschillen tussen cohorten voor Nederlands en Wiskunde weergegeven.

Tabel 3: Onconditionele verschillen per vak tussen cohorten voor drie opleidingstypen (zie Appendix Vergelijking 2; standaardfouten tussen haakjes).

| Nederlands                            |        |        |             |        |        |              |        |        | Wiskunde     |        |        |              |        |        |              |        |        |
|---------------------------------------|--------|--------|-------------|--------|--------|--------------|--------|--------|--------------|--------|--------|--------------|--------|--------|--------------|--------|--------|
| (I)VBO                                |        |        | MAVO        |        |        | HAVO/VWO     |        |        | (I)VBO       |        |        | MAVO         |        |        | HAVO/VWO     |        |        |
| C '89                                 | C '93  | C '95  | C '89       | C '93  | C '95  | C '89        | C '93  | C '95  | C '89        | C '93  | C '95  | C '89        | C '93  | C '95  | C '89        | C '93  | C '95  |
| Vaste parameters                      |        |        |             |        |        |              |        |        |              |        |        |              |        |        |              |        |        |
| 2.30                                  | 488.7  | 7.46   | -0.16       | 506.0  | 7.16   | 1.99         | 520.7  | 6.75   | -1.66        | 492.1  | 0.23   | -1.47        | 504.9  | 0.94   | 0.85         | 516.6  | -0.60  |
| (0.55)                                | (0.49) | (1.04) | (0.50)      | (0.49) | (0.83) | (0.70)       | (0.66) | (1.11) | (0.44)       | (0.42) | (0.82) | (0.31)       | (0.33) | (0.49) | (0.44)       | (0.42) | (0.67) |
| Random parameters                     |        |        |             |        |        |              |        |        |              |        |        |              |        |        |              |        |        |
| Tussen-leerlingen-variantie ( $S^2$ ) |        |        |             |        |        |              |        |        |              |        |        |              |        |        |              |        |        |
| 181.5                                 | 226.5  | 291.4  | 196.4       | 246.5  | 246.1  | 198.6        | 259.5  | 231.3  | 110.2        | 110.2  | 119.5  | 66.36        | 68.49  | 69.75  | 81.91        | 83.60  | 75.78  |
| (4.48)                                | (6.00) | (12.7) | (4.11)      | (5.44) | (9.43) | (6.30)       | (8.97) | (13.0) | (2.60)       | (2.60) | (5.30) | (1.39)       | (1.61) | (2.69) | (2.61)       | (3.06) | (4.26) |
| Tussen-scholen-variantie ( $S^2$ )    |        |        |             |        |        |              |        |        |              |        |        |              |        |        |              |        |        |
| 25.28 (3.12)                          |        |        | 39.5 (3.85) |        |        | 34.63 (5.13) |        |        | 26.34 (2.81) |        |        | 22.80 (2.04) |        |        | 15.65 (2.21) |        |        |

Uit Tabel 3 blijkt dat de gemiddelde score voor Nederlands voor cohort '93-3 voor (I)VBO-leerlingen 488.7 is. Voor cohort '89-3 is de gemiddelde tekstbegripscore (488.7 + 2.30 =) 491.0. Dit verschil is significant (aangezien de parameter schatting (2.30) groter is dan 1.96 \* se). En, de gemiddelde score van cohort '95-3 is (488.7 + 7.46 =) 496.16. Er is voor (I)VBO-leerlingen dus een (significante) daling in de gemiddelde score tussen cohort '89 en cohort '93, en een (significante) stijging van de gemiddelde score tussen cohort '93 en cohort '95 (zie ook Figuur 1). Voor MAVO-leerlingen kan er geen verschil aangetoond worden in de gemiddelde tekstbegripscore tussen de cohorten '89-3 en '93-3, maar het gemiddelde in cohort '95-3 is wel hoger dan dat in beide andere cohorten. Voor HAVO/VWO-leerlingen is er geen verschil in gemiddelde score voor tekstbegrip tussen de drie cohorten.

Zoals verwacht zou mogen worden zijn de gemiddelde tekstbegripscores van (I)VBO-leerlingen in alle drie de cohorten lager dan de gemiddelde scores van MAVO-leerlingen in de desbetreffende cohorten, die op zich weer lager zijn dan de gemiddelde tekstbegripscores van HAVO/VWO-leerlingen ( $\chi^2_1 > 196.8$ ;  $df = 1$ ;  $p < .001$ ).

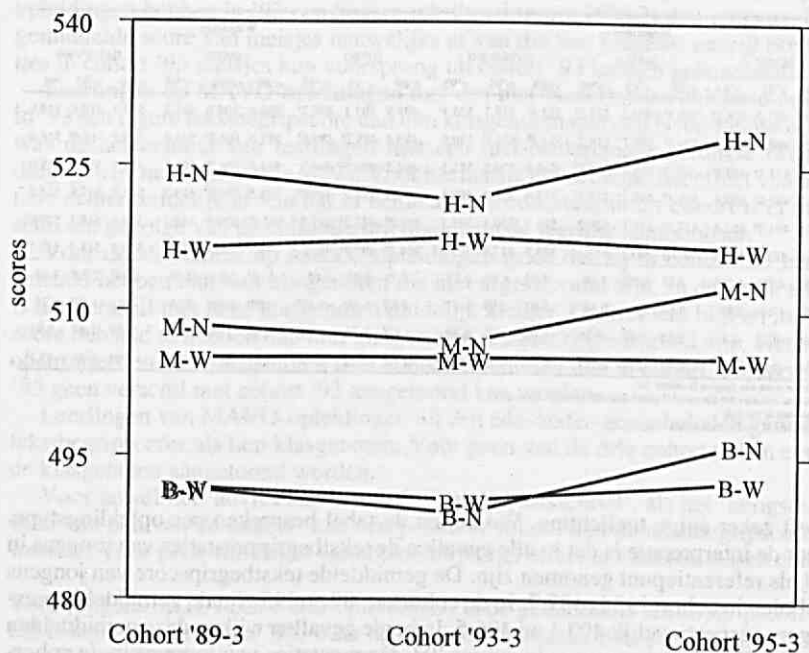
De random parameters moeten anders geïnterpreteerd worden dan de vaste parameters. Zijn de vaste parameters geschat als afwijking van het gemiddelde voor de desbetreffende (combinatie van) opleidingstypen, de random parameters zijn directe schattingen van de varianties in de desbetreffende cohorten (zie Appendix (2)). De verschillen tussen (I)VBO-leerlingen variëren van 181.5 voor cohort '89, tot 291.4 voor cohort '95. De verschillen tussen leerlingen (binnen scholen) nemen toe van cohort tot cohort ( $\chi^2_1 > 21.2$ ;  $p < .001$ ). Voor MAVO-leerlingen constateren we eveneens een duidelijke toename in de tussen-leerlingen-variantie in tekstbegripscores tussen cohort '89 en cohort '93 ( $\chi^2_1 = 23.3$ ;  $p < .001$ ), terwijl de verschillen tussen leerlingen niet toenemen van cohort '93 naar cohort '95 ( $\chi^2_1 = 0.80$ ;  $p = .37$ ). Voor HAVO/VWO-leerlingen constateren we net zoals bij MAVO-leerlingen een toename in de variantie tussen cohort '89 en cohort '93 ( $\chi^2_1 = 30.9$ ;  $p < .001$ ), terwijl het verschil tussen de varianties in cohort '93 en cohort '95 niet significant is.

Voor Wiskunde is de gemiddelde score voor (I)VBO-leerlingen in cohort '89-3 lager dan die in cohort '93-3, terwijl er geen verschil aangetoond is tussen de cohorten '93-3 en '95-3 (zie Figuur 1). Voor MAVO-leerlingen kan een zelfde patroon aangetoond worden als voor (I)VBO-leerlingen; een stijging in gemiddelde wiskunde prestaties tussen de cohorten '89-3 en '93-3, en geen verschil tussen '93-3 en '95-3. Voor HAVO/VWO-leerlingen verschilt de gemiddelde prestatie niet van cohort tot cohort (zie Figuur 1).

Vanzelfsprekend zijn de wiskunde-prestaties van (I)VBO-leerlingen gemiddeld genomen lager dan die van MAVO-leerlingen ( $\chi^2_1 = 139.87$ ;  $p < .001$ ), en MAVO-leerlingen hebben een lagere gemiddelde score dan HAVO/VWO-leerlingen ( $\chi^2_1 = 314.24$ ;  $p < .001$ ; zie Figuur 1).

De verschillen tussen (I)VBO-leerlingen (binnen scholen) zijn constant in de drie cohorten. Hetzelfde geldt voor MAVO- en HAVO/VWO-leerlingen. Wel blijken de verschillen tussen MAVO-leerlingen kleiner dan die tussen HAVO/VWO-leerlingen ( $\chi^2_1 = 29.19$ ;  $p < .001$ ), en die tussen HAVO/VWO-leerlingen kleiner dan die tussen (I)VBO-leerlingen ( $\chi^2_1 = 124.34$ ;  $p < .001$ ). Kortom, de populatie (I)VBO-leerlingen is zeer heterogeen.

Figuur 1: Gemiddelde prestaties per opleidingstype per cohort voort Nederlands (N) en Wiskunde (W), (B: VBO; M: MAVO; H: HAVO/VWO).



In Figuur 1 zijn de gemiddelde prestaties grafisch weergegeven. Duidelijk blijkt dat de gemiddelde prestaties voor Nederlands in cohort '93 iets lager zijn dan in cohort '89. De gemiddelde prestatie voor het derde cohort ('95) is iets hoger. De figuur geeft géén inzicht in de verschillen in prestaties, aangezien geen maat voor de spreiding is opgenomen. Als we ons echter bedenken dat de 90 procent waarschijnlijkheidsintervallen zo'n 50 punten bedragen voor de onderscheiden opleidingstypen, dan is wel duidelijk dat de verschillen in gemiddelden tussen de cohorten uiterst klein zijn.

Zoals eerder opgemerkt zijn de verschillen in gemiddelden tussen de drie cohorten wellicht toe te schrijven aan verschillen op diverse verklarende variabelen. Vandaar dat ook een model geanalyseerd is waarin de gemiddelde score geschat wordt conditioneel op de invloed van negen achtergrondvariabelen, te weten: sekse, culturele oriëntatie, opstomers, afstromers,

één-ouder-gezin, het hoogste opleidingsniveau in het gezin, het advies aan het einde van de basisschool, het gemiddelde advies, en de proportie meisjes op school. In Tabel 4 zijn de (geschatte) effecten van deze variabelen weergegeven. Conform het model worden per verklarende variabele (per combinatie van opleidingstypen) drie effecten geschat: de (gemiddelde) afwijking voor deze variabele van het algemeen gemiddelde, en de afwijkingen van deze variabelen in de cohorten '89 en '95 van 1993. Voor de overzichtelijkheid zijn voor de dummy-variabelen de (geschatte) gemiddelden voor verschillende subpopulaties per cohort weergegeven, en voor de continue variabelen de regressiegewichten per cohort.

Tabel 4: Verschillen tussen onderscheiden subpopulaties en regressiegewichten voor drie verklarende variabelen (Culturele, Oriënt.: culturele oriëntatie; Opl. Ouder: hoogste opleidingsniveau van ouders; Gem. Opl. Niv.: gemiddelde hoogste opleidingsniveau van ouders per school).

|                  | Nederlands |        |        |        |        |        |          |        |        | Wiskunde |        |        |        |        |       |          |        |        |
|------------------|------------|--------|--------|--------|--------|--------|----------|--------|--------|----------|--------|--------|--------|--------|-------|----------|--------|--------|
|                  | (I)VBO     |        |        | MAVO   |        |        | HAVO/VWO |        |        | (I)VBO   |        |        | MAVO   |        |       | HAVO/VWO |        |        |
|                  | c'89       | c'93   | c'95   | c'89   | c'93   | c'95   | c'89     | c'93   | c'95   | c'89     | c'93   | c'95   | c'89   | c'93   | c'95  | c'89     | c'93   | '95    |
| Jongens          | 492.1*     | 487.7  | 496.5* | 508.7* | 504.4  | 511.5* | 523.4*   | 518.3  | 525.7* | 493.9    | 493.1  | 496.2* | 506.6  | 506.8  | 507.3 | 518.4    | 518.6  | 516.9  |
| Meisjes          | 490.7*     | 490.8* | 496.5* | 508.3* | 508.7* | 514.3  | 525.5*   | 524.1* | 530.9  | 488.4    | 488.2* | 489.0* | 503.0  | 503.7* | 503.6 | 516.7    | 516.6* | 516.4  |
| Cult. Oriëntatie | 489.5      | 483.4* | 493.9  | 505.2  | 500.1* | 512.3* | 520.6    | 520.8  | 521.1  | 492.2    | 490.7* | 494.3  | 504.3  | 504.7* | 505.2 | 518.1    | 518.7  | 518.0  |
| Afstromers       | 487.8      | 483.0* | 489.6  | 505.4* | 498.7* | 510.8* | 525.4    | 522.9  | 527.8  | 490.3    | 488.0* | 489.0* | 505.5  | 504.3* | 503.2 | 516.0    | 517.8  | 514.2  |
| Opstromers       | 497.6      | 491.7* | 502.4  | 513.5* | 512.3* | 520.2  | 524.5    | 520.5  | 528.3  | 496.2*   | 497.2* | 497.8  | 507.4* | 509.4* | 510.3 | 518.4    | 518.3  | 520.6* |
| 1-ouder gezin    | 493.0      | 487.9  | 496.7  | 509.9  | 503.5  | 513.1  | 523.3    | 517.8  | 525.8  | 494.2*   | 491.3* | 495.2  | 506.2  | 506.4  | 505.7 | 519.8    | 518.4  | 513.1* |
| Advies bo*       | 4.42       | 4.21*  | 6.75*  | 4.05   | 4.08*  | 4.10   | 4.79     | 4.70   | 2.33   | 5.79     | 5.02*  | 5.08   | 4.18*  | 2.77*  | 2.88  | 4.14*    | 2.37*  | 3.69   |
| Opleid. Niv.*    | 0.84       | 1.16*  | 2.36   | 0.64   | 0.81*  | 0.10   | -0.01    | -0.09  | 1.97*  | 0.58     | 0.83*  | -0.47* | 0.49   | 0.38*  | 0.56  | 0.41     | 1.19*  | 0.75   |
| Gem. Opl.*       | 1.60       | -0.30  | -2.41  | 0.54   | 1.05   | -0.26  | 0.05     | 1.46   | 0.36   | 3.83*    | 0.65   | -0.30  | 2.49   | 1.05   | 0.61  | 0.29     | 1.81   | -0.92  |
| Prop. Meisjes*   | 1.64*      | -2.93  | 1.48   | 0.35   | 0.18   | -2.24  | 2.10     | 14.65* | -8.90* | -0.48*   | -4.02* | -4.73  | 1.15   | -2.39  | -7.76 | -7.99    | 1.26   | -13.1* |

a: significant t.o.v. de gemiddelde score van jongens in cohort '93;

b: significant ander verschil tussen cohorten dan het verschil tussen dezelfde cohorten bij jongens;

c: regressiegewicht wijkt significant af van nul;

d: regressiegewicht wijkt significant af van dat voor cohort '93.

e: continue variabele.

Tabel 4 behoeft zeker enige toelichting. We zullen de tabel bespreken per opleidingstype. Belangrijk voor de interpretatie is dat in alle gevallen de tekstbegriffscores van jongens in het cohort '93 als referentiepunt genomen zijn. De gemiddelde tekstbegriffscore van jongens op (I)VBO-scholen in cohort '93 is 487.7. In de cohorten '89 en '95 zijn de gemiddelde prestaties van jongens respectievelijk 492.1 en 496.5. In beide gevallen wijken deze gemiddelden af van de gemiddelde tekstbegriffscore in cohort '93. De prestaties van jongens in de cohorten '89 en '95 zijn dus hoger dan in cohort '93 (zie ook Figuur 2). De gemiddelde prestatie voor tekstbegrip van meisjes in cohort '93 is significant hoger dan die van jongens in hetzelfde cohort. Deze voorsprong van meisjes is echter noch in cohort '89, noch in cohort '95 aantoonbaar (zie ook Figuur 2). Voor de andere onderscheiden subpopulaties in dit opleidingstype kan weliswaar een verschil aangetoond worden met de gemiddelde tekstbegriffscores van jongens, maar deze verschillen zijn constant over de drie cohorten. Er kan derhalve niet aangetoond worden dat de verschillen tussen deze subpopulaties kleiner (dan wel groter) geworden zijn.

Voor de vier resterende variabelen in Tabel 4 moeten de effecten enigszins afwijkend geïnterpreteerd worden. Weergegeven is het regressiegewicht, die de mate van verandering in score per eenheid van de afhankelijke variabele aangeeft. Zo bedraagt de invloed van de variabele 'advies basisonderwijs' in cohort '93 in het (I)VBO 4.21; per eenheid in toename advies is de (verwachte) tekstbegriffscore 4.21 punten hoger. In cohort '89 is de relatie tussen advies



in tekstbegripscore hetzelfde als in cohort '93, maar in cohort '95 is het verband tussen 'advies basisonderwijs' en score van leerlingen sterker.

Het effect van het opleidingsniveau van de ouders is gerelateerd aan de leesprestaties van (I)VBO-leerlingen; hoe hoger het opleidingsniveau, hoe hoger de tekstbegripscores van de leerlingen. Deze relatie varieert niet over de drie cohorten. Tussen het gemiddelde opleidingsniveau van de ouders van de leerlingen op een school en de tekstbegripscores kan geen verband aangetoond worden, noch voor cohort '93, noch in één van de beide andere cohorten.

In cohort '93 is er geen relatie tussen de proportie meisjes op een school en de leesprestaties. Alleen voor cohort '89 kan aangetoond worden dat er een positief verband is tussen de proportie meisjes op een (I)VBO-school en de leesprestaties.

Ook voor jongens op MAVO-opleidingen blijkt de gemiddelde tekstbegripscore in de cohorten '89 en '95 hoger dan die in cohort '93. De gemiddelde prestaties van jongens in cohort '93 zijn dus lager dan die in de cohorten '89 of '95 (zie ook Figuur 2). Meisjes op MAVO-opleidingen hebben in '93 een hogere tekstbegripscore (508.7) dan jongens. In '89 wijkt de gemiddelde score van meisjes nauwelijks af van die van jongens, terwijl blijkens de prestaties in cohort '95 meisjes hun voorsprong uit cohort '93 hebben gehandhaafd.

Leerlingen op MAVO-opleidingen met een niet-Nederlandse culturele oriëntatie hebben in '93 een lagere tekstbegripscore dan hun klasgenoten met een Nederlandse oriëntatie. In '89 was de achterstand van leerlingen met een niet-Nederlandse culturele oriëntatie verhoudingsgewijs net zo groot als in '93. Voor het laatste cohort wijkt het effect van culturele oriëntatie echter duidelijk af van dat in beide andere cohorten; in dit cohort is er nauwelijks verschil ten gevolge van de culturele oriëntatie van de leerling aantoonbaar.

Voor de afstromers op MAVO-opleidingen geldt dat zij in cohort '93 een lagere score behaald hebben dan hun klasgenoten die niet afgestroomd zijn. In de beide andere cohorten is het verschil met deze klasgenoten duidelijk kleiner. Opstomers blijken in '93 een hogere score behaald te hebben dan hun klasgenoten die niet opgestroomd zijn. Het verschil tussen opstomers en hun klasgenoten is in cohort '89 kleiner dan in cohort '93, terwijl voor cohort '95 geen verschil met cohort '93 aangetoond kan worden.

Leerlingen van MAVO-opleidingen uit een één-ouder-gezin behalen gemiddeld dezelfde tekstbegripscores als hun klasgenoten. Voor geen van de drie cohorten kan een verschil met de klasgenoten aangetoond worden.

Voor zowel het 'advies aan het einde van de basisschool', als het 'hoogste opleidingsniveau van één van de ouders' kan een positieve relatie met de tekstbegripscores aangetoond worden. Voor geen van beide variabelen blijkt het effect te variëren tussen cohorten.

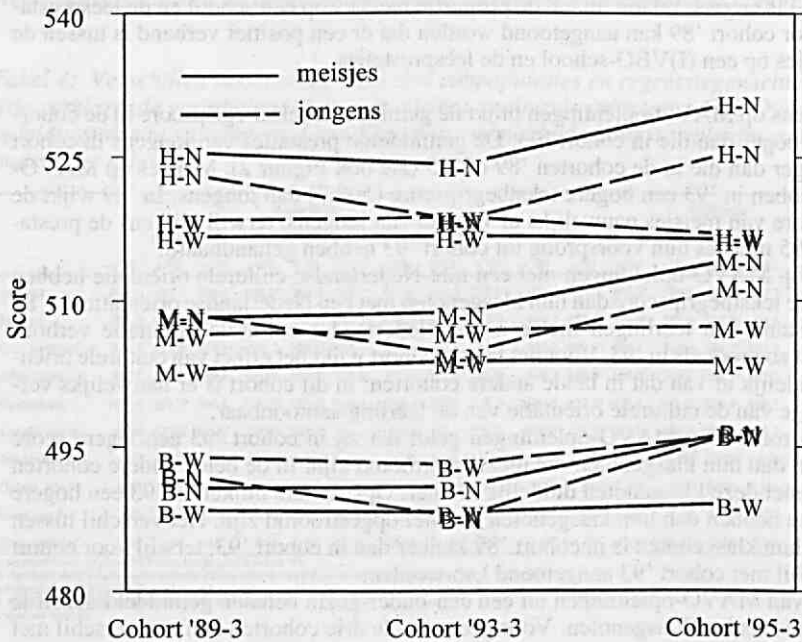
Voor jongens op HAVO/VWO-opleidingen is de gemiddelde tekstbegripscore in '93 518.3. Dit is lager dan in '89 of '95, waar de gemiddelde prestatie respectievelijk 523.4 en 525.7 is (zie Figuur 2). Meisjes op HAVO/VWO-opleidingen hebben in '93 gemiddeld een hogere tekstbegripscore dan jongens in hetzelfde cohort. In cohort '89 is het verschil tussen de gemiddelde score van jongens en meisjes kleiner dan in '93, en in cohort '95 is het verschil tussen de tekstbegripscores van jongens en meisjes net zo groot als in '93 (zie Figuur 2).

HAVO/VWO-leerlingen met een niet-Nederlandse culturele oriëntatie doen qua leesprestaties niet onder voor hun klasgenoten, noch in cohort '93, noch in cohort '89 of in cohort '95. Ook voor op- en afstromers kan in HAVO/VWO-opleidingen geen verschil aangetoond worden met leerlingen die niet op- of afstromen. Dit geldt voor alle drie de cohorten.

Concentreren we ons vervolgens op de wiskunde prestaties dan kan op grond van de gegevens in Tabel 4 geconcludeerd worden dat de gemiddelde wiskunde prestatie van jongens in de derde klas van het (I)VBO niet veranderd is tussen cohort '89 en cohort '93. In cohort '95 is de gemiddelde prestatie echter wel hoger. Meisjes in de derde klas van het (I)VBO hebben in de

cohorten '89 en '93 voor wiskunde een vergelijkbare achterstand op jongens, en in cohort '95 is deze achterstand (verhoudingsgewijs) nog iets groter geworden (zie ook Figuur 2).

*Figuur 2: Gemiddelde prestaties per opleidingstype per cohort voor Nederlands (N) en Wiskunde (W) voor jongens (gestippelde lijn) en meisjes (ononderbroken lijn). (B: VBO; M: MAVO; H: HAVO/CVWO)*



(I)VBO-leerlingen met een niet-Nederlandse culturele oriëntatie hebben in alle drie de cohorten een vergelijkbare achterstand (op jongens). Leerlingen die in de eerste drie leerjaren naar een lager opleidingstype afgestroomd zijn, hebben in de cohorten '89 en '93 een vergelijkbare lagere score voor wiskunde (dan jongens), deze achterstand is in cohort '95 nog verder toegenomen. Opstomers in het (I)VBO hebben in de beide eerste cohorten een hogere wiskundescore dan jongens. Wel is het verschil in cohort '93 iets kleiner geworden dan dat is gebleken in cohort '89. In cohort '95 ten slotte hebben opstomers in het (I)VBO hun voorsprong geconsolideerd. Was in cohort '89 de gemiddelde wiskunde-prestatie van (I)VBO-leerlingen uit één-oudergezinnen niet te onderscheiden van die van hun klasgenoten uit twee-oudergezinnen, in de beide daarop volgende cohorten hebben zij een achterstand ten opzichte van hun klasgenoten.

In het (I)VBO is een duidelijk, maar over cohorten constante invloed van het advies aan het einde van het basisonderwijs waar te nemen. Voor het opleidingsniveau van de ouders van de leerlingen kan een positieve relatie met de wiskunde-prestaties in de beide eerste cohorten aangetoond worden. In het laatste cohort is de relatie met het opleidingsniveau van de ouders veel zwakker (en zelfs niet significant). Van het gemiddelde opleidingsniveau van de ouders van de leerlingen van een school kan alleen in cohort '89 een positief effect aangetoond worden. Het percentage meisjes op een school is in de cohorten '93 en '95 negatief gerelateerd aan de wiskunde-prestaties. Naast het genoemde individuele effect van sekse, is er

in deze cohorten ook een effect van de groepssamenstelling; hoe groter het percentage meisjes hoe lager de prestaties (van jongens en meisjes).

Voor jongens en meisjes in MAVO-opleidingen kan een constant verschil in gemiddelde prestatie tussen de drie cohorten aangetoond worden; de gemiddelde wiskunde prestaties van meisjes zijn in alle drie de cohorten lager dan die van jongens (zie ook Figuur 2). Leerlingen met een niet-Nederlandse culturele oriëntatie hebben net zoals afstromers in het MAVO in alle drie de cohorten een zelfde achterstand op hun klasgenoten. Opstromers hebben daarentegen hun voorgrond in cohort '89 verhoogt in cohort '93, en dit verschil geconsolideerd blijkt hun prestaties in cohort '95. Leerlingen uit één-oudergezinnen tenslotte doen qua wiskunde prestaties in geen van de drie cohorten onder voor hun klasgenoten uit twee-oudergezinnen.

Van het advies aan het einde van de basisschool zijn de effecten te traceren op de wiskunde prestaties in de derde klas van MAVO-opleidingen; hoe hoger het advies, hoe hoger de prestaties. In het eerste cohort ('89) is het sterker dan in de beide andere cohorten. Het opleidingsniveau van de ouders heeft in alle drie de cohorten een vergelijkbare positieve relatie met de wiskunde prestaties. Voor de twee resterende variabelen, het gemiddelde opleidingsniveau en de proportie meisjes, is in geen van de cohorten een effect op de prestaties traceerbaar.

De wiskunde prestaties van jongens op HAVO/VWO-opleidingen zijn in alle drie de cohorten hoger dan die van meisjes (zie Figuur 2). De prestaties van leerlingen met een niet-Nederlandse culturele oriëntatie en de gemiddelde prestatie van afstromers verschillen niet van de gemiddelde prestaties van jongens. Voor opstromers kan alleen in cohort '95 aangetoond worden dat zij een hogere prestatie behaald hebben dan hun klasgenoten die niet opgestroomd zijn. Voor leerlingen uit één-oudergezinnen valt alleen in het laatste cohort een achterstand aan te tonen. Voor alle drie de cohorten is een effect van het advies aan het einde van de basisschool aantoonbaar. Dit effect is in cohort '89 wel sterker dan in de beide andere cohorten. De invloed van het opleidingsniveau van de ouders op de wiskunde prestaties is positief en constant over de cohorten. Van het gemiddelde opleidingsniveau is geen invloed op de prestaties traceerbaar. De proportie meisjes op een opleiding heeft alleen in het laatste cohort een negatieve invloed op de wiskunde prestaties van de leerlingen.

Natuurlijk is het niet alleen van belang of aangetoond kan worden dat een verschil significant is, hetzij tussen verschillende cohorten, hetzij tussen verschillende groepen leerlingen. Maar waar het natuurlijk werkelijk om gaat is de vraag naar de grootte van deze effecten. Vandaar dat in de onderstaande tabel de verschillen tussen de onderscheiden groepen nogmaals weergegeven zijn, maar nu in termen van effectgrootte.

Tabel 5: Verschillen tussen onderscheiden subpopulaties in termen van effectgrootte.

|                  | Nederlands       |                   |                   |                  |                   |                  |                  |                  |                  | Wiskunde         |                   |                  |                  |                   |      |          |                   |                   |
|------------------|------------------|-------------------|-------------------|------------------|-------------------|------------------|------------------|------------------|------------------|------------------|-------------------|------------------|------------------|-------------------|------|----------|-------------------|-------------------|
|                  | (I)VBO           |                   |                   | MAVO             |                   |                  | HAVO/VWO         |                  |                  | (I)VBO           |                   |                  | MAVO             |                   |      | HAVO/VWO |                   |                   |
|                  | c'89             | c'93              | c'95              | c'89             | c'93              | c'95             | c'89             | c'93             | c'95             | c'89             | c'93              | c'95             | c'89             | c'93              | c'95 | c'89     | c'93              | '95               |
| Jongens          | .29 <sup>b</sup> | n.v.t.            | 0.52 <sup>a</sup> | .27 <sup>b</sup> | n.v.t.            | .42 <sup>a</sup> | .31 <sup>a</sup> | n.v.t.           | .44 <sup>b</sup> | n.s.             | n.v.t.            | .26 <sup>b</sup> | n.s.             | n.v.t.            | n.s. | n.s.     | n.v.t.            | n.s.              |
| Meisjes          | n.s.             | .20 <sup>a</sup>  | .36 <sup>b</sup>  | n.s.             | .25 <sup>a</sup>  | .33              | n.s.             | .33 <sup>a</sup> | .41 <sup>b</sup> | n.s.             | -.42 <sup>a</sup> | n.s.             | n.s.             | -.32 <sup>a</sup> | n.s. | n.s.     | -.20 <sup>a</sup> | n.s.              |
| Cult. Oriëntatie | .40 <sup>b</sup> | -.28 <sup>a</sup> | .62 <sup>b</sup>  | .32 <sup>b</sup> | -.25 <sup>a</sup> | .72 <sup>b</sup> | n.s.             | n.s.             | n.s.             | n.s.             | -.21 <sup>a</sup> | .31 <sup>b</sup> | n.s.             | -.22 <sup>a</sup> | n.s. | n.s.     | n.s.              | n.s.              |
| Afstromers       | .32 <sup>a</sup> | -.30 <sup>a</sup> | .39 <sup>b</sup>  | .42 <sup>b</sup> | -.33 <sup>a</sup> | .72 <sup>b</sup> | n.s.             | n.s.             | .29 <sup>b</sup> | n.s.             | -.43 <sup>a</sup> | n.s.             | n.s.             | -.26 <sup>a</sup> | n.s. | n.s.     | n.s.              | n.s.              |
| Opstromers       | .39 <sup>b</sup> | .25 <sup>a</sup>  | .64 <sup>b</sup>  | n.s.             | .47 <sup>a</sup>  | .47 <sup>b</sup> | .25 <sup>a</sup> | n.s.             | .47 <sup>b</sup> | n.s.             | .35 <sup>a</sup>  | n.s.             | .20 <sup>b</sup> | .27 <sup>a</sup>  | n.s. | n.s.     | n.s.              | .24 <sup>b</sup>  |
| 1-ouder gezin    | .33 <sup>b</sup> | n.s.              | .53 <sup>b</sup>  | .39 <sup>b</sup> | n.s.              | .57 <sup>b</sup> | .33 <sup>b</sup> | n.s.             | .47 <sup>b</sup> | .24 <sup>b</sup> | -.15 <sup>a</sup> | .33 <sup>b</sup> | n.s.             | n.s.              | n.s. | n.s.     | n.s.              | -.54 <sup>b</sup> |

a: effectgrootte t.o.v. gemiddelde score jongens in hetzelfde cohort '93;

b: effectgrootte t.o.v. zelfde subpopulatie in cohort '93;

n.s.: niet significant hetzij t.o.v. jongens in cohort '93, hetzij t.o.v. zelfde subpopulatie in cohort '93.

De in Tabel 5 gerapporteerde effectgroottes geven in de kolom voor cohort '93 de effect-grootte aan voor het verschil met de gemiddelde prestaties van jongens in dat opleidingstype. In de beide andere kolommen is de effectgrootte voor dezelfde subpopulatie aangegeven t.o.v. het gemiddelde van deze groep in cohort '93. Zo is de effectgrootte voor het verschil in tekst-begripscores van jongens in het (I)VBO tussen de eerste twee cohorten .29, en de effectgrootte voor het verschil tussen de beide laatste cohorten .52. Voor de leesvaardigheid van (I)VBO-leerlingen blijkt dat er duidelijke verschillen tussen de cohorten aantoonbaar zijn, maar dat die - behoudens voor meisjes - relatief gelijk zijn voor de onderscheiden deelpopulaties. Ook wordt in één oogopslag duidelijk dat het hier niet om bijzonder grote effecten gaat, en dat er, net zoals voor het tekstbegrip in het MAVO, een duidelijke minimum voor cohort '93 te constateren is.

Voor Wiskunde zijn er voornamelijk verschillen tussen de onderscheiden subpopulaties aantoonbaar. Slechts in een enkel geval is er een (significant) verschil tussen de cohorten, zodat de berekening van effectgroottes in de meeste gevallen achterwege kan blijven.

De bovenstaande analyses en gesuggereerde interpretaties zijn alleen zinvol als de drie steekproeven, i.c. de drie cohorten, vergelijkbaar zijn. Dat wil zeggen: drie aselechte steekproeven uit de populatie leerlingen. Om hier enig inzicht in te verkrijgen kan zowel de respons in de drie steekproeven, als de intraklassecorrelatie vergeleken worden. Deze laatste zijn met name relevant omdat de steekproeven volgens verschillende principes getrokken zijn. Voor de beide VOCL-cohorten zijn immers steekproeven eersteklassers getrokken, waarvan ten behoeve van deze analyse de prestaties in de derde klas vergeleken zijn. Terwijl in het laatste cohort ('95-3) een steekproef scholen is getrokken, waarbij de kans op trekking evenredig was aan de grootte van de school.

In Tabel 6 is de respons per steekproef (cohort) weergegeven. Uit deze tabel blijkt dat in cohort '89 maar liefst 85.6 procent van de aangeschreven scholen aan het VOCL-cohort onderzoek meegewerkt heeft. Van deze scholen heeft in cohort '93 slechts 31.8 procent wederom medewerking verleend. Omdat dit aantal als te gering beschouwd werd, is een tweede steekproef aangeschreven (waarvan 53% meegewerkt heeft), zodat het totaal aantal deelnemende scholen 333 bedraagt. In de derde steekproef hebben 71 van de 120 aangeschreven scholen deelgenomen. Het geheel overziend is daarmee de respons in het eerste cohort duidelijk hoger dan die in de beide andere cohorten, en de respons in de tweede steekproef duidelijk lager<sup>2</sup>.

Tabel 6: Aantallen aangeschreven en deelnemende scholen per cohort.

|                         | Cohort '89  | Cohort '93  | Cohort '95 |
|-------------------------|-------------|-------------|------------|
| 1e ronde: aangeschreven | 440         | 381         | 120        |
| 1e ronde: medewerking   | 381 (85.6%) | 121 (31.8%) | 71 (59.2%) |
| 2e ronde: aangeschreven |             | 400         |            |
| 2e ronde: medewerking   |             | 212 (53.0%) |            |

Combineren we dit met de gegevens in Tabel 1, waarin o.a. tot uiting gekomen is dat de verdeling van leerlingkenmerken over de verschillende steekproeven nogal verschilt van steekproef tot steekproef. We kunnen derhalve slechts concluderen dat de verschillen tussen de steekproeven een alternatieve verklaring zijn voor de geobserveerde verschillen tussen de drie cohorten.

Een tweede manier om inzicht te krijgen in de verschillen tussen de drie steekproeven is na te gaan of de proportie tussen scholen variantie varieert van steekproef tot steekproef (zie Tabel 7).

Tabel 7: Intraklassecorrelatie per opleidingstype per cohort.

|          | Nederlands   |              |              | Wiskunde     |              |              |
|----------|--------------|--------------|--------------|--------------|--------------|--------------|
|          | Cohort '89-3 | Cohort '93-3 | Cohort '95-3 | Cohort '89-3 | Cohort '93-3 | Cohort '95-3 |
| (1)VBO   | .070         | .158         | .078         | .192         | .230         | .091         |
| MAVO     | .144         | .200         | .065         | .267         | .263         | .167         |
| HAVO/VWO | .154         | .180         | .049         | .232         | .148         | .067         |

In Tabel 7 is de intraklassecorrelatie ( $\rho = S^2_{\text{school}} / S^2_{\text{totaal}}$ ) per opleidingstype per cohort weergegeven. Deze zijn met name voor de beide VOCL-cohorten bijzonder groot. Voor bij voorbeeld MAVO-opleidingen is de intraklassecorrelatie in Cohort '93 maar liefst .20 voor Nederlands, en .26 voor Wiskunde. In alle combinaties van vak, opleidingstype en cohort is de intraklassecorrelatie in cohort '95 het laagste. De intraklassecorrelatie, zoals geobserveerd in de beide cohorten '89 en '93, is ook opmerkelijk veel hoger dan in vergelijkbare studies (die in dezelfde periode uitgevoerd zijn; vergelijk bijvoorbeeld: Kuhlemeier & Van den Bergh, 1989 voor Nederlands), waar de intraklassecorrelaties veeleer in de orde van grootte van het derde cohort zijn). Ondanks de zorgvuldige steekproefprocedures heeft hier wellicht een systematische vertekening plaatsgevonden.

#### 4. CONCLUSIE

We hebben gedemonstreerd dat voor tekstbegrip Nederlands de prestaties van derdeklassers in cohort '93 lager zijn die van derdeklassers in hetzij cohort '89, hetzij cohort '95. Ook in diverse subpopulaties wordt deze trend zichtbaar. Hierbij moet aangetekend worden dat, in termen van effectgroottes, het om kleine tot uiterst kleine effecten gaat. Een oppervlakkige conclusie zou echter zijn dat de invoering van de basisvorming geleid heeft tot verwarring bij docenten, waardoor het onderwijs in deze periode minder adequaat geweest is.

Voor het vak Wiskunde blijken voornamelijk verschillen tussen subpopulaties binnen cohorten aantoonbaar. Deze verschillen tussen subpopulaties blijken goeddeels dezelfde in de andere cohorten. Uitzonderingen zijn bij voorbeeld jongens in het (1)VBO die in cohort '95 een iets hogere prestatie behaald hebben, en het zelfde geldt voor opstomers op het HAVO/VWO, die ook in cohort '95 een iets hogere prestatie verwezenlijkt hebben. Een eerste conclusie zou dan ook kunnen zijn dat leerlingen, nu de basisvorming eenmaal goed is ingevoerd, of wel op hetzelfde niveau gebleven zijn als voor de basisvorming, of vooruitgegaan zijn wat betreft hun prestaties voor het vak Wiskunde.

De tweede verwachting bij de invoering van de basisvorming was dat de basisvorming een panacee zou zijn tegen de ongewenst grote verschillen tussen leerlingen; na invoering van de basisvorming zouden de verschillen tussen leerlingen kleiner worden, zo was de verwachting. Op grond van de in deze studie beschikbaar gekomen gegevens moet geconcludeerd worden dat deze hoop ijdel is gebleken; de verschillen tussen leerlingen lijken in de loop van de tijd eerder groter geworden dan dat zij afgenomen zijn.



Probleem bij deze studie is de vergelijkbaarheid van de steekproeven. Er zijn aanwijzingen dat de beide VOCL-cohorten (i.e. cohort '89-3 en '93-3) niet optimaal vergelijkbaar zijn. Bovendien kan men twijfels hebben over de vergelijkbaarheid van de VOCL-gegevens en die uit het derde cohort. Immers, een dergelijke afname in intraklassecorrelatie is ongekend. Dit bemoeilijkt de interpretatie van de vergelijking van de prestaties. Immers, kunnen er geen verschillen aangetoond worden omdat de steekproeven niet volledig vergelijkbaar zijn, of zijn er geen verschillen tussen de onderscheiden populaties? Dit neemt niet weg, dat indien er bijzonder grote effecten van de basisvorming geweest zouden zijn, deze ook in de huidige studie aangetoond zouden moeten kunnen worden (vergelijk Van den Bergh, Brandsma & Bosker, 2000).

Een laatste punt van commentaar betreft de gebruikte instrumenten. Het is een alom bekend gegeven dat de prestaties van leerlingen variëren al naar gelang het gebruikte meetinstrument. Of, anders gezegd: het meetinstrument zelf is een (vaak veronachtzaamde) bron van variatie (zie o.a. Kuhlemeier & Van den Bergh, 1998 voor Nederlands, en Kuhlemeier, Kleintjes & Van den Bergh voor andere vakken waaronder Wiskunde). Conclusies zijn derhalve vrijwel altijd instrumentgebonden. Het is dan ook betreurenswaardig dat in de onderhavige studie geen opzet zoals bij peilingsonderzoek gebruikelijk gevolgd kon worden. Bij peilingsonderzoek wordt immers een vak in zijn gehele breedte onderzocht, waarbij leerlingen steekproefsgewijs (verschillende) steekproeven items voorgelegd krijgen. Hier hebben we ons noodgedwongen moeten beperken tot relatief smalle toetsen, die het curriculum van de beide vakken niet in hun geheel (kunnen) dekken.

Ondanks verschillende steekproef en meettechnische problemen kunnen we samenvattend concluderen dat de beoogde doelen - een algehele prestatieverhoging, en een duidelijke verkleining van de verschillen tussen leerlingen - op geen enkele wijze gerealiseerd zijn. Wellicht kan aangevoerd worden dat de metingen te snel na invoering plaatsgevonden hebben, maar als vijf jaar na invoering geen resultaat gevonden kan worden, dan ligt het meer voor de hand dat een dergelijk effect ook niet opgetreden is. Het is ons inziens veel waarschijnlijker dat de verwachtingen destijds - om politieke redenen - veel te hoog opgevoerd zijn. Zeker wanneer bedacht wordt dat de basisvorming - vrijwel - budgettair neutraal ingevoerd moest worden.

## NOTEN

1 Hoe realistisch dergelijke verwachtingen zijn laten wij in het midden (zie o.a. Van den Bergh & Kuhlemeier, 1990). Wij willen er alleen op wijzen dat een algehele prestatieverhoging, i.c. voor alle onderdelen van een vak, minder waarschijnlijk is als er tevens gepropageerd wordt dat er binnen een vak ook een duidelijke verschuiving van de aandacht plaats dient te vinden (bij voorbeeld van meer aandacht voor de mondelinge taalvaardigheid en het gebruik van taal in communicatieve situaties).

2. Een logit model waarin aangenomen wordt dat de frequentie van respons in verhouding tot het totaal aantal aangeschreven scholen *niet* varieert van steekproef tot steekproef past dan ook niet bij de geobserveerde aantallen ( $\chi^2 = 244.92$ ;  $df = 2$ ;  $p < .001$ ).

## LITERATUUR

- Bergh, H. van den, Brandsma, H.P., Bosker, R.J., mmv Meijer, J. & Elshout-Mohr, M. (2000). *Veranderingen in prestatieniveaus in de periode 1991-1997. Een verkenning van de gevolgen van de invoering van de basisvorming*. Universiteit Twente: Center for applied research on education.
- Bergh, H. van den & Kuhlemeier, H. (1990). De haalbaarheid van eindtermen Nederlands in de basisvorming. *Pedagogische Studiën*, 67, 1-15.

- Cohen, J. (1988). *Statistical power analysis in the behavioural sciences*. Hillsdale (NJ): Lawrence Erlbaum Associates.
- Inspectie van het onderwijs (1999). *Werk aan de basis. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het Onderwijs.
- Kuhlemeier H. & Bergh, H. van den (1989). *De proefpeiling Nederlands: een onderzoek naar de haalbaarheid van peilingsonderzoek in het voortgezet onderwijs*. Arnhem: Cito.
- Kuhlemeier, H. & Bergh, H. van de (1998). Relationships between language skills and task effects. *Perceptual and Motor Skills*, 86, 443-463.
- Kuhlemeier, H., Kleintjes, F. & Bergh, H. van den (1999). *Effect van toets, toetsvorm en vraagtype op de moeilijkheid van de afsluitingstoetsen basisvorming*. Arnhem: Cito.
- Ministerie van Onderwijs en Wetenschappen (1989). *De ontwikkeling van eindtermen voor het basisonderwijs en de basisvorming in het voortgezet onderwijs (deel 1)*. Zoetermeer: Ministerie van Onderwijs en Wetenschappen.
- Verhelst, N.D., Glas, C.A.W., & Verstralen, H.H.F.M. (1993). *OPLM: One parameter logistic model*. Computer program and manual. Arnhem: Cito.
- Werf, M.P.C. van der, Kuypers, H. & Lubbers, M.J. (1999). *Achtergrondkenmerken van leerlingen en opbrengsten van het voortgezet onderwijs*. Groningen: GION.

## APPENDIX

Stel  $Y_{ij}$  is de score van leerling  $i$  in school  $j$ , en de dummy-variabelen  $X89_{ij}$ ,  $X93_{ij}$ , en  $X95_{ij}$  geven de respectievelijke cohorten aan (i.e. de dummy-variabelen staan aan (zijn één) als een score in een bepaald cohort geobserveerd is, wanneer een score in een van beide andere cohorten geobserveerd is, staat de dummy-variabele uit (is nul)). We kunnen de geobserveerde scores dan schrijven als:

$$\begin{aligned}
 Y_{ij} = & X89_{ij} (\beta_1 + e_{1ij} + u_{10j}) + \\
 & X93_{ij} (\beta_2 + e_{2ij} + u_{20j}) + \\
 & X95_{ij} (\beta_3 + e_{3ij} + u_{30j}),
 \end{aligned} \tag{1}$$

$(i = 1, 2, \dots, I; J = 1, 2, \dots, J).$

Per cohort wordt één gemiddelde geschat ( $\beta_1$ ,  $\beta_2$  en  $\beta_3$ ) en twee residuele scores. De eerste residuele score ( $e_{1ij}$ ,  $e_{2ij}$ , en  $e_{3ij}$ ) geeft telkens de afwijking van leerling  $i$  van het gemiddelde van school  $j$  aan, de tweede residuele score ( $u_{10j}$ ,  $u_{20j}$ , en  $u_{30j}$ ) geeft de afwijking van school  $j$  ten opzichte van het geschatte gemiddelde. Gewoontegetrouw nemen we aan dat de residuele scores normaal verdeeld zijn met verwachte waarde nul, en een variantie van respectievelijk  $S^2_{e1ij}$ ,  $S^2_{e2ij}$ , en  $S^2_{e3ij}$ , en  $S^2_{u10j}$ ,  $S^2_{u20j}$ ,  $S^2_{u20j}$ , en  $S^2_{u30j}$ .

Hoewel het model uit Vergelijking 1 beschouwd kan worden als één basismodel, is het feitelijk nauwelijks informatief. Immers, er wordt geen rekening gehouden met verschillen tussen opleidingstypen, die naar alle verwachting groot zullen zijn. En, daarnaast wordt aangenomen dat de varianties niet variëren tussen opleidingstypen. Vandaar dat een tweede model geconstrueerd is, waarin wel verschillen tussen opleidingstypen gespecificeerd zijn. Stel,  $VB_{ij}$ ,  $MA_{ij}$ , en  $HV_{ij}$ , zijn dummy-variabelen die aangeven of de leerling wiens score geanalyseerd wordt in een (I)VBO-klas, een MAVO-klas of een HAVO/VWO-klas zit (i.e.  $VB_{ij}$ , staat alleen 'aan' als een leerling in een (I)VBO-klas zit, etc.). Het te analyseren model, waarin zowel de verschillen tussen cohorten, als de verschillen tussen opleidingstypen geschat worden, kan dan geschreven worden als:

$$\begin{aligned}
Y_{ij} = & \beta_1 * VB_{ij} + \beta_4 * MA_{ij} + \beta_7 * HV_{ij} + \\
& X89_{ij} (\beta_2 * VB_{ij} + \beta_5 * MA_{ij} + \beta_8 * HV_{ij}) + \\
& X95_{ij} (\beta_3 * VB_{ij} + \beta_6 * MA_{ij} + \beta_9 * HV_{ij}) + \\
& [VB_{ij} (X89_{ij} * e_{1ij} + X93_{ij} * e_{2ij} + X95_{ij} * e_{3ij} + u_{10j}) + \\
& MA_{ij} (X89_{ij} * e_{4ij} + X93_{ij} * e_{5ij} + X95_{ij} * e_{6ij} + u_{20j}) + \\
& VH_{ij} (X89_{ij} * e_{7ij} + X93_{ij} * e_{8ij} + X95_{ij} * e_{9ij} + u_{30j})],
\end{aligned} \tag{2}$$

(i = 1, 2, ..., I; j = 1, 2, ..., J).

In Vergelijking 2 worden negen vaste parameters geschat ( $\beta_1$  tot  $\beta_9$ ). Deze kunnen als volgt geïnterpreteerd worden. De eerste drie ( $\beta_1$ ,  $\beta_4$  en  $\beta_7$ ) geven de gemiddelde score van derdeklassers per opleidingstype in cohort '93 aan. De tweede set van drie vaste parameters ( $\beta_2$ ,  $\beta_5$  en  $\beta_8$ ) geven per opleiding het verschil aan tussen cohort '89 en cohort '93. Het derde en laatste trio vaste parameters ( $\beta_3$ ,  $\beta_6$  en  $\beta_9$ ) representeren het verschil tussen cohort '93 en cohort '95. Naast deze gemiddelde scores worden twaalf variantiecomponenten geschat:  $S^2_{e1ij}$  tot  $S^2_{e9ij}$ , en  $S^2_{u10j}$ ,  $S^2_{u20j}$  en  $S^2_{u30j}$ . Met andere woorden: het model staat toe dat de variantie tussen leerlingen verschilt tussen opleidingstypen en tussen cohorten; de variantie tussen scholen kan wel verschillen tussen de drie onderscheiden opleidingstypen, maar is constant over de cohorten.

Het model volgens Vergelijking 2 kan eenvoudig uitgebreid worden met verklarende variabelen. Hierbij kan het effect van zo'n verklarende variabele geschat worden per cohort. Om redenen van eenduidigheid zijn de effecten van de toegevoegde verklarende variabelen geschat voor cohort '93. De effecten van deze variabelen in de beide andere cohorten zijn, analoog aan de gemiddelde scores per cohort, geschat als afwijking van het effect van deze variabele in het cohort '93.

*Oud-minister Wim Deetman stelt na lezing evaluatierapport basisvorming vast:*

## Onderwijsveld verdient compliment en stimulans

Interview door Piet Helmholt

“Een aantal scholen voor voortgezet onderwijs is bewust, actief en succesvol bezig met de realisering van de basisvorming. Een ingrijpend veranderingsproces, dat geleidelijk plaats vindt en gelukkig niet ten koste gaat van de leerlingprestaties. Anders gezegd: de mammoet-tanker van de onderwijsinnovatie vaart - dankzij de inzet van de professionals in de scholen - stabiel uit de haven en begint op koers te komen. Dat verdient een compliment en waar nodig stimulans.” Dat zegt oud-onderwijsminister Wim Deetman na lezing van het inspectierapport.

Volgens Deetman kun je tot die conclusie echter alleen komen als je bij bestudering van de rapportage ook ‘de juiste bril’ opzet: “Je kunt met deze evaluatiegegevens heel snel op de loop gaan. Zo in de trant van: er is nog veel niet gerealiseerd. Wie dat doet heeft geen oog voor het object van de evaluatie: het onderwijsveld. Onderwijsvernieuwing vraagt een enorme inspanning van scholen. Leraren, directies, bestuurders en begeleiders realiseren het in een maatschappij die niet statisch is. Voortdurend is er verandering en vernieuwing. De samenleving is op tal van punten in beweging en dat heeft een enorme doorwerking naar de scholen. Denk alleen maar aan de gevolgen van ICT in de scholen. Bovendien: bij het vaststellen van de wet op de basisvorming was veel onduidelijk. Ik denk aan zaken als de eindtermen, de vakken, scholing en bijscholing en leerplanontwikkeling. Als je dát voor ogen houdt, past niets anders dan het geven van een schouderklop aan de mensen in het onderwijsveld. De mammoet-tanker blijkt goed in beweging te komen. De politiek en de onderwijsinspectie moeten zich dat terdege realiseren.”

### SUGGESTIES

Evenals oud-staatssecretaris Jacques Wallage is Wim Deetman van mening “dat de politiek in de evaluatiegegevens geen aanleiding moet zien voor grote en fundamentele aanpassingen van de basisvorming. Het schip is amper de haven uit. Een hernieuwd politiek debat over de structuurverandering maakt het onderwijs kapot. Het onderwijs kan niet reageren op een politiek die vandaag dit wil en morgen weer wat anders wenst. Geef scholen en de mensen voor de klas daarentegen nu de gelegenheid om de dingen waar ze mee bezig zijn verder te ontwikkelen. Natuurlijk kun je voor onderdelen wel aanpassingen aanbrengen en op een aantal fronten zal ook bijsturing moeten plaatsvinden om overbelasting tegen te gaan, maar koerswijziging moet zeker niet. De overheid moet scholen ondersteunen om tot verbetering en verfijning te komen. Ik zie in dat verband ook een nadrukkelijke taak voor de onderwijsinspectie. Inspecteurs moeten - voor zover dat al niet gebeurt - adviseren. Ze moeten met suggesties ter verbetering komen. Zo in de trant van ‘we hebben op die en die school gezien, dat het kan, misschien kunnen jullie daar van leren’. Er is bij de onderwijsinspectie veel know-how. Daar moet het onderwijsveld beter van worden. Scholen hebben zeker behoefte aan handreikingen en suggesties.”

“De laatste jaren” - zo vervolgt Wim Deetman - “zie je in het onderwijsveld - maar ook elders in de samenleving - een overheid die meer op afstand komt te staan. Wetgever en politiek kunnen maar heel beperkt sturen. Maar dat betekent niet dat de overheid - en dan heel speci-

fiek - de inspectie scholen niet zou moeten stimuleren om in de goede richting te gaan.

Inspecteurs nemen heel veel waar en het is uitstekend dat ze dit fenomeen van de basisvorming gezamenlijk ook heel systematisch benaderen. Dat is perfect. Maar de Inspectie heeft niet slechts tot taak het vergaren van statistiekgegevens. De inspecteur moet met de scholen die hij onder zijn hoede heeft goede contacten onderhouden en meewerken aan het bevorderen van de kwaliteit van het onderwijs. Ik ben dan ook van mening, dat een inspecteur - naast het werken in teams - een vast aantal scholen heeft en daar een relatie mee onderhoudt."

## MIDDEL

Volgens oud-minister Deetman, hebben scholen daar behoefte aan: "Het evaluatierapport over de basisvorming kan daarbij een goed middel zijn. Dit rapport moet bij scholen niet op de boekenplank worden bijgezet. Het moet gebruikt worden als middel om te komen tot kwaliteitsverbetering. De inspecteur moet zich opstellen als bondgenoot van de scholen. Weliswaar op afstand, maar betrokken bij onderwijsverbetering. Natuurlijk moeten er ook regels worden gecontroleerd, maar daarnaast moet er een volwassen relatie bestaan tussen de inspecteur en de scholen. Samen probleemoplossend bezig zijn. Ook de Inspectie moet zich realiseren dat er in het onderwijsveld gewerkt wordt door professionele mensen. Het onderwijs is een sector waarin de meeste hoogopgeleiden werkzaam zijn. Docenten zijn mensen met een HBO- of universitaire opleiding. Professionals dus. Behandel ze dan ook zo en heb er vertrouwen in. Aan de andere kant zeg ik tegen scholen: wees bereid om te leren van datgene wat de Inspectie je voorhoudt."

## ICT

Terugkijkend naar de invoering van deze onderwijsinnovatie zegt Wim Deetman: "Toen ik - samen met Nel Ginjaar - het wetsvoorstel had ingediend, zeiden we tegen elkaar: één ding staat vast; over vijftig jaar ziet het er anders uit, dan we nu met stelligheid denken. Kijk - je formuleert een aantal doelstellingen, maar de kindertjes en de onderwijzers heb je niet aan een touwtje. Bovendien verandert de samenleving in hoog tempo. Wie had toen kunnen denken dat bijvoorbeeld ICT zo'n vlucht zou nemen. We hadden er toen een vage notie van. Voor mij staat vast dat ICT op een geweldige manier in de komende jaren gaat interveniëren. Het onderwijs - zelfs het klassikale onderwijs - wordt in de komende tien vijftien jaar op z'n kop gezet door ICT. Dat staat voor mij als een paal boven water. De Inspectie doet er verstandig aan in een vervolgevaluatie het ICT-gebeuren en de consequenties ervan voor het onderwijsleerproces mee te nemen."



*Oud-staatssecretaris Wallage vindt inspectierapport 'opbeurend' en noemt*

## **Evaluatie basisvorming van grote kwaliteit**

Interview door Piet Helmholt

"Met de evaluatie van de basisvorming heeft de Inspectie van het onderwijs uitstekend werk geleverd. Ik vind het een prestatie van grote kwaliteit. Methodologisch gezien zit het knap in elkaar en je mag dit evaluatieproject zondermeer wetenschappelijk verantwoord noemen." Dat zegt de Groninger burgemeester Jacques Wallage, de man die als staatssecretaris van Onderwijs nadrukkelijk aan de politieke wieg van de basisvorming stond.

Volgens de oud-bewindsman doen zijn politieke vrienden er verstandig aan om nu niet een hernieuwde discussie te beginnen over de basisvorming: "In plaats daarvan moet de staatssecretaris stevig het heft in handen nemen als het gaat om de voortgaande innovatiestrategie. De kracht van het inspectierapport is dat het laat zien dat er beweging is ontstaan. Er zijn voldoende aanknopingspunten om nu verder te gaan. Laat de Inspectie - in overleg met de afzonderlijke scholen - nu afspraken gaan maken over het vervolgtraject. Die afspraken worden vervolgens vastgelegd in een ontwikkelingsplan, dat de status krijgt van een contract. Doelstelling daarbij is: de vervolmaking van de basisvorming. Ik heb er alle vertrouwen in dat het dan moet lukken. Onderwijsinnovatie is iets van de lange adem. Het inspectierapport is in mijn ogen dan ook heel opbeurend en biedt zeker voldoende perspectief voor de toekomst."

### **PEDAGOGISCH VERANTWOORD**

Oud-staatssecretaris Wallage typeert het inspectierapport als een 'pedagogisch verantwoord verhaal': "De tolerantie waarmee de Inspectie naar het geheel kijkt valt op. Het had ook veel harder gekund. Daar is terecht niet voor gekozen. De Inspectie heeft niet vanuit een ivoren toren en op afstand het proces geëvalueerd. Het rapport getuigt ervan dat men met de beide benen in het onderwijsveld is gaan staan. Rekening houdend met het gegeven dat onderwijsinnovatie tijd en energie kost. Ik heb daar veel waardering voor. Het gaat immers ook om een eerste meting van een invoeringsproces, dat jaren duurt. Je kunt nu eenmaal niet verwachten dat op dit moment alle doelstellingen zijn gerealiseerd. De echte meerwaarde van de basisvorming zal pas over een langere termijn meetbaar zijn. Zo heb ik het rapport ook gelezen en dan ben ik erg optimistisch op grond van de bevindingen van de Inspectie. Met de opstellers van het evaluatierapport stel ik vast dat er alle reden is om de weg te vervolgen en dit als een heel belangrijke eerste fase te zien. Vergeet niet dat er bij de invoering van de basisvorming veel verzet was. De besluitvorming had een hoog politiek karakter. Er was een enorme beduchtigheid rond de vraag of het allemaal op schoolniveau wel zou lukken. In het onderwijsveld was er - om het maar zacht uit te drukken - een heel gemengde stemming. Nu is objectief vastgesteld dat er al heel wat is gebeurd. Het is de kracht van dit inspectierapport dat wetenschappelijk is vastgesteld dat er beweging is ontstaan en het geeft positieve aanknopingspunten om nu verder te gaan."

### **OPTIMALE SCHOOLGROOTTE**

Het verbaast Wallage niet dat juist de grotere scholen meer aspecten van de basisvorming hebben gerealiseerd dan de kleinere: "Ik heb zelf altijd de bredere scholengemeenschap ver-

dedigd omdat er vlees op de botten moest komen. Je hebt een zekere kritische massa nodig om tot een professionele organisatie te komen. Je moet voldoende volume hebben om je rooster te kunnen maken. Er is iets als een optimale schoolgrootte, ook uit onderwijskundig oogpunt. Ik denk daarbij aan scholen met een leerlingenaantal van minimaal 800. In de afgelopen jaren zijn veel scholen gefuseerd, zonder dat in alle gevallen het gewenste resultaat werd bereikt. Dat blijkt ook uit het evaluatierapport. Het kwetsbaarste resultaat werd geboekt op nevenvestigingen, die men uiteindelijk na de fusie gewoon als autonome scholen liet voort-sudder. Ik had gehoopt dat de vorming van een scholengemeenschap zou leiden tot het formeren van een brede onderbouw in één gebouw. Dan heb je het voordeel dat je kunt werken aan een pedagogisch klimaat voor 12-14-jarigen in de volle breedte in één gebouw. In plaats daarvan zie je dat brede scholen in de marktpositie van hun havo/vwo-afdeling zo veel last hebben van het feit dat ze een VBO/MAVO-afdeling aan boord hebben, dat ze deze 'bevlakte broertjes' maar liever in een apart gebouw situeren. Dat zegt iets over de immense concurrentieverhoudingen waaronder scholen moeten werken. We zijn in Den Haag principiëler geweest dan veel scholen tot dusver. Een aantal scholen heeft de ruimte gebruikt om zo gese-gmenteerd mogelijk te blijven werken, terwijl men in de politiek hoopte dat scholen de eerste jaren vooral zouden gebruiken voor het realiseren van een onderwijsaanbod, waarin kinderen zich breder kunnen oriënteren. Technisch-organisatorisch gezien zijn de voorwaarden geschapen, maar in de praktijk van alle dag is de basisvorming het kwetsbaarst gebleken daar waar geen optimaal gebruik is gemaakt van die ruimte."

### SLAP AFTREKSEL

Jacques Wallage denkt daarbij terug aan de start van de basisvorming: "We hebben met vallen en opstaan van de basisvorming vooral een onderwijsinhoudelijke vernieuwing gemaakt. Om politieke redenen kwamen de structurelementen tussen haakjes te staan. Mij restte destijds niets anders dan om in het bestaande stelsel - met handhaving van de eindexamens en de structuur - een nieuw programma en een nieuwe werkwijze in te voeren. Dat is van het begin af aan een hachelijke onderneming geweest. In de Eerste Kamer deed men het voorkomen alsof het de grootste onderwijsrevolutie aller tijden was. Tegen het CDA heb ik ooit gezegd: deze basisvorming is meer van u dan van mij. Voor voorstanders van geïntegreerd voortgezet onderwijs is basisvorming een slap aftreksel. Wim Deetman heeft met het advies van de Wetenschappelijke Raad voor het Regeringsbeleid (WRR) een zeker fundament gelegd voor deze basisvorming door iedere structuurdiscussie tussen haken te plaatsen. Ik weet nog steeds niet of de onderwijsinhoudelijke impulsen niet sterker waren geweest als we een aparte onderbouwschool hadden gemaakt. Als dat was gebeurd, waren alle seinen op een andere kleur gezet. Maar het is niet gebeurd. De basisvorming is beleidsarm ingevoerd. We wisten ook dat dat consequenties zou hebben voor de invoering. Belangrijk werd vooral: hoe motiveer je leerkrachten om de onderwijskundige idealen, die in de wet staan, in de school tot leven te brengen. Ik geloof dat uit het inspectierapport in geen enkel opzicht blijkt, dat de toenmalige keuzes voor pedagogische noties niet goed zijn geweest. Integendeel en dat vind ik bemoedigend."

### TWIJFELS

De oud-staatssecretaris geeft toe dat hij bij de indiening van het wetsvoorstel basisvorming de nodige twijfels had: "In die tijd heb ik tegen Jos van Kemenade eens gezegd: is het eigenlijk alle moeite wel waard. Mag ik mijn handtekening - en daarmee ook een beetje die van jou - er wel onder zetten? Brengt het allemaal wel voldoende verandering teweeg in het onderwijs? De dynamiek in de samenleving is zoveel groter dan in het onderwijs. Moet het aan-

passingsproces niet een veel sterkere impuls krijgen? Uiteindelijk heb ik de wet wel verdedigd en in het Staatsblad gebracht. Voor mij stond vast: de politiek moet ophouden met discussiëren. Je kunt het onderwijs niet aandoen om nog eens vijftien of twintig jaar te bakkeleien over de hervorming van het voortgezet onderwijs en besluitvorming uit te stellen. Bovendien: het is een compromis. Structuurdoorbreking zit er niet in, maar er is wel een geweldige uitdaging om programmatisch en pedagogisch het roer om te gooien. In de derde plaats wist ik dat het veld weliswaar niet stond te trappelen van ongeduld, maar dat er wel profesie genoeg is om het geleidelijk aan ter hand te nemen."

### INHOUDELIJKE AMBITIE

Voor burgemeester Wallage blijft het een vraag "of de onderwijskundige pretenties van de basisvorming waargemaakt kunnen worden in een onderwijsstelsel, waarin het eindexamen zo vooruit schaduwt. Maar een hernieuwde politieke discussie daar over openen is niet verstandig. De bewindslieden moeten inzetten op het verbeteren van de resultaten, die als matig of onvoldoende worden beoordeeld. Uit de beschrijving van de Onderwijsinspectie blijkt hoe groot de variëteit is in het veld. Het is opmerkelijk dat scholen met dezelfde faciliteiten zo verschillend kunnen omgaan. We moeten misschien nog meer dan in het verleden leren van de succesvolle school. De verschillen tussen goede en slechte scholen zitten behalve in de schaal ook in aspecten als onderwijskundig leiderschap, teamvorming, visie-ontwikkeling en de inhoudelijke betrokkenheid van ouders. Heel belangrijk blijkt het wel of niet hebben van inhoudelijke ambitie in de schoolleiding. Als er iets kwetsbaar blijkt te zijn, dan is het wel het onderwijskundig handelen van docenten. Uit de inspectierapportage blijkt dat duidelijk. Dat is ook zo'n voordeel van deze evaluatie. Er wordt nu - voor het eerst - dwars door de school heen gekeken. Een deel van de in kaart gebrachte problemen zijn niet die van de basisvorming, maar het zijn de problemen van onderwijs, dat professioneel niet aan de maat is. In een dergelijke situatie is elke innovatieopdracht er één te veel. Veel van de lacunes die nu boven tafel komen, waren ook aan het licht gekomen als er geen enkele innovatie was geweest. Het zegt meer over de staat van het onderwijs, dan over de noodzaak en mogelijkheid tot innovatie."

### TROTS

Oud-staatssecretaris Wallage ziet in het inspectierapport zijn mening bevestigd dat goede scholen bewust met visieontwikkeling bezig zijn: "Ik vind het slecht dat er nog steeds scholen zijn, die geen verifieerbare onderwijsdoelen formuleren, die gezamenlijk waar worden gemaakt. Dergelijke scholen hebben geen missie en kunnen die dan ook niet doorvertalen naar de onderwijspraktijk. Het ontbreekt hen aan professionele kaders. Je ziet bij die scholen dan ook dat de leraar iemand is die amechtig probeert de klas voor te blijven en sterk methodegebonden aan het werk is. Het is uitstekend dat de Inspectie de lacunes laat zien. Eerlijk en duidelijk. Maar laat het voor de politiek en het veld een stimulans zijn om te werken aan verbeteringen. We moeten af van de kledderige sfeer. De onderwijstraditie heeft een ongehoord vermogen tot klagen. Behalve boeren kennen we geen andere beroepsgroep, die in een zaaltje bijeengekomen, vooral de nadruk legt op de problemen. Je hebt ontzettend veel energie nodig om daar doorheen te komen. Laten we - met het rapport in de hand - nu eerst maar eens vaststellen waar we goed in zijn. Als je internationaal de kwaliteit van ons onderwijs meet, mogen we best trots zijn. Leerlingbegeleiding nemen we kennelijk serieus en er zijn meer van die aanknopingspunten. Ik zou willen beginnen met voor zowel de bovenbouw als de onderbouw de winst te incasseren. Als we iets willen verbeteren aan het beroepsbeeld van de docent, ligt hier het startpunt. Er gebeuren in het onderwijs veel goede dingen!"

## UITDAGEN

Volgens Wallage zou de Onderwijsinspectie daarbij een belangrijke rol kunnen vervullen: "Door bijvoorbeeld het in kaart brengen van types scholen: een leerling-georiënteerde school, een examen-georiënteerde school e.d. Je kunt vervolgens die plaatjes aan scholen voorleggen, gekoppeld aan de vraag: wat willen jullie? Wat is het perspectief van jouw school? Scholen moeten door de Inspectie worden uitgedaagd. Ik vind dat we zover moeten gaan dat de Inspectie nu met individuele scholen afspraken moet gaan maken over het opstellen van een veranderplan. Een beschrijving, waarin de school precies aangeeft wat men gaat veranderen, waarom en wanneer. Ik hoop dat het bevoegd gezag en de Minister dan ook tegen scholen, die achter blijven, durft op te treden. Er mag dan wel een grotere bestuurlijke vrijheid zijn ontstaan in de afgelopen jaren, de gemeenschap eist terecht wel kwaliteit. We hebben een eeuw lang een naar de inhoud toe afstandelijke overheid gehad. Die wissel gaat nu om. We hebben autonomere scholen, maar het gaat de Minister nu - in het hart van de zaak - om goede kwaliteit in die instellingen. Daarom moeten we in mijn ogen contracten gaan sluiten met autonome scholen over de vraag welke inhoudelijke prestaties men wil gaan leveren. In dat opzicht is het inspectierapport goud waard. Langs de lijnen van dit rapport kunnen we een innovatiestrategie uitzetten en targets kiezen voor de komende vijf jaren. Daarbij moet je de moed hebben om heldere en concrete doelstellingen te formuleren en er ook en gericht de nodige faciliteiten voor in te zetten."

## ONTWIKKELINGSPLAN

Jacques Wallage noemt de Inspectie in dat verband 'contractpartij': "Scholen maken - in opdracht van de Inspectie - een ontwikkelingsplan. Dat is een contract tussen de school (ouders, leerlingen en docenten), het bevoegd gezag en de Minister. Doelstelling is: het vervolmaken van de basisvorming. Over een dergelijk contract kan best onderhandeld worden. Ook de Inspectie mag best zeggen: wij hadden verder willen gaan, maar dit is er uiteindelijk uit gekomen. In onderwijsland hebben we te maken met verdragende mechanismen, maar de samenleving stelt kwaliteitseisen. Zo worden de vakken techniek en verzorging noodzakelijk geacht. Scholen, die in dat opzicht achter blijven, moeten worden aangepakt. De authentieke rol van de Minister - het toezien op wat in scholen gebeurt en wat de samenleving wil - moet weer voldoende uit de verf komen. Een belangrijke les uit de evaluatie van de basisvorming is voor mij dan ook dat de Minister veel meer nog degene moet zijn die tegen het onderwijs zegt: dit is wat de samenleving van u verwacht."



# De evaluatie van de basisvorming geëvalueerd

Roel. J. Bosker, *Toegepaste Onderwijskunde, Universiteit Twente*

## 1. DE EVALUATIE-DIMENSIES GEËXPlicITEERD

Evaluatie-onderzoek onderscheidt zich op een aantal wezenlijk aspecten van fundamenteel sociaal-wetenschappelijk onderzoek (zie bijv. Rossi e.a., 1999 of Swanborn, 1999). Bij een bespreking van de evaluatie van de basisvorming is het daarom essentieel die aspecten te benoemen, en, naast het algemene methodologische oordeel dat in de kern van de zaak gebaseerd is op criteria als betrouwbaarheid en validiteit, deze een prominente plaats te geven. In het navolgende zal ik met name stilstaan bij:

1. de organisatie van de evaluatie
  2. de doelstellingen, de belanghebbende partijen en het waarderingskader
  3. de implementatie van de basisvorming
  4. de evaluatie van een volledig dekkende, dat wil zeggen alle scholen voor voortgezet onderwijs in Nederland betreffende innovatie
  5. de doelmatigheid van het beleidsprogramma
  6. de conclusies die naar aanleiding van de evaluatie worden getrokken.
- Slechts terzijde zal ik ook nog iets te berde brengen over de algemene methodologische eisen van betrouwbaarheid en validiteit.

## 2. DE ORGANISATIE VAN DE EVALUATIE

De evaluatie van de basisvorming heeft, anders dan bij de evaluatie van andere grootschalige innovaties in het Nederlandse onderwijs, een lange aanloop gekend met vele commissies en veel plannen, zoals we in het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer kunnen lezen. Dit heeft uiteindelijk geresulteerd in een situatie waarbij de Inspectie van het Onderwijs de facto de rol van evaluator toebedeeld heeft gekregen, zoals ze die rol eerder ook had bij de evaluatie van de Wet op het Basisonderwijs. Daar is veel voor te zeggen (de voordelen komen in het verloop van het betoog meermalen aan bod), maar er is ook iets op tegen: hoe onafhankelijk kan de inspectie zich een oordeel vormen én vervolgens uitspreken over een onderneming waar de politiek zich zo mee verbonden heeft? Die principiële, voor sommigen wellicht academische vraag wordt actueel bij lezing van het rapport 'Werk aan de Basis' (Inspectie van het Onderwijs, 1999). Waar is het vraagstuk van de gelijke kansen gebleven, en waarom is dat als een lang gekoesterde doelstelling voor elke voorgenomen vernieuwing in het voortgezet onderwijs en uiteindelijk ook voor de basisvorming, niet als te evalueren doelstelling opgenomen? Bij lezing van het inleidende artikel van Peters-Sips kan men het langzaam naar de achtergrond en tenslotte geheel uit beeld zien geraken. Het onafhankelijkheidsvraagstuk komt ook naar voren bij lezing van de conclusies, die toch wel erg verhullend geformuleerd zijn. Ik kom hier later nog op terug, en memoreer het nu om aan te geven dat het mijns inziens geen academische kwestie is. Een evaluatie behoort extern te worden uitgevoerd, en hoe men het ook wendt of keert, de Inspectie van het Onderwijs is en (wat er ook gebeuren moge in de toekomst), was dat zeker ten tijde van deze evaluatie-onderneming, niet een onafhankelijk van de politiek opererende instantie, daar zij veel onder de ministeriële verantwoordelijkheid. De evaluatieliteratuur en de evaluatiestandaarden van de Joint Committee on Standards for Evaluation (1994) stellen namelijk dat "*The persons conducting the evaluation should be both trustworthy and competent to perform the*



*evaluation, so that the evaluation findings achieve maximum credibility and acceptance*". En het is met name het 'trustworthy' gedeelte dat onafhankelijkheid impliceert. Om elke schijn van belangenverstrengeling of politieke beïnvloeding te vermijden ware een constructie van een onafhankelijk evaluatieconsortium, onder leiding van een onafhankelijk deskundige, met daarbinnen een rol voor de Inspectie van het Onderwijs, te prefereren geweest. De oorspronkelijk ingestelde coördinatiecommissie evaluatie voortgezet onderwijs (de CCE) voldeed aan dit desideratum, de uiteindelijk gekozen constructie niet.

### 3. DE EERSTE EVALUATIEFASEN

Wie na dit wat 'zure' begin nu een verdere opeenhoping van geconstateerde tekortkomingen verwacht komt bedrogen uit. In veel opzichten is de evaluatie namelijk voorbeeldig. Te beginnen in het onderkennen en betrekken van de verschillende belanghebbenden, 'the stakeholders', die voorafgaand aan de evaluatie geconsulteerd worden over de doelstellingen en uitwerkingsaspecten van de basisvorming alswel over de evaluatie zelf. Hetgeen in het kader van de evaluatie in kaart is gebracht lijkt mij dekkend, op één uitzondering na. Wellicht is namelijk in deze fase de doelstelling zoals de WRR die nog voor de basisvorming formuleerde als 'vermindering van ongelijkheid van kansen in het onderwijs door vermindering van een te vroege definitieve school- en/of beroepskeuze' (WRR, 1986, p. 260) zoek geraakt? Weliswaar zou men kunnen verwijzen naar de wetstekst, waarin alleen nog maar 'uitstel van studie- en beroepskeuze' aan de orde komt, maar de ontstaansgeschiedenis van deze doelstelling is hierin gelegen dat dit eigenlijk een intermediair doel - zo niet een middel - is voor het uiteindelijke te bereiken ideaal van toegankelijk onderwijs.

Sterk is de gemaakte keuze voor een accent op de vraag wat er van de basisvorming terecht is gekomen in de onderwijspraktijk, zonder overigens de effectvraag zelf te verwaarlozen. Effecten zonder deugdelijke implementatie zijn immers niet aannemelijk, en de conclusies zouden wel heel erg schraal worden als er geen enkele indicatie gegeven zou kunnen worden van mogelijke implementatiefouten of -tekortkomingen. Degelijk en voorbeeldig, maar dat is de Inspectie van het Onderwijs inmiddels wel toevertrouwd, is het opgestelde en gehanteerde waarderingskader: zonder duidelijke criteria met daarop helder geformuleerde standaarden kan er immers geen (navolgbaar) waarde-oordeel worden uitgesproken. En de kern van elke evaluatie is dat dat uiteindelijk gebeurt. Het is overigens wel een saillant detail dat de scholen zelf de normen strenger zouden hebben gesteld, dat wil zeggen dat ze de standaarden op de evaluatie-criteria hoger geformuleerd wilden zien. Ware het gebeurd, de evaluatie-uitkomsten zouden dramatisch zijn geweest.

### 4. DE IMPLEMENTATIE VAN DE BASISVORMING

Een belangrijke rol in de gehele evaluatie hebben de lesbezoeken gespeeld. Ruim 7000 lessen zijn door inspecteurs geobserveerd en beoordeeld. De kernvraag van elke innovatie is in hoeverre deze doordringt op de werkvloer, in de klas dus, en deze lesbezoeken vormden de bouwstenen voor dit onderdeel van de evaluatie. Hier blijkt op een tweetal aspecten het grote voordeel van de betrokkenheid van de Inspectie van het Onderwijs bij de evaluatie: de non-respons is nihil (een droomwereld voor de opdrachtonderzoekers, die overigens ook zouden dromen van de kennelijk voor de evaluatie ter beschikking staande middelen en menskracht) en de inhoudelijke en vakdidactische deskundigheid van de observatoren staat buiten kijf en maken daarmee ook hoog-inferentiële observaties mogelijk. Maar er is ook een onmiskenbaar nadeel: de basis van de oordeelsvorming over de lespraktijk is de observatie. Net als bij het Integraal en het Regulier School Toezicht (afgekort als IST en RST), waar dezelfde weg wordt bewandeld, dringt zich dan de vraag op wat de validiteit is van de aldus verzamelde gegevens. Meten

we hier de dagelijkse lespraktijk? Het antwoord op deze vraag is negatief. Wat hier gemeten wordt is hetgeen een docent vermag als hij of zij zijn best doet. We meten met andere woorden niet het gemiddelde instructiegedrag, maar het plafond. De werkelijke implementatie van de basisvorming wordt met andere woorden overschat. Het hoger leggen van de evaluatie-standaarden biedt hier geen soelaas, omdat men steeds de tegenwerping kan maken dat tussen de waargenomen en de werkelijke alledaagse instructiepraktijk een enorm gat zou kunnen bestaan. Men zou dus kunnen zeggen dat dit deel van de evaluatie antwoord geeft op de vraag of de docenten de basisvorming *kunnen hanteren*, maar niet of ze hem *de facto hanteren*.

Een tweede opmerking betreft de verbanden die worden gezocht tussen de mate van implementatie en de opbrengsten voor leerlingen. De gerapporteerde associaties zijn niet sterk, hetgeen geen bevreemding wekt daar Van den Bergh, Peters-Sips & Zwarts in dit themanummer immers rapporteren dat docenten wel, maar scholen amper van elkaar verschillen in de mate van invoering van de basisvorming. Juist omdat de Inspectie van het Onderwijs een waarderingskader hanteerde was het echter mogelijk (en dat is het nog steeds) om na te gaan of scholen die de basisvorming volgens het waarderingskader voldoende hebben ingevoerd hogere opbrengsten realiseren bij hun leerlingen dan scholen die in dit opzicht ver beneden de maat blijven. De contrasten zijn dan scherper. Mocht deze vergelijking niet goed uitvallen, dan is er een stevig debat te voeren over de causale ketenstructuur die aan het basisvormingsbeleid ten grondslag ligt.

## 5. DE DEKKINGSGRAAD VAN DE BASISVORMING

Een probleem dat bij de evaluatie moest worden opgelost betrof het gegeven dat er geen vergelijkingsgroep voorhanden was; de basisvorming is immers over de volle breedte ingevoerd. Het evaluatiemodel waarbij een controlegroep (geen basisvorming) met een experimentele groep (wel basisvorming) wordt vergeleken kon dus niet worden toegepast. Over de evaluatie van volledig dekkende programma's leert de literatuur ons dat een vergelijking van de situatie vóór met die van na de invoering een goed alternatief is. Dat is hetgeen in de cohort-vergelijkingen (zie het artikel van Van den Bergh, Bosker & Brandsma in dit themanummer) aan de orde is. Maar veel informatie over lespraktijk of over de beheersing van de algemene vaardigheden kon niet ontleend worden aan eerder verricht onderzoek, en het is hier dat het geconstrueerde beoordelingskader zijn vruchten afwerpt. Niet de vraag of het beter gaat dan vroeger, maar of het onderwijs in de basisvorming voldoet aan heldere standaarden gedefinieerd op door iedereen onderschreven criteria, is hier aan de orde.

## 6. DOELMATIGHEIDSASPECTEN

In de evaluatie als geheel wordt weinig aandacht besteed aan de efficiency van het onderwijs en de veranderingen daarin als gevolg van de basisvorming. Gegeven de geconstateerde onderwijsresultaten is de doelmatigheidskwestie in evaluatie-onderzoek gericht op de vraag of dezelfde resultaten niet tegen lagere kosten te realiseren zouden zijn geweest, of dat met een andere even dure innovatie betere resultaten zouden kunnen worden geboekt. Het vergt op zich al een studie of deze vragen te beantwoorden zijn, maar de kritische bespreking van Boersma in dit themanummer over de mogelijke overladenheid van het curriculum in de basisvorming raakt wel aan dit probleem. In essentie beweert hij immers dat een minder inspanningen vergende dieptestrategie bij het aan de orde stellen van hetgeen in de kerndoelen geïmpliceerd wordt tot dezelfde en in sommige opzichten mogelijk zelfs betere resultaten zou leiden dan de geconstateerde breedtestrategie (waarbij alle mogelijk leerstofinhouden, ook die slechts zijdelings raken aan de kerndoelen of die wel in methoden voorkomen maar niet bedoeld zijn in de kerndoelen aan de orde worden gesteld).

De doelmatigheidskwestie kunnen we ook op een andere manier aan de orde stellen: hoe efficiënt worden de leerlingen naar hun diploma geleid, ofwel hoe verhoudt zich de feitelijke tot de nominale studieduur? Het aantal dubblures is na de invoering van de basisvorming opvallend afgenomen. De kans dat een leerling ongeschonden de eerste drie leerjaren in het voortgezet onderwijs overleeft is in zes jaar tijd gestegen van 0,70 naar 0,80. Dat is een forse verbetering van de doelmatigheid van het voortgezet onderwijs, die bovendien niet gepaard lijkt te zijn gegaan met een daling van het onderwijspeil.

## 7. DE CONCLUSIES TEGEN HET LICHT GEHOUDEN

Anders dan de bijdragen in dit themanummer, waar op een zeer prudente manier de gegevens worden gepresenteerd en conclusies omzichtig worden geformuleerd, laat lezing van het evaluatiedocument bij uitstek, namelijk het algemeen rapport (Inspectie van het Onderwijs, 1999), licht zien tussen de gepresenteerde bevindingen en de vervolgens uitgesproken slotsom. De eerste hoofdconclusie luidt (idem, p.122): *"Het peil van het jeugdonderwijs is in de afgelopen vijf jaar toegenomen."* De cohortvergelijkingen (Van der Werf e.a., 1999) die voor deze conclusie mede als basis dienen, laten echter zien dat er geen verhoging van het peil is gerealiseerd. De deelname aan het HAVO en VWO mag dan zijn toegenomen, gegeven echter de startpositie van de leerlingen bij entree in het voortgezet onderwijs zijn er in de tijd geen verschuivingen in prestaties te constateren (zie de bijdrage van Van den Bergh, Bosker & Brandsma in dit themanummer). De winst die is geboekt ten gevolge van de toegenomen deelname aan de bovenstroom van het Nederlandse onderwijs (HAVO en VWO) is, zo zou men het voorgaande anders kunnen stellen, de vrucht van het primair en niet van het secundair onderwijs. Bovendien, zoals Creemers (1999) aangeeft, roept het waarderingskader zoals dat gebruikt is bij de uitgevoerde peilingen (zie het artikel van Kleintjes, Zwarts & Peters-Sips elders in dit themanummer) op zijn minst vragen op: als we met de basisvorming een minimum beheersingsniveau nastreven, waarom dan een beheersingsniveau per opleidingstype geformuleerd en waarom is het dan akkoord als één op de drie inhoudsdomeinen niet voldoende beheerst wordt? Een slotsom die dichter bij de bevindingen ligt is mijns inziens dat het peil niet gedaald is, en dat is op zich reeds een opmerkelijke bevinding gegeven de omstandigheid dat de invoering van de basisvorming nagenoeg budgettair neutraal is verlopen, dat de evaluatie al vijf jaren na de start heeft plaatsgevonden, en dat de doelmatigheid gestegen is.

Komen we bij de tweede conclusie: *"Voor nogal wat leerlingen is er feitelijk sprake van uitstel van studie- en beroepskeuze. Inhoudelijk kan er op dit punt nog veel verbeterd worden door beter aan te sluiten op de individuele mogelijkheden en interesses van leerlingen"* (idem, p.123). Deze conclusie wordt gestaafd met de bevinding dat de onderwijsstroom waarin de leerlingen terecht komen meer dan vroeger afwijkt van de basisschool, en het gegeven dat er nogal wat op- en afstroombepalingen plaatsvinden. Maar helaas wordt deze conclusie niet verbonden met het oogmerk waarop dat uitstel van studie- en beroepskeuze gerealiseerd moest worden: een verbetering van de toegankelijkheid van het Nederlandse onderwijs, ofwel een meer gelijkwaardige participatie van achterstandsgroepen. De evidentie die in het rapport zelf wordt aangedragen is overigens negatief: de kansen van de achterstandsgroepen zijn er niet beter op geworden, en er is nog steeds ongelijkheid in het voortgezet onderwijs. Het is merkwaardig dat bij de eindconclusies hierover niets meer gezegd wordt, temeer daar de integratie van de eerste fase, gelet op de nog steeds zeer snelle toedeling van leerlingen aan de categorale opleidingstypen (zie ook Creemers-van Wees e.a., 1999 en Doolaard e.a., 1999) niet goed van de grond is gekomen. Nu laten grootschalige onderzoeken (Dekkers e.a., 2000; Van der Werf e.a., 1999; Van den Bergh, Bosker & Brandsma in dit themanummer) weliswaar zien dat het voortgezet onderwijs niet veel meer ongelijkheid produceert dan reeds tot stand gekomen in het primair onderwijs, een zekere reductie in deze fase (herschikking in een heterogene brugperiode) treedt ook niet op.

Ten slotte de derde conclusie die op de doelstelling van modernisering en harmonisering van het onderwijsprogramma betrekking heeft: "Het onderwijsprogramma heeft zich in de eerste vijf jaar in lijn met deze doelstellingen ontwikkeld. In het (1)VBO zijn de omstandigheden voor verdere ontwikkelingen van het aanbod niet gunstig" (idem, p.123). Inderdaad zijn de basisvormingsvakken ingevoerd, en dus lijkt de modernisering en harmonisering van het aanbod over de opleidingen gerealiseerd. Maar met de ontwikkeling van de algemene vaardigheden is het nog niet goed gesteld (zie van den Bergh e.a. 2000) en ook de vakkenintegratie laat zeer te wensen over (zie ook Cremers-van Wees e.a., 1999).

## 8. DE SLOTSOM

De evaluatie van de basisvorming is op hoofdlijnen zeer goed uitgevoerd. Alle relevante aspecten zijn in kaart gebracht en gemeten, er is ook waarderingskader ontworpen en de deelonderzoeken zijn grondig en systematisch uitgevoerd. Rond de validiteit van de lesobservaties heb ik enige twijfels geuit (nader onderzoek lijkt hier geboden), en bedenkingen zijn er ten aanzien van de normen zoals die ten aanzien van de peiling van het niveau zijn gehanteerd. Alhoewel het aangedragen materiaal over het algemeen goed is (zoals ook de bijdragen in dit themanummer), zijn naar mijn mening de conclusies bij het rapport te positief geformuleerd. Bij toekomstige beleidsevaluaties dient meer institutionele onafhankelijkheid van de evaluerende instantie gegarandeerd te worden, waarbij mijn voorkeur uitgaat naar een onafhankelijk evaluatieconsortium waarin de Inspectie van het Onderwijs, het CBS, het Cito en de onderzoeksinstituten participeren onder leiding van een onafhankelijke deskundige.

## NOOT

1. De Inspectie van het Onderwijs gebruikt de begrippen standaarden, normen en criteria op een andere wijze dan ik doe in dit artikel. Daarom ter verduidelijking: met criterium bedoel ik een variabele waarop we de prestatie gaan meten (de Inspectie van het Onderwijs gebruikt hiervoor het woord standaard). De waarde op dat criterium waarboven we vinden dat het te bestuderen object voldoende scoort, duid ik aan met standaard of norm.

## LITERATUUR

- Bergh, H. van den, Bosker, R.J., Brandsma, H.P., Meijer, J., & Elshout-Mohr, M. (2000). *Algemene vaardigheden en de basisvorming* (aangeboden ter publicatie).
- Creemers, B.P.M. (1999). Het succes van de basisvorming. Commentaar op het rapport van de inspectie. In M. Peters-Sips, J. van der Linden en A. Wald (Red.) *Verder werken aan de basis* (pp. 31-34). Utrecht: Inspectie van het Onderwijs.
- Cremers-van Wees, L.M.C.M., Akkermans, L.M.W., & Brandsma, H.P. (1999). *Ontwikkelingen in de basisvorming in de periode 1990-1998: modernisering en harmonisering in de basisvorming*. Enschede: OCTO.
- Dekkers, H.P.J.M., Bosker, R.J., & Driessen, G. (2000). Complex inequalities of educational opportunities. A large-scale longitudinal study on the relation between gender, social class, ethnicity and school success. *Educational Research and Evaluation*, 6(1), 1-24.
- Doolgaard, S., Creemers-van Wees, L.M.C.M., & Bosker, R.J. (1999). *De basisvorming in 1996; beschrijving en vergelijking met de periode voor invoering*. Enschede: OCTO.
- Inspectie van het Onderwijs (1999). *Werk aan de basis. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het Onderwijs.
- Joint Committee on Standards for Evaluation (1994). *The program evaluation standards* (2<sup>nd</sup> ed). Thousand Oaks / London / New Delhi: Sage.
- Rossi, P.H., Freeman, H.E., & Lipsey, M.W. (1999). *Evaluation. A systematic approach* (6<sup>th</sup> ed.). Thousand Oaks / London / New Delhi: Sage.

- Swanborn, P. (1999). *Evalueren. Het ontwerpen, begeleiden en evalueren van interventies: een methodische basis voor evaluatie-onderzoek*. Amsterdam: Boom.
- Werf, M.P.C. van der, Kuyper, H., & Lubbers, M.J. (1999). *Achtergrondkenmerken van leerlingen en opbrengsten van het voortgezet onderwijs*. Groningen: GION.
- Wetenschappelijke Raad voor het Regeringsbeleid (1986). *Basisvorming in het onderwijs*. Den Haag: Staatswetenschappelijke Raad voor het Regeringsbeleid.



# Voor herhaling vatbaar

Gert Rijlaarsdam<sup>1</sup>, *Instituut voor de Lerarenopleiding, Universiteit van Amsterdam*

## 1. INLEIDING

De eerste keer dat ik met de resultaten van het inspectieonderzoek naar de basisvorming werd geconfronteerd was in najaar 1999. Het vakblad voor Leraren in Levende Talen bereidde naar aanleiding van de rapportages van de Inspectie een themanummer voor waarin panels van leraren in de Moderne Vreemde Talen en Nederlands de rapportages van commentaar voorzagen (Levende Talen, nr. 545, december 1999). Ik mocht bij die panelbijeenkomsten aanwezig zijn. De panelleden waren in het algemeen niet verrast over de uitkomsten: dat de gemiddelde talenles vakdidactisch onder de maat was, was wat reeds werd gevreesd. Dat er kerndoelen niet aan de orde kwamen, ook dat kon al haast niet anders. Maar dat de leerlingenprestaties zo goed waren - in de rapporten werd gemeld dat 90 procent van de leerlingen voldoende prestaties leverde - dat wilde er bij deze leraren niet in. Slecht talenonderwijs dat leidde tot zoveel voldoende prestaties....dat kon niet waar zijn. Hier moest iets niet kloppen: of de metingen van de prestaties klopten niet, of de theorie dat lessen aan bepaalde vakdidactische vereisten moeten voldoen wil er sprake zijn van goed onderwijs (en dus: goede prestaties) klopte niet. Of er was met beide iets mis. In mijn onderstaand commentaar op de artikelen over het inspectieonderzoek is bovenstaande ervaring leidraad. Mijn conclusie is dat het inspectieonderzoek een bijdrage levert aan de operationalisering van vakdidactische theorie over 'de goede les' en dat dit veel aanknopingspunten biedt voor verder onderzoek. Een tweede conclusie is dat het inspectieonderzoek tekort schiet in het leggen van verbanden tussen onderwijsaanbod en leerlingenprestaties, en dat is uitermate jammer. De kans op een validering van de wankelende theorie over effectief onderwijs, en vooral de wellicht nog wankeler vakdidactische theorie, is daarmee verkeken. Een volgende keer beter, hoop ik...

Voor ik mijn twee thema's aan de orde stel, moeten me twee dingen van het hart. Ten eerste, met het inspectieonderzoek is geenszins de *invoering* van de basisvorming geëvalueerd (zie het artikel van Peters-Sips in het themanummer over de beleidsmatige voorgeschiedenis). Zou een operatie als de invoering van de basisvorming in het bedrijfsleven op gelijke wijze zijn uitgevoerd, met overeenkomstige resultaten, dan hadden de aandeelhouders het management naar huis gestuurd, en zou de hele invoering als casus van 'bad practice' in managementinnovatieboeken zijn terechtgekomen. Het is onvoorstelbaar dat zo'n grote vernieuwingsoperatie met zo weinig investering in het uitvoerend personeel politiek draagvlak kon creëren. Ten tweede, nu er toch een evaluatieonderzoek is gedaan, een peiling naar de stand van zaken op dit moment, komt de zwarte piet van het mislukken van de basisvorming ten onrechte bij scholen en leraren te liggen. Het onderwijs blijkt niet te zijn wat het wezen moest. De Inspectie en de beleidsmakers, met de hulp van meta-deskundigen van het onderwijs als onderwijskundigen en vakdidactici - hebben het onderwijs de lat gelegd. Maar wat is hun zeggingskracht als ze niet in staat zijn om te bepalen of, en in welke mate het onderwijs sinds de invoering van de basisvorming vooruit is gegaan (zie het onthullende artikel over de crash van het VOCL- onderzoek van Van den Bergh, Bosker & Brandsma in dit themanummer), als ze niet in staat zijn om aanbodonderzoek en prestatieonderzoek aan elkaar te koppelen, als ze zelfs niet in staat blijken om goed prestatieonderzoek uit te voeren? De meta-deskundigen hebben een heel belangrijke bijdrage te leveren aan de kwaliteit van het onderwijs, maar een eerste voorwaarde is dat de kwaliteit van het eigen werk 'boven de norm' is. Natuurlijk is ook dat werk mensenwerk, met al zijn mogelijke tekortkomingen, maar het is altijd in dienst en ten dienste van het onderwijs zelf. Een leraar is ten minste 26-29 lesuren in touw

met activiteiten die direct het onderwijsleerproces betreffen: hij geeft lessen, van zekere kwaliteit, met bepaalde resultaten. Hoeveel tijd van de week beïnvloedt een meta-specialist direct de kwaliteit van onderwijs? Ik weet het: ik ben ook zo'n meta-specialist. Diep in mijn hart vind ik het beschamend om voor mezelf te constateren dat veel van mijn tijd zit in het opzetten en begeleiden van heel leuke onderzoeks- en ontwikkelprojecten, die veel mensen van de straat houden, aardige ICO-publikaties kunnen opleveren, fijn ver congresbezoek genereert, maar over de daadwerkelijke bijdrage aan de onderwijspraktijk tast ik in het duister<sup>2</sup>. Als het gaat om uitspraken over de kwaliteit van het onderwijs, moeten we ons als meta-specialisten enigszins terughoudend en bescheiden opstellen tenzij we een duidelijke bijdrage aan die kwaliteit leveren, door een hoge mate van interactie met leraren. Want die leraren zijn de spil waar het om draait: niet de scholen, niet de systemen, niet de vakdidactiek, niet de onderwijskunde. Dat vind ik een van de helderste bevindingen van het inspectieonderzoek: er zijn vooral goede en minder goede leraren, nauwelijks goede of minder goede scholen. Op de kwaliteit van die leraren moeten we ons als meta-specialisten en beleidsmakers richten: het inspectieonderzoek geeft ons daarvoor enige aanknopingspunten. Ik kom daar in het slot van deze bijdrage op terug.

## 2. DE LERAAR IS DE KWALITEIT

Flink wat kerndoelen worden niet of in onvoldoende mate aangeboden. Soms is dat terug te voeren op de schoolboeken (voor het vak Nederlands dekt het gemiddelde schoolboek kennelijk minder dan de helft van de kerndoelen), vaker op de leraren. Vooral de nieuwe kerndoelen komen niet veel aan bod. De algemene doelen komen al helemaal niet aan bod, op een uitzondering na. Merk op dat het steeds gaat om 'aan-bod komen': of er daadwerkelijk en effectief onderwijs gegeven wordt in de algemene vaardigheden die wél aan bod komen, komt niet ter sprake. Merkwaardigerwijze worden inhouden (kerndoelen) en uitvoering (didactiek) los van elkaar onderzocht: waar lessen die wel of lessen die niet deugden over gingen, dat komen we niet te weten.

De kwaliteit van de lessen in de basisvorming viel niet mee. Aan de kwaliteit van de meting wil ik niet veel afdoen: een enkel beoordelingsaspect (de 'standaard' Differentiatie, zie Van den Bergh, Zwarts & Peters-Sips in dit themanummer) is waarschijnlijk niet goed gemeten, en de standaarden correleren nogal, maar er zijn voldoende gegevens gerapporteerd waaruit blijkt dat we de resultaten kunnen vertrouwen. Vooral de aspecten 'Vakdidactiek' en 'Actief leren' scoren slecht. Deze bevinding geldt nogal wat schoolvakken. Merk overigens op dat de operationalisering van het aspect 'vakdidactiek' terecht voor alle vakken verschillend is (behalve voor de drie Moderne Vreemde Talen), en dat vergelijking op een en dezelfde schaal dus wat moeilijk is, hoewel de inspectie dat wel doet.

Het inspectieonderzoek stelde aldus vast dat het met de basisvorming (nog) niet zo goed gesteld is. Belangrijker vond ik de vaststelling dat het de docent is die de kwaliteit uitmaakt en niet de school. Men vindt dit bijvoorbeeld terug in het artikel van Van den Bergh & Zwarts (1999) dat zij speciaal voor talenleraren schreven. Voor een goede docent maakt het niet uit op wat voor een school hij werkt: hij geeft goed onderwijs, ondanks de feilen van de schoolorganisatie. We moeten dus heel veel middelen zetten op de kwaliteit van de docent, en omstandigheden scheppen waarin docenten kunnen groeien in hun professionaliteit, we moeten hun gereedschappen in handen geven waarmee zij goed onderwijs kunnen ontwerpen en uitvoeren.

Eerst maar iets over de kerndoelen. Over de gigantische versnippering en het gebrek aan samenhang. Ten eerste, het aantal kerndoelen voor de periode 1993-1998 was nogal aan de ruime kant - de herziening van 1998 heeft weinig soelaas geboden. Bij Moderne Vreemde Talen bijvoorbeeld ging het om 19 doelen, waarvan sommige zijn onderverdeeld in tien of

meer onderdelen. Zo gezien levert dat al gauw een lijstje op van 70 (zegge en schrijf: zeventig) kerndoelen en -doeltjes. Elke les een doeltje aan bod laten komen? Welke eerst? Dat is het tweede probleem: de kerndoelen vormen per vak nauwelijks een geheel. Voor leerboekenmakers en docenten is het een schier onmogelijke opgave structuur in de lijst van kerndoelen aan te brengen. Op basis van welke systematiek? De kans op grabbeltononderwijs is erg groot - een lesje over zus, een lesje over zo, zonder onderlinge samenhang.

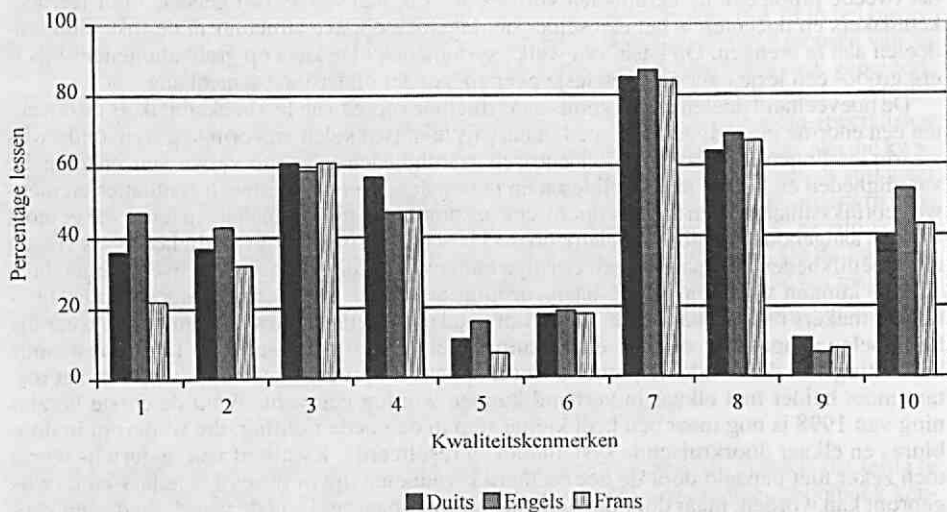
De hoeveelheid doelen en het gemis aan structuur vergen van leerboekenmakers en docenten een enorme inventiviteit om goed onderwijs te ontwikkelen en vorm te geven, onderwijs waarin leerlingen cumulatief en concentrisch vaardigheden en kennis verwerven, en leren die vaardigheden en kennis te consolideren en te transfereren naar nieuwe leersituaties en nieuwe gebruikssituaties. Een niet-hiërarchische opsomming van kerndoelen - in feite: wat er moet worden aangeboden (aanbodsdoelen), hoewel beschreven in leerlingvaardigheden - is vragen om moeilijkheden, en daarmee een ernstige onderwijskundige en vakdidactische deficiëntie.

Wat kunnen we doen om de basisvorming helderder te maken, en docenten en leerboekenmakers meer in handen te geven om goed onderwijs te ontwikkelen? Toch weer die kerndoelen aanpakken: selectie en ordening, met de nadruk op selectie. In de aanstaande herziening van de kerndoelen moet zeker nog eens 25 procent geschrapt worden; het res-tant moet helder met elkaar in verband kunnen worden gebracht. Want de eerste herziening van 1998 is nog maar een heel kleine stap in de goede richting, die wederom in dublures en elkaar doorkruisende systematieken resulteerde. Kwaliteit van onderwijs wordt toch zeker niet bepaald door de hoeveelheid kerndoelen die in twee of drie jaar onderwijs gepropt kan worden, maar door de diepgang en wendbaarheid van de vaardigheden en kennis van leerlingen. Als de onderwijskunde en de vakdidactiek in die operatie slagen, dan beschikken leerboekenmakers en docenten over een inhoudelijk samenhangend geheel aan doelen waarmee te werken valt, en dat hen in staat stelt de vrije ruimte echt als vrije ruimte te benutten. Een serieus landelijk vakdidactisch - onderwijskundig onderzoek zou deze leerdoelselectie en -ordening moeten opleveren. Liever geen 'adviescommissie' waarin ieder gelegitimeerd zijn stokpaarden mag berijden en er dan via slim taalgebruik in slaagt - geheel volgens poldermodeliaanse traditie - de bonte mengelmoois optisch als eenheid te presenteren.

Een tweede invalshoek om het onderwijs hulp te bieden bij het opzetten en uitvoeren van basisvormingsonderwijs wordt geboden door de vakdidactische theorie die de Inspectie heeft gehanteerd. Vooral op dit beoordelingsaspect scoren veel schoolvakken immers onvoldoende. Ik laat hier even achterwege dat het inspectieonderzoek de kans heeft laten liggen om die theorie te valideren - ik kom daar hieronder op terug. Waar bestaat die theorie uit? Voor de Moderne Vreemde Talen worden tien kenmerken van de standaard 'vakdidactiek' genoemd. De tien kenmerken zijn:

1. De leraar hanteert vanaf het begin in overwegende mate de doeltaal als voertaal
2. Stimuleert de leerlingen de doeltaal te gebruiken
3. Geeft communicatief onderwijs
4. Geeft functioneel grammaticaonderwijs, waarbij grammatica geen doel op zich vormt
5. Bevordert dat de communicatie in de klas tussen hem en de leerlingen en tussen leerlingen onderling zich kenmerkt door het regelmatig voorkomen van 'negotiation of meaning' en andere in het kader van de les functionele communicatiestrategieën
6. Bewerkstelligt dat leerlingen bezig zijn met de taal in levenschte situaties waarin de nodige informatie wordt opgezocht en allerlei informatie wordt vergeleken
7. Spreekt de doeltaal correct uit en gebruikt een juiste intonatie
8. Bevordert een correcte uitspraak bij leerlingen
9. Verwijst naar andere vakken/leergebieden of naar leerstof uit andere bronnen
10. Motiveert de leerlingen voor het vak en wekt interesse voor de inhoud.

*Figuur 1: Mate waarin kenmerken van de vakdidactiek in de lessen worden gerealiseerd (in procent)*



Dat de scores niet florissant zijn - het gebruik van de taal in kwestie door docent en leerling in de klas is bijvoorbeeld wel heel erg infrequent - laten we even voor wat het is. De vraag die nu eerst aan de orde is: waarom zijn die scores zo laag, en wat kunnen we docenten bieden om vakdidactisch gezien betere lessen te verzorgen? Zijn die scores zo laag omdat docenten vakdidactisch onverantwoord lesgeven, of is de vakdidactische theorie misschien minder valide dan deze lijkt? Een aanwijzing voor de laatste verklaring is de geringe interne consistentie van de schaal vakdidactische kenmerken (Van den Bergh, Peters-Sips & Zwarts in dit themanummer).

Net als hierboven bij de kerndoelen gesteld werd, zou je zeggen dat er teveel kenmerken worden genoemd en dat ze niet samenhangend worden gepresenteerd. Probleem bij dit soort rijtjes is dat het, net zoals bij de kerndoelen, limitatieve wensenlijstjes zijn. De onderlinge afhankelijkheid van de kenmerken verdwijnt uit het oog en contextgebonden conditionele relaties worden niet geëxpliciteerd. Moet een leraar bijvoorbeeld de doeltaal hanteren als leerlingen deze nog onvoldoende beheersen (kenmerk 1 en 2)? Moet elke les aan dezelfde kenmerken voldoen? Is de lesvorm niet mede afhankelijk van de lesinhoud, en van de voorafgaande en de volgende les over het onderwerp? Exploreren van nieuwe kennis of vaardigheden vergt een andere lesopzet dan het consolideren of transfereren ervan. Ik zou het veel helderder vinden als er per schoolvak enkele effectieve patronen of scenario's van lessen worden onderscheiden. (Van Hout-Wolters, Elshout-Mohr & Broekkamp (1999) noemen dat instructie-episodes.) Zulke patronen zijn herkenbare didactische eenheden voor docenten. De zinvolheid van vakdidactische kenmerken moet blijken uit de beschrijving van zo'n scenario of patroon, alsmede uit het gebleken effect op het leerresultaat. Die patronen moeten ook echt in de werkelijkheid voorkomen en niet uitsluitend leunen op prescriptieve bureau-uitspraken.

Kortom, dat de scores uit de peiling van vakdidactische kenmerken rond de veertig à zestig procent blijven steken, is op zichzelf nog niet alarmerend. Je kunt je nauwelijks onderwijs indenken waarvan de lessen op alle aspecten rond de negentig procent scoren. Voordat dit soort lijstjes wordt gepresenteerd als 'vakdidactische theorie' en de peilingresultaten worden gerapporteerd, moeten alle betrokkenen - Inspectie, onderzoekers, docenten - beschikken over



relevant voorbeeldmateriaal: eerst moet aangetoond worden dat het bestaat of bestaanbaar is. Als het goed is, beschikt de Inspectie na dit eerste onderzoek over uitstekend materiaal van docenten die voorbeeldige lessen gaven: laat de inspectie die voorbeelden erbij geven, als materialisering van verregaand jargon als 'Bevordert dat de communicatie in de klas tussen hem en de leerlingen en tussen leerlingen onderling zich kenmerkt door het regelmatig voorkomen van "negotiation of meaning" en andere in het kader van de les functionele communicatiestrategieën'. Als die voorbeelden er niet zijn, heeft het weinig zin om de kenmerken als eis te stellen. Opschoning van redundantie in het lijstje zou overigens al heel wat helpen.

Toch nog even terug naar de vraag waarom voor zoveel schoolvakken de vakdidactiek zo laag scoort. Ik sluit niet uit dat een lage score op een kenmerk een aanwijzing is voor invaliditeit van de vakdidactische theorie. Maar weinig daarvan kan steunen op empirische validatie. Maar er kunnen ook andere dingen aan de hand zijn. In Nederland zijn docenten zwaar belast met uitvoering: bijna driekwart van hun werktijd gaat op aan lesuitvoering, en daar komt ook nog eens nakijkwerk en vergadertijd bij. Het voorbereiden van lessen, al is het maar een gedegen inzicht in de structuur van het schoolboek, schiet er snel bij in. Tijd voor studie en training wordt nauwelijks in werktijd georganiseerd (lesuitval!). Vakdidactiek is een ver oord, ver verwijderd van de dagelijkse praktijk van lessen en leerlingen. Het scheppen van vakdidactisch sterkere lessen vergt extra inspanning, bovenop het beschikbare materiaal in de schoolboeken. Ook het systematisch nagaan van de meerwaarde van de didactische ingreep kost tijd. Uit het inspectieonderzoek kan worden geconcludeerd dat de condities niet zodanig zijn dat die inspanningen geleverd kunnen worden. En kennelijk is de effectiviteit (validiteit) van die vakdidactische 'theorie' niet zodanig dat die veel docenten stimuleert die inspanning te leveren.

### 3. VOOR HERHALING VATBAAR

Wat mij betreft herhaalt de Inspectie dit soort onderzoeken regelmatig. Vandaar de alluderende titel van dit artikel. Dat het onderzoek voor verbetering vatbaar is, spreekt vanzelf. Naar perfectie kunnen we alleen maar streven. Ik besluit dit commentaar dan ook met enkele aanbevelingen voor een volgend inspectieonderzoek.

#### *Prestatiemetingen*

Het onderdeel dat het slechtst is onderzocht zijn de leerlingprestaties. Zeker, de prestaties van een groot aantal leerlingen zijn onderzocht, de steekproeftrekking enzovoort waren behoorlijk, maar op essentiële punten schiet het onderzoek heel erg tekort. Eigenlijk is het treurig dat in een rapport over de kwaliteit van de basisvorming, waarin de boodschap aan scholen, beleid en docenten zo negatief is, het Nederlandse toetsinstituut zo vrijuit gaat. Op twee punten ga ik in: de steekproeftrekking en de rapportage.

#### *De steekproeftrekking*

Niet verantwoord wordt waarom de steekproeftrekking en de toewijzing van taken zodanig is gebeurd dat er over leerlingen heen geen schoolvakscore kon worden vastgesteld. Ik heb geen idee waarom men de procedures uit het landelijk peilingsonderzoek niet heeft toegepast. De kans om kwaliteit van lessen - bijvoorbeeld de vakdidactische kwaliteit - en de leerlingprestaties met elkaar te verbinden, is daarmee verkeken: echt een gemiste kans om onderwijskundige en vakdidactische theorie te valideren, en spitsere rapportages te presenteren.

#### *De rapportage*

In de rapportage moeten beschrijving en normering losgekoppeld worden, zoals dat ook voor de kwaliteit van de lessen is gebeurd. De meetmethode is erg ondoorzichtig, en erg afhankelijk van het toevallige groepje docenten en 'andere deskundigen' die de normering uitvoerden. Ik zou graag leerlingenscores op toetsen en taken willen zien en daarnaast de normering



van die scores. De kwaliteit van de normering is nu immers geen voorwerp van analyse geweest: de hele rapportage leunt tezeer op een onderdeel van de procedure waarvan betrouwbaarheid en validiteit onbekend zijn. Uit peilingsonderzoek waarbij ik eind jaren tachtig betrokken was, leerden we hoe moeilijk het normeren was. We vroegen docenten twee beoordelingen uit te voeren, op basis van enkele leerlingproducten die we hen voorlegden. De ene vraag was hoeveel procent van de leerlingen een schrijfprodukt van een bepaald niveau moest kunnen schrijven aan het eind van de basisschool. De andere vraag was hoeveel procent van de leerlingen dat niveau zouden halen. De mate waarin voorspeld niveau en werkelijk niveau uiteen liep, verschilde heel sterk tussen docenten. Als je zo'n discrepantie mag opvatten als een validiteitsaanwijzing, dan hadden heel wat beoordelaars de normeringstaak niet hoeven uit te voeren. Het *is* gewoon een heel erg moeilijke taak. Verder lijkt het me uitermate belangrijk dat (1) veel meer docenten betrokken worden bij de normering dan nu het geval is geweest: evenzeer als er wordt gegeneraliseerd over leerlingen, moet er gegeneraliseerd worden over normeerders-docenten, en (2) de instructie aan de normeerders sterk verbeterd wordt. En ten slotte, het is voor de evaluatie van de basisvorming minder interessant om te weten hoeveel leerlingen per schoolvak aan de normen voldoen: het gaat erom hoeveel leerlingen de basisvorming 'halen', ofte wel, hoeveel procent van de leerlingen voor alle schoolvakken van de basisvorming (een acceptabel deel van) de kerndoelen heeft gehaald. Dat was toch het doel van de basisvorming?

#### *De kwaliteit van de lessen*

De Inspectie heeft een bijzondere prestatie geleverd om zoveel lessen zo nauwkeurig te observeren. Gelukkig heeft de Inspectie ervoor gekozen om beschrijving en normering van elkaar los te koppelen. Hierboven gaf ik al aan dat ik hoop dat de standaarden zich zullen ontwikkelen tot scenario's in plaats van losstaand gepresenteerde kwaliteitskenmerken, en dat ten behoeve daarvan stevig onderwijskundig en vakdidactisch onderzoek kan plaatsvinden. In elk geval biedt de operationalisering zoals nu gepubliceerd een goed referentiekader voor de discussie over wat goede lessen zijn. Op twee punten zou ik graag een aanbeveling doen: de normering van de lessen en de rapportage. Wat de normering betreft, zou ik graag een procedure zien die overeenkomsten vertoont met de normering van de leerlingprestaties zoals hierboven beschreven: (1) inzet van docenten als mede-normeerders, en (2) normering aan de hand van docentprestaties, d.w.z. lessen.

Natuurlijk moeten er docenten betrokken worden bij de normering van lessen: het is toch gek dat leraren wel leerlingprestaties mogen normeren in dit inspectieonderzoek, maar niet de lessen. Daarom moet de normering ook achteraf plaatsvinden en niet zoals dat nu gebeurt door veel deskundigen - een heel enkele docent - te betrekken bij het normeren vooraf (zoals de Inspectie dat ook nu weer dit voorjaar deed voor het nieuwe beoordelingskader). In plaats daarvan zou ik liever zien dat panels van docenten video-opnamen van representatieve lessen of delen daarvan beoordelen op kernaspecten. Deze lessen zijn immers ook te schalen, net als de leerlingprestaties, bijvoorbeeld op het vakdidactische beoordelingsaspect 'mate van communicatief taalonderwijs' (voor de moderne vreemde talen). Op basis van die schaal worden lessen opgenomen die een schaalpunt representeren. In de rapportage kan dan duidelijk gemaakt worden, veel duidelijker dan nu het geval is, hoe een goede-matige- zwakke les eruit ziet, gegeven een bepaalde invalshoek. Er kan ook voor worden gekozen zogenaamde 'normlessen' voor te leggen aan het panel, en het panel te vragen hoeveel procent van de lessen een vergelijkbaar patroon zou moeten bevatten. Het grote voordeel van zo'n procedure - nadelen zijn ook gemakkelijk te genereren - is dat er via een overtuigender rapportage meteen instrumenten beschikbaar zijn voor scholen en docenten om binnen de school en de vaksectie het eigen onderwijs tegen het licht te houden. De hele educatieve infrastructuur kan hiermee zijn voordeel doen, en een bijdrage leveren aan het verbeteren van de onderwijskwaliteit.

#### 4. TOT BESLUIT

Ik hoop dat het vaststellen van de onderwijskwaliteit een periodiek onderzoek wordt, waarin peilingsonderzoek en onderwijsprocesonderzoek nauw op elkaar wordt afgestemd. Voorwaarde is wel dat docenten veel meer dan nu betrokken worden bij het onderzoek, en dat de rapportage meer mogelijkheden biedt aan scholen en docenten om de gegevens te benutten voor eigen kwaliteitsonderzoek en -verbetering. Die betrokkenheid van docenten zou ook in de analyse- en interpretatiefase tot uiting moeten komen: op vragen als 'waarom scoren deze kenmerken van vakdidactiek zo laag' moeten zij hun antwoorden kunnen uitwisselen. Voorwaarde is ook dat er perspectief bestaat op passend onderwijsbeleid: in evaluatieonderzoek is het niet ongebruikelijk om vooraf beleidsconsequenties te formuleren bij bepaalde mogelijke uitslagen. Wat zullen de beleidsconsequenties zijn als blijkt dat docenten vinden dat hoogstens 80 procent van de kerndoelen haalbaar zijn? Weer vier jaar wachten op ingewikkelde adviescommissies, maar intussen wel via memo's aan scholen beleidswijzigingen doorgeven over onontwikkeld onderwijsbeleid? Verandering op verandering op wijziging op wijziging stapelen? Of ferm zeggen: jullie hebben gelijk, we schrappen kerndoelen. En ferm zeggen: als de lessen niet goed genoeg zijn, dan trekken we 100 Mf uit voor in-service training op school, want de leraar is de kern van de kwaliteit.

Maar zo zal het wel niet werken, vrees ik. Wel hoge eisen stellen, maar niet in de condities voorzien. Het wordt weer herfst, merk ik.

#### NOOT

1. Met dank aan collega's Michel Couzijn, Martine Braaksmā en Theo Witte voor hun commentaar op een eerdere versie van dit stukje. De inzichten die ik in dit stukje verwoord, komen lang niet altijd uit eigen koker: discussies met dezen en genen, zoals met leden van de redactie van *Levende Talen*, hebben aan mijn oordeelsvorming bijgedragen.

2. Gelukkig kan ik ook nog bedenken dat ik mijn medewerking gaf aan het totstandkomen van een basisvormingsleergang, en nascholing rond basisvorming hielp opzetten en uitvoerde, maar dat zijn toch verre afgeleiden van mijn weektijdbesteding.

#### LITERATUUR

- Elshout-Mohr M., Van Hout-Wolters B., & Broekkamp H. (1999). Mapping situations in classroom and research: eight types of instructional-learning episodes. *Learning and Instruction*, 9 (1), 57-75.
- Rijlaarsdam, G. (eindredactie), den Hartog, G., van Gool, Ch., van Schooten, E. & Riemersma, A. (1999). Talen in de basisvorming. Voer voor discussie. *Levende Talen*, 545 (dec.), p. 694-787.
- Van den Bergh, H. & Zwarts, M. (1999). De kwaliteit van de lessen Nederlands en Moderne Vreemde Talen. Verschillen en mogelijkheden tot verbetering. *Levende Talen*, 545, (dec.), p. 779-783.

Commentaar op het artikel 'Kwaliteit van het vakspecifieke aanbod' elders in dit themanummer.

## Oorzaken en aanpak van overladenheid van het operationele curriculum in de basisvorming

Kerst Th.Boersma, *Centrum voor  $\beta$ -didactiek, Universiteit Utrecht; SLO, Enschede*

### 1. INLEIDING

De Inspectie concludeert in haar evaluatierapport (Inspectie van het Onderwijs, 1999a) dat het feitelijk onderwijsaanbod in de basisvorming strikt genomen niet voldoet aan de kenmerken van goed onderwijs. Geen enkele school biedt in alle vakken alle kerndoelen aan en daarmee voldoet geen enkele school aan de wettelijke verplichting (zie het artikel 'Kwaliteit van het vakspecifieke aanbod' elders in dit themanummer). Voor de algemene vaardigheidsdoelen geldt in versterkte mate wat voor de kerndoelen geldt: ze worden nog te weinig aangeboden op een wijze die als voldoende te bestempelen is (zie het artikel 'Kwaliteit van het vakoverstijgende aanbod: het voorkomen van de algemene vaardigheden').

Als oorzaak noemt de Inspectie de overladenheid van het programma. *"De kerndoelen zijn soms nogal abstract of zo gedetailleerd dat een voldoende behandeling meer tijd vraagt dan beschikbaar is. Op schoolniveau is zeker sprake van overladenheid, getuige het feit dat geen enkele school alle kerndoelen kon behandelen. Ook het feit dat de scholen meer tijd voor de basisvorming hebben ingeruimd dan geadviseerd, vormt een indicatie voor de geconstateerde overladenheid."*

Overladenheid van het curriculum is een ernstig probleem, omdat het een negatieve invloed kan hebben op scholen, leraren, leerprestaties van leerlingen en de invoering van de basisvorming. Overladenheid van het programma kan er toe leiden dat leraren ervaren dat de ruimte voor vernieuwing te gering is, dat ze tijd te kort komen en te veel in te weinig tijd moeten doen. Met een overladen programma zullen leraren geneigd zijn minder voor actieve vormen van leren te kiezen en meer voor zelfstudie en reproductie. Door overladenheid zijn leraren meer gericht op de vraag hoe zijzelf binnen de beschikbare tijd door de stof heenkomen en minder op de vraag hoe ze kunnen bijdragen aan optimalisering van de leerprocessen van hun leerlingen.

Overladenheid heeft ook negatieve invloed op de leerresultaten van leerlingen. Al meer dan dertig jaar laat allerlei onderzoek zien dat 'time on task' een van de meest bepalende factoren op het rendement van leren is. Een tekort aan leertijd staat ook op gespannen voet met constructivistische opvattingen over leren. Zo stelt Lijnse (1997) dat kennisoverdracht, in combinatie met programmatische tijdsdruk, bij de meeste leerlingen onherroepelijk tot begripsforcering zal leiden en dus tot weinig of niet hanteerbare kennis. We illustreren dat met het volgende voorbeeld.

Uit literatuur is bekend dat veel leerlingen problemen hebben met het aanvankelijk deeltjesbegrip. Ze zien in veel gevallen niet dat moleculen opgevat moeten worden als modellen en derhalve niet over dezelfde eigenschappen beschikken als stoffen. Moleculen zijn niet geel en hebben geen geur. De meeste leerstrategieën die tot voor kort beschikbaar waren, slaagden er niet in om dit en soortgelijke misvattingen te voorkomen. Pas recentelijk is daarvoor een succesvolle leerstrategie uitgewerkt (Vollebregt, 1998). Daarvoor waren echter wel 12 lessen vereist, waarin een reeks zorgvuldig gedefinieerde leeractiviteiten werden aangeboden. In de meeste scholen zou deze lessenreeks, door de druk op het programma, niet kunnen worden uitgevoerd. Toch zijn dat kennelijk voorwaarden waaraan voldaan moet worden om

een van de voor de natuurwetenschappelijke vakken meest basale begrippen met succes te kunnen aanleren.

De Inspectie heeft voor de evaluatie van de basisvorming uitvoerig onderzoek gedaan naar de kwaliteit van het onderwijsaanbod. Overladenheid als zodanig was echter niet expliciet in de vraagstelling opgenomen en is dan ook niet expliciet onderzocht. De conclusie van de Inspectie dat de teleurstellende resultaten wijzen op overladenheid van het programma is desalniettemin behoorlijk onderbouwd, ook al is nader onderzoek naar overladenheid en de oorzaken ervan gewenst. Om de overladenheid te reduceren adviseerde de Inspectie een onderscheid te maken tussen verplichte en niet-verplichte kerndoelen.

In deze discussiebijdrage wordt ingegaan op de vraag hoe overladenheid kan worden verklaard en welke maatregelen genomen kunnen worden om de overladenheid terug te brengen of te voorkomen. Daarbij geven we ook een antwoord op de vraag of we van de voorgestelde aanpassing van de regelgeving wel het gewenste effect kunnen verwachten.

## 2. OORZAKEN VAN OVERLADENHEID

### *Overladenheid als verklaring voor de onvoldoende dekking met de kerndoelen*

Alvorens nader in te gaan op de vraag of de analyse van de inspectie wel volstaat, is het goed te benadrukken dat de inspectie niet gemeten heeft in hoeverre de constatering dat de kerndoelen niet in 75 procent van de geadviseerde tijd worden gerealiseerd, toe te schrijven is aan het aantal kerndoelen. Dat kan ook niet rechtstreeks gemeten worden. Kerndoelen zijn doelstellingen op macroniveau, terwijl overladenheid tot uitdrukking komt op het microniveau. Misschien doen leraren wel de verkeerde dingen. Daar zijn hier en daar enkele aanwijzingen voor. Zo constateerde inspectie dat biologie leraren nog steeds aandacht besteden aan het thema 'mens en gezondheid', terwijl dat niet meer bij biologie hoort, maar bij verzorging (Inspectie van het Onderwijs, 1999b; p. 22).

Op grond van de constatering dat de kerndoelen niet binnen de geadviseerde tijd worden gerealiseerd, kunnen we dus niet zonder meer concluderen dat er sprake is van overladenheid. Van feitelijke overladenheid is sprake als we vaststellen dat leraren en/of leerlingen meer moeten dan binnen de daarvoor gegeven condities kan. Daarnaast is het natuurlijk goed mogelijk dat er geen sprake is van feitelijke overladenheid, maar dat leraren en leerlingen wat moet wel als zodanig percipiëren. Om te bepalen of sprake is van feitelijke overladenheid moeten we eerst de inhoud en omvang van het operationele curriculum bepalen en vervolgens analyseren of dat een voldoende dekking met de kerndoelen heeft. En om te bepalen of sprake is van gepercipieerde overladenheid zullen we leraren en leerlingen naar hun mening moeten vragen.

De vraag is natuurlijk of het verstandig is om maatregelen, zoals reductie van het aantal kerndoelen, te nemen voordat overladenheid voldoende is vastgesteld. De aanwijzingen daarvoor in het inspectieonderzoek zijn echter voldoende om overladenheid als probleem serieus te nemen. We moeten ons dan afvragen welke oorzaken we daarvoor kunnen aanwijzen. Als we inzicht in mogelijke oorzaken hebben, kunnen we ook aanbevelingen doen voor maatregelen die effectief zouden kunnen zijn bij terugbrengen ervan.

### *Het aantal kerndoelen als oorzaak van overladenheid*

Als we de redenering van de Inspectie volgen, is de eerste vraag natuurlijk of we over gegevens beschikken die er op zouden kunnen wijzen dat de kerndoelen wel binnen de beschikbare onderwijstijd kunnen worden gerealiseerd. Verondersteld zou kunnen worden dat dergelijke gegevens bij de ontwikkeling van de kerndoelen zijn verzameld. Een evaluatie van de



wijze waarop kerndoelen werden ontwikkeld (Boersma & Smit, 1997) laat zien dat dergelijke gegevens niet beschikbaar zijn.

De opdrachten van OCenW voor ontwikkeling van voorstellen voor kerndoelen voor de basisvorming en voor de eerste revisie zijn uitgevoerd door de SLO (Letschert, 1998). De opdrachten waren tamelijk precies ten aanzien van het stelsel van regelgeving en globaal ten aanzien van de inhoud van vakken. Ook conclusies van in de Tweede Kamer gevoerde discussies, zoals het besluit het vak verzorging in te voeren en de toezegging om extra aandacht aan het milieu te schenken, werden in de opdracht opgenomen.

Voorstellen voor kerndoelen werden ontwikkeld door commissies die bestonden uit ervaren leraren, vakdidactici en leerplanontwikkelaars. De SLO vervulde daarbij de rol van projectleider, om samenhang tussen de programma's te waarborgen, en om de voorstellen binnen het gestelde tijdspad op te kunnen leveren. Daarnaast vervulden SLO-medewerkers binnen de commissies de rol van secretaris. De ontwikkelde programma's werden, voordat ze aan OCenW werden opgeleverd, voor raadpleging aan belangengroepen (leraren, opleiders, uitgever) voorgelegd. De gegeven commentaren werden gewogen en, waar dat door de commissie van belang werd geacht, verwerkt. Na oplevering van de voorstellen bood OCenW de voorstellen, voor formele advisering, aan de Onderwijsraad aan. In deze wijze waarop kerndoelen werden ontwikkeld zijn de volgende punten van cruciaal belang.

- De opdrachten voor ontwikkeling en revisie van kerndoelen zijn uitgevoerd binnen de door OCenW gestelde condities. Onduidelijkheden in de opdracht werden geaccepteerd en werkende weg, in overleg met OCenW, opgelost.
- De kerndoelen werden in vakcommissies ontwikkeld en gereviseerd, op basis van inzicht, ervaring en overleg; de kerndoelen zijn daardoor het resultaat van onderhandeling.
- Bij de ontwikkeling van kerndoelen kon geen gebruik gemaakt worden van empirische gegevens over knelpunten in de klassenpraktijk en veranderingswensen van leraren, een tekort dat al door de WRR (1986) was gesignaleerd.
- Bij door OCenW gewenste vernieuwingen, bijvoorbeeld ten aanzien van algemene vaardigheden, die met de ontwikkeling en revisie van kerndoelen werden meegenomen, kon geen gebruik worden gemaakt van empirische gegevens over de uitvoerbaarheid en leerbaarheid.

De conclusie is dat de wijze van ontwikkeling van kerndoelen niet garandeert dat de ontwikkelde kerndoelen ook uitvoerbaar zijn voor leraren en dat ze tot voor leerlingen realiseerbare doelstellingen leiden. Bij de revisie van de kerndoelen in 1997 werd gesignaleerd dat procedures ontbraken om te bepalen of wat in de concept-kerndoelen werd voorgesteld in de praktijk wel kon worden uitgevoerd. Gegevens over de uitvoerbaarheid zijn niet beschikbaar.

Een andere manier om een beeld te krijgen van de invloed van kerndoelen op de overladenheid van het curriculum van de basisvorming zou zijn om gebruik te maken van gegevens die verzameld zijn in de periode vóór de invoering van de basisvorming. Gegevens daarover ontleenen we aan het onderzoek van Kuiper naar de feitelijke lespraktijk voor de natuurwetenschappelijke vakken, dat werd uitgevoerd nadat de Wetenschappelijke Raad voor het Regeringsbeleid (1986) gewezen had op het ontbreken van empirische gegevens over de bestaande praktijk in de eerste leerjaren van het voortgezet onderwijs (Kuiper & Alting, 1990). In dat onderzoek werden ook gegevens verzameld over de vraag of leraren over voldoende tijd beschikken en wat de oorzaken van een eventueel tekort aan tijd zijn. Als mogelijke oorzaken werden naast overladenheid ook lesuitval, het niveau van de leerlingen en concentratie, inzet en motivatie van de leerlingen onderscheiden. In Tabel 1 is aangegeven welk percentage van de Biologie- en Natuurkundeleraren het tekort aan tijd in het MAVO en in het HAVO/VWO (bij Biologie in leerjaar 1 en 2 en bij Natuurkunde in leerjaar 2) aan deze oorzaken toeschrijft.



Tabel 1: Oorzaken voor tekort aan tijd (in % van het aantal leraren dat aangeeft over onvoldoende tijd te beschikken). Als oorzaken zijn aangegeven (1) het programma is te overladen; (2) er is te veel lesuitval; (3) het gemiddeld niveau van de leerlingen is te laag; (4) het gebrek aan concentratie, inzet en motivatie van leerlingen; (5) anders.

|                                 | 1  | 2  | 3  | 4  | 5  | n  |
|---------------------------------|----|----|----|----|----|----|
| <b>Biologie (leerjaar 1)</b>    |    |    |    |    |    |    |
| MAVO                            | 67 | 13 | 7  | 7  | 6  | 10 |
| HAVO-VWO                        | 58 | 4  | 4  | 4  | 30 | 26 |
| <b>Biologie (leerjaar 2)</b>    |    |    |    |    |    |    |
| MAVO                            | 47 | 6  | 0  | 0  | 47 | 15 |
| HAVO-VWO                        | 62 | 4  | 0  | 0  | 34 | 26 |
| <b>Natuurkunde (leerjaar 2)</b> |    |    |    |    |    |    |
| MAVO                            | 32 | 0  | 32 | 18 | 18 | 22 |
| HAVO-VWO                        | 60 | 10 | 15 | 10 | 5  | 20 |

Tabel 1 (ontleend aan Kuipers & Alting, 1990) geeft een goed inzicht in de oorzaken die docenten aangeven voor het tekort aan tijd. Overladenheid van het programma is de meest genoemde oorzaak. De verklaring daarvoor kunnen we uiteraard niet toeschrijven aan de kerndoelen. Doordat overladenheid ook in de periode voorafgaand aan de invoering van de basisvorming door leraren als oorzaak voor gebrek aan tijd wordt aangewezen, moeten we concluderen dat overladenheid zeker niet alleen aan het aantal kerndoelen kan worden toegeschreven. Deze conclusie sluit uiteraard niet uit dat met de invoering van de basisvorming het probleem nog is vergroot.

#### *Methodetrouw en ruime interpretaties van kerndoelen als oorzaken van overladenheid*

We moeten dus zoeken naar andere verklaringen. Een plausibele verklaring voor overladenheid lijkt de methodetrouw van leraren. Als leraren de methode als programma opvatten en de methode omvat meer dan op grond van de beschikbare tijd realiseerbaar is, dan kunnen we verwachten dat leraren niet volledig door het boek komen. Er zijn empirische gegevens beschikbaar die deze veronderstelling ondersteunen.

Zo rapporteert Kuiper (1993) dat voor 75 procent (HAVO-VWO) tot 86 procent (MAVO-LBO) van de biologieleraren de jaarplanning en het leerboek nagenoeg identiek zijn aan elkaar (Kuiper, 1993, p. 61). Uit in het kader van de Third International Mathematics and Science Study uitgevoerd onderzoek blijkt dat gedurende het eerste leerjaar in de basisvorming 60 tot 73 procent van de wekelijkse lestijd van Biologieleraren voor 76-100 procent is gebaseerd op het leerboek (Kuiper et al., 1997; p. 163). In de evaluatie van de basisvorming wordt zelfs gesproken over methodeslaafsheid van leraren (Inspectie van het Onderwijs, 1999b, p. 13).

Een mogelijke verklaring voor overladenheid vinden we dus in de manier waarop leraren hun methode gebruiken. De vraag is natuurlijk of dat een voldoende verklaring is en of we een deel van de verklaring voor overladenheid ook niet aan de gebruikte methoden kunnen toeschrijven. Methodetrouw impliceert niet alleen dat leraren hun methode volgen, maar ook dat zij daar vertrouwen in hebben. Ze verwachten dat de methode dekkend is met de kerndoelen en dat onderwerpen die zij de moeite waard vinden er in aan de orde komen. Uitgevers houden uiteraard rekening met hun gebruikers. We kunnen dan ook verwachten dat methoden niet alleen een ruime interpretatie van de kerndoelen bevatten (zie Boersma, 2000, p. 43-56), maar ook extra onderwerpen die leraren belangrijk vinden. We kunnen dus verwachten dat methoden meer bevatten dan binnen de beschikbare tijd aan de orde kan worden gesteld. De

ruimte die kerndoelen voor interpretatie bieden en ruime interpretatie waar uitgevers voor kiezen, kan verklaren dat leraren in de problemen komen als zij hun methode als programma gebruiken. Leraren zullen selectiever met hun methode moeten omgaan. Dat het aantal kerndoelen dat probleem kan versterken ligt voor de hand. Overladenheid heeft, deze redenering volgend, niet een enkelvoudige oorzaak.

#### *Overladenheid en versnippering*

Het probleem van de brede interpretatie hangt samen met een ander punt. Voor een aantal vakken, zoals de natuurwetenschappelijke vakken, aardrijkskunde en geschiedenis is het traditie om bij selectie van vakinhoud vooral breed te denken en niet diep. Doordat een groot aantal onderwerpen aan de orde moet komen, verwerven leerlingen tamelijk oppervlakkige kennis die - er van uitgaand dat de onderwerpen voor de basisvorming daarop zijn geselecteerd - ook niet door hen in voor henzelf betekenisvolle contexten kan worden toegepast. Curricula met die karakteristieken zijn in de VS wel vergeleken met the Great Salt Lake van Utah. Dat is 83 miles lang, 51 miles breed en 13 feet diep (Freedman, 1998). Het hoeft geen betoog dat dergelijke curricula niet in overeenstemming zijn met de uitgangspunten van de basisvorming.

Een alternatief voor deze opvatting over de inhoud van curricula, en dat wordt al sinds Bruner (1963; zie ook Forrester, 1992) bepleit, is om te kiezen voor diepte. Diepte wordt gerealiseerd door een beperkt aantal sleutelbegrippen of -vaardigheden te selecteren en die in een grote verscheidenheid van contexten uit te werken. Dat leidt er ook toe dat de samenhang en structuur van vakken wordt versterkt. Leerlingen (en leraren) zijn van mening dat veel van de genoemde vakken uit een opeenstapeling van feiten bestaan, zonder veel structuur (zie bijvoorbeeld Roebertsen, 1996; Boersma, 1997). Ook de Inspectie van het Onderwijs (1999a) constateert dat sprake is van versnippering; de adviezen van de Inspectie aan de overheid zijn er dan ook op gericht om de overladenheid en versnippering van het leerstofaanbod terug te dringen (Inspectie van het Onderwijs, 1999a, p. 125). De oplossing voor versnippering van het leerstofaanbod is niet in eerste instantie een reductie van het aantal kerndoelen, maar de keuze voor een dieptebenadering die zich richt op een beperkt aantal sleutelbegrippen en -vaardigheden. Het hoeft geen betoog dat zo'n omslag aanzienlijk minder eenvoudig is te realiseren dan een reductie van het aantal kerndoelen.

#### *Overladenheid als systeemkenmerk*

Op grond van het voorafgaande lijkt het aannemelijk dat een groot aantal factoren bijdraagt aan de overladenheid van het curriculum van de basisvorming. Die gedachte komen we ook tegen in een van de commentaren op de situatie in het studiehuis.

*"De studiehuisen die het afgelopen jaar in Nederland verzezen zijn, leiden aan een teveel van alles. Er zit te veel lesstof in de vakken, en er zitten te veel vakken in het rooster. Er zit te veel werk in een studielastuur, en er zitten te veel studielasturen in een jaar. Van de leerlingen worden te veel zelfstandigheid en te veel vaardigheden gevraagd en van de leraren wordt te veel extra werk gevraagd waarvoor ze niet betaald krijgen. Eigenlijk is er van alles te veel in het studiehuis, behalve geld."* (Van der Plas, 2000; p. 26).

Zowel in het studiehuis als in de basisvorming hebben we kennelijk te maken met een situatie waarin er van alles te veel lijkt te zijn en waarin we niet een enkelvoudige oorzaak van overladenheid kunnen aanwijzen. Vooralsnog doen we er dan ook goed aan overladenheid als systeemkenmerk te zien. Het is een probleem waaraan door veel actoren onbedoeld wordt bijgedragen.

De mogelijke verklaringen voor overladenheid (Goodlad, Klein & Tye, 1979; Van den Akker, 1988; Kuiper, 1993; Boersma & Looy, 1997) vatten we samen in tabel 2.

Tabel 2: Mogelijke oorzaken van overladenheid, geordend naar verschijningsvormen van het curriculum.

| <i>Verschijningsvormen</i>                     | <i>Oorzaken</i>  |
|--|--|
| <i>Het geschreven curriculum (macroniveau)</i> | 1. <i>Het aantal vakken is te groot.</i>   |
|  | 2. <i>De minimumtabel omvat te weinig uren.</i>  |
|  | 3. <i>Het aantal kerndoelen en algemene vaardigheden is te groot.</i>  |
|  | 4. <i>Vakleerplannen hebben onvoldoende structuur en richten zich te veel op breedte en te weinig op diepte.</i>                           |
|  | 5. <i>Een minimale interpretatie van de kerndoelen en algemene vaardigheden ontbreekt.</i>   |
| <i>Het geschreven curriculum (microniveau)</i> | 6. <i>Auteurs van methoden beperken zich onvoldoende tot een minimale interpretatie van kerndoelen.</i>                                    |
|  | 7. <i>Methoden hebben onvoldoende structuur en richten zich te veel op breedte en te weinig op diepte.</i>                                 |
| <i>Het geïnterpreteerde curriculum</i>         | 8. <i>Leraren toetsen te weinig de dekking van de methode met de kerndoelen en algemene vaardigheden.</i>                                  |
| <i>Het voorbereide curriculum</i>              | 9. <i>Leraren bepalen in onvoldoende mate welke gedeelten van de methode aan de orde moeten komen en welke kunnen worden overgeslagen.</i> |
| <i>Het operationele curriculum</i>             | 10. <i>Leraren en leerlingen werken de methode niet volledig door.</i>   |

### 3. CONCLUSIES

Als we overladenheid als systeemkenmerk zien, ligt het voor de hand er van uit te gaan dat alle betrokken actoren, leraren, scholen, uitgeverijen, SLO en OCenW, een bijdrage aan de oplossing van het probleem moeten leveren.

We kunnen nu twee benaderingen kiezen. De eerste benadering is dat we eerst nader onderzoek naar overladenheid doen. Dergelijk onderzoek is belangrijk, omdat daarmee inzicht in een van de grote curriculumproblemen kan worden verworven. Van belang is in ieder geval dat we nauwkeuriger gegevens verzamelen over de tijd die per kerndoel wordt besteed.

De tweede benadering die we zouden kunnen kiezen is om te bezien of we, ons basierend op de gesuggereerde oorzaken, niet een aantal maatregelen kunnen voorstellen waarmee in ieder geval winst te boeken is, ook al is nog niet in alle opzichten duidelijk of daarmee het probleem daar volledig mee wordt verklaard. De effecten van de interventies die we voorstellen moeten dan uiteraard worden onderzocht. Een cruciaal punt in onze analyse van de oorzaken van overladenheid is het probleem dat leraren en uitgevers hebben om af te bakenen wat gezien de kerndoelen minimaal aan de orde zou moeten komen. Daarop aansluitend stellen we de volgende drie met elkaar samenhangende maatregelen voor.

1. Ontwikkel een minimale interpretatie van de kerndoelen, omdat daarmee de bandbreedte van mogelijke interpretaties kan worden ingeperkt. Voorwaarde daarbij is echter wel dat

- de minimale interpretatie breed gedragen wordt door het veld en niet opgevat kan worden als aanscherping van de regelgeving. Scholen die hun eigen interpretatie aan de kerndoelen willen geven hebben daartoe het volste recht. Desgewenst kunnen afzonderlijke interpretaties voor VMBO en HAVO/VWO worden ontwikkeld.
2. Om te kunnen bepalen of kerndoelen binnen de beschikbare tijd passen is het noodzakelijk een schatting te maken van de tijd die nodig is voor onderwijs dat op basis van de minimale interpretatie wordt gegeven. Pas dan is ook te overzien of voorstellen voor revisie van kerndoelen binnen de beschikbare tijd kunnen worden uitgevoerd. Daartoe moeten empirische gegevens over besteding van tijd per kerndoel worden verzameld.
  3. Op grond van minimale interpretaties van kerndoelen kunnen ook handleidingen bij bestaande methoden worden ontwikkeld. Leraren kunnen daarmee bewaken of hetgeen ze willen aanbieden in overeenstemming is met de eisen die vanuit de kerndoelen worden gesteld. Uitgevers zouden deze handleiding ook in de docentenhandleiding van hun methode kunnen opnemen.

De ontwikkeling van het studiehuis, waarbij de sturing meer dan voorheen bij de leerlingen wordt gelegd, is niet te rijmen met de stand van zaken in de basisvorming. Om een eigen identiteit te kunnen ontwikkelen, en dat geldt zowel voor scholen, leraren als leerlingen, is het noodzakelijk over vrije ruimte te beschikken. Het terugdringen van het beslag dat de kerndoelen op het programma leggen tot de beoogde 75 procent is dan ook noodzakelijk om de in gang gezette vernieuwingen te kunnen realiseren. Wij zijn dan ook van mening dat voor de verdere invoering van de basisvorming een reductie van de overladenheid noodzakelijk is. Om die reden is het gewenst om ten minste de voorgestelde maatregelen te nemen en om die maatregelen te ondersteunen met begeleidend onderzoek. Dat onderzoek moet zich enerzijds richten op de besteding van onderwijstijd en de problemen die zich daarbij voor leraren en leerlingen voordoen, en anderzijds op de effectiviteit van de voorgestelde maatregelen.

## LITERATUUR

- Akker, J.J.H. van den (1988). *Ontwerp en implementatie van natuuronderwijs*. Lisse: Swets & Zeitlinger.
- Boersma, K.Th. (1997). *Systeendenken en zelfsturing in het biologie-onderwijs*. Oratie Universiteit Utrecht.
- Boersma, K.Th. (2000). Het leerplan van de basisvorming als probleem. In: M. Peters-Sips, J. van der Linden & A. Wald (red.). *Verder werken aan de basis. Basisvorming bundelt krachten*. Verslag Conferentie Basisvorming, 18 november 1999 te Ede. Utrecht: Inspectie van het Onderwijs.
- Boersma, K.Th. & Looy, F. (1997). *Een praktijktheorie voor leerplanontwikkeling*. Enschede: SLO.
- Boersma, K.Th. & Smit, G. (1997). *De revisietrajecten voor kerndoelen basisonderwijs en basisvorming en examenprogramma's VBO/MAVO*. Een casestudy onderzoek. Enschede: SLO.
- Freedman, D. (1998). Science Education. How curriculum and instruction are evolving. *Curriculum Update*, fall 1998.
- Bruner, J.S. (1963). *The Process of Education*. New York: Vintage Books.
- Forrester, J.W. (1992). *System Dynamics and Learner-Centered-Learning in Kindergarten through 12<sup>th</sup> Grade Education*. [http://sysdyn.mit.edu/cle/lom\\_syseducation.html](http://sysdyn.mit.edu/cle/lom_syseducation.html)
- Goodlad, J.L., Klein, M.F. & Tye, K.A. (1979). *Curriculum inquiry: The study of curriculum practice*. New York: McGraw Hill.
- Inspectie van het Onderwijs (1999a). *Werk aan de basis. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (1999b). *Biologie in de basisvorming. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het Onderwijs.
- Kuiper, W.A.J.M. (1993). *Curriculumvernieuwing en lespraktijk. Een beschrijvend onderzoek op het terrein van de natuurwetenschappelijke vakken in het perspectief van de basisvorming*. (proefschrift Universiteit Twente)
- Kuiper, W.A.J.M. & Altink, A. (1990). *Biologie, natuurkunde, scheikunde en kennis der natuur: de feitelijke lespraktijk in beeld*. Enschede: OCTO.

- Kuiper, W.A.J.M., Bos, K.Tj. & Plomp, Tj. (1997). *Wiskunde en de natuurwetenschappelijke vakken in leerjaar 1 en 2 van het voortgezet onderwijs*. Nederlands aandeel in TIMSS populatie 2. Enschede: OCTO.
- Letschert, J.F.M. (1998). *Wieden in een geheime tuin. Een studie naar kerndoelen in het Nederlandse basis-onderwijs*. Enschede: SLO. (proefschrift Universiteit Utrecht)
- Lijnse, P.L. (1997). Vakdidactiek, het vergeten fundament van het studiehuis. *Tijdschrift voor didactiek van de Wiskunde en Natuurwetenschappen*, 14/1, 72-91.
- Plas, E. van der (2000). Het studiehuis - van alles teveel behalve geld. *Volkskrant Magazine 15-4- 2000*, 33, 19-26.
- Roebertsen, H. (1996). *Integratie en toepassing van biologische kennis. Ontwikkeling en onderzoek van een curriculum rond het thema 'lichaamsprocessen en vergifti'*. Utrecht: CDB Wetenschappelijke bibliotheek.
- Vollebregt, M. (1998). *A problem-posing approach to teaching an initial particle model*. Utrecht: CDB Wetenschappelijke bibliotheek.
- Wetenschappelijke Raad voor het Regeringsbeleid (1986). *Basisvorming in het onderwijs*. 's-Gravenhage: Staatsuitgeverij.



# Vorm aan de basis, een terugblik op de invoering en evaluatie van de basisvorming

Mia Verhagen, *Open Schoolgemeenschap Bijlmer (OSB), Amsterdam*  
Anke Hartevelt, *Petrus Canisius College, Alkmaar*

## 1. VOORAF

Op verzoek van de gastredactie van dit themanummer over de evaluatie basisvorming kijken wij in dit artikel terug op hoe deze evaluatie in de scholen heeft uitgepakt. Wij beschrijven eerst de ontwikkeling van de basisvorming tot op dit moment. Vervolgens bespreken we enkele knelpunten op het gebied van schoolontwikkeling en onderwijsaanbod die scholen zijn tegengekomen bij de invoering en bij de evaluatie van de basisvorming.

## 2. WAAR GING HET OM IN DE BASISVORMING?

Het was de beoogde onderwijsvernieuwing die de samenleving zou moeten verbeteren en de ongelijkheid in startposities zou kunnen wegwerken. *"Een algehele verhoging van het peil van het jeugdonderwijs, het uitstellen van het moment van de studie- en beroepskeuze en modernisering en gedeeltelijk harmoniseren van het onderwijsprogramma"*, zo luiden de drie hoofddoelen in de Wet op de Basisvorming in 1992 (zie ook het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer).

Het idee van een verlengde en gezamenlijke brugperiode beschreef minister van Kernen al in 1975 in zijn Contourennota en hij stimuleerde vervolgens de middenschoolexperimenten, waar wij, schrijvers van dit artikel, beiden jarenlang werkzaam zijn geweest. In een interview in de bundel 'Inspiratie' lichtte Van Kemenade de doelstellingen van het middenschoolonderwijs nog eens toe: *"Drie aspecten van onderwijsvernieuwing waren belangrijk: allereerst noem ik de sociale motivering, dat wil zeggen het streven om alle kinderen, ongeacht herkomst en geslacht, gelijke kansen te geven op deelname aan het onderwijs en ontplooiing van hun talenten. Ten tweede ging het om een culturele motivering, gericht op verbreding van het vormingsaanbod, op doorbreking van traditioneel negentiende-eeuwse hiërarchie van vormingsidealen en op voorkoming van een vroegtijdige voorsortering van kinderen naar onderscheiden en in schooltypen neergeslagen culturele modellen. Het derde aspect is dat van een pedagogisch-didactische motivering, erop gericht om de leerling in zijn of haar verschillende ontwikkelingsfasen meer tot zijn recht te laten komen en dit te laten gebeuren in de grote verscheidenheid van ontwikkelingsmogelijkheden van leerlingen."* (Van Kemenade, 1997).

Er is na 1975 jarenlang gediscussieerd over de voorgestelde vernieuwingen, het werd een politiek heet hangijzer. Over het resultaat van deze discussie: het rapport van de WRR over de basisvorming werd verschillend geoordeeld (Wetenschappelijke Raad voor het Regeringsbeleid, 1986). Sommigen vonden de middenschoolgedachte er te sterk in terug en anderen vonden het, gezien de uitzonderingen, zoals de regeling voor het geven van de klassieke talen op de gymnasia of geen verplichting van een tweede moderne vreemde taal in het VBO, een verbaal compromis in de oude strijd rond de middenschool.

Invoering van de basisvorming werd echter een feit en vanaf 1993 moesten scholen voor voortgezet onderwijs hun brugperiode verlengen, techniek en verzorging als nieuw vak invoeren en vooral op een nieuwe manier met leren bezig zijn. Fusies moesten er vervolgens voor zorgen dat de basisvorming kon gedijen in de brede scholengemeenschappen want die zouden pas echt het idee van uitstel van studie- en beroepskeuze uit kunnen voeren.

### 3. EVALUATIE

De evaluatie werd bij wet vastgelegd: na vijf jaar zou de inspectie een evaluatie uitvoeren en er zou gekeken worden of en hoe de veranderingen in de eerste fase van het voortgezet onderwijs zich hadden voltrokken.

120 Scholen, waaronder onze beide scholen waar we respectievelijk werkzaam waren/zijn, kregen in het voorjaar van 1998 een visitatie van een team van inspecteurs, er werden lessen bijgewoond; gesprekken gevoerd; er werd schriftelijk onderzoek gedaan en het Cito nam proeftoetsen af. Er werd gekeken naar de invoering van de basisvorming, maar ook naar de kwaliteit en het rendement van het onderwijs. Scholen kregen vervolgens uitgebreid mondeling en schriftelijk verslag van de bevindingen en tevens een reeks aanbevelingen om aan de slag te gaan.

Een van de conclusies van de Inspectie in het evaluatierapport 'Werk aan de basis' was dat de verschillen tussen scholen groot zijn als het gaat om de kernkwaliteit, de kwaliteit van de lessen. Bij zo'n 20 procent van de scholen is de kwaliteit van de lessen veelal onvoldoende en is de uitval van voornamelijk allochtone leerlingen groot. Aan onderwijsvernieuwing zijn deze scholen niet of nauwelijks toegekomen. De meeste scholen (60%) hebben een middelmatige kwaliteit en de inhoudelijke onderwijsvernieuwing in de basisvorming is op de meeste scholen onvoldoende van de grond gekomen.

Naast schoolleiders en docenten kreeg ook de overheid een aantal aanbevelingen waarop door Staatssecretaris Adelmund op 15 september 1999 met een brief werd gereageerd bij de aanbieding van het evaluatierapport van de Inspectie aan de Tweede Kamer. Deze reactie dient als richtsnoer bij de noodzakelijke discussie over de te maken keuzes voor de toekomstige ontwikkeling van de basisvorming.

Behalve de Inspectie hebben ook onderzoeksinstituten onder auspiciën van NWO rapporten gepubliceerd over, een weliswaar veel beperkter, onderzoek waarin de vergelijking wordt gemaakt tussen resultaten en schoollopen van leerlingen die in 1989 en in 1993 in de brugklas begonnen. De onderzoekers hebben daardoor de effecten van de basisvorming ook kunnen bekijken en de balans van de basisvorming kunnen opmaken. Zij hebben positieve effecten geconstateerd zoals de afname van het zittenblijven, er is meer opstroom en er is meer persoonlijke begeleiding. Negatieve effecten die genoemd worden zijn de enorme verschillen in het aanbod per schooltype, geen afname van de sociale ongelijkheid, de vervroeging van de selectie door dakpanklassen en nauwelijks toepassingsgericht onderwijs.

Uit de evaluatie van de basisvorming komen terechte bevindingen en aanbevelingen. Inmiddels zijn mogelijkheden en wensen voor verbetering van de basisvorming op korte en lange termijn bijeengebracht in de bundel 'Verder werken aan de basis' (Peters-Sips, Van der Linden en Wald, 2000). De Inspectie, de Vereniging voor Onderwijs Research en de Vereniging voor het management in het Voortgezet Onderwijs droegen bouwstenen aan voor dit vervolg.

Toch valt het terugkijkend naar deze invoeringsperiode, het scholenveld niet zomaar te verwijten dat de implementatie van de basisvorming onvoldoende ver gevorderd is gezien de knelpunten die de scholen zoal zijn tegengekomen bij die invoering op het gebied van schoolontwikkeling en onderwijsaanbod.

#### 4. SCHOOLONTWIKKELING

Het heeft vanaf het begin ontbroken aan voldoende ontwikkeltijd en -ruimte, dus aan geld. De regelgeving was en is niet faciliterend genoeg voor een substantiële verkleining van de lestaak, waardoor ontwikkeltijd voor alle docenten vrijgemaakt had kunnen worden. Schoolleiders hadden daarnaast te weinig tijd voor uitzetten van onderwijsinhoudelijk beleid met name door de fusies. Het ontstaan van grote scholengemeenschappen is door de overheid gestimuleerd en het aantal zelfstandige scholen is tussen 1992 en 1997 gedaald van 1454 naar 691. Veel tijd en aandacht werd besteed aan het structureren van organisaties van de gefuseerde school. Een neveneffect van de fusies was het gebrek aan continuïteit in de schoolleiding en docententeams, hetgeen ook de invoering van vernieuwing vertraagde.

Daar komt bij dat schoolleiders behalve met fusies ook te maken hadden met een aantal andere grote operaties zoals de invoering van de lump sum, decentralisatie van huisvesting,erschikking en invoering ICT. Schoolleiders hadden en hebben meer tijd en geld nodig om al deze veranderingen te kunnen implementeren.

Een ander belangrijk knelpunt was dat alleen de onderbouw zich moest vernieuwen terwijl de bovenbouw nog een aantal jaren met de oude examens moest werken. De onzekerheid van docenten over de juiste effecten van de basisvorming met het oog op de oude examens veroorzaakte mede een stroeve invoering van de basisvorming.

Nu de Tweede fase en het VMBO zijn ingevoerd, is het logisch dat er in de hele school een integraal vernieuwingsproces op gang komt. De noodzaak om het onderwijs te veranderen dient zich nu vanuit de nieuwe innovaties sterker aan. Met terugwerkende kracht ziet men duidelijker wat er in de periode van de start van de basisvorming had moeten gebeuren.

De invoering Tweede fase nodigt uit tot het veranderen van het 'starre denken' in lessen van 50 minuten en een verdeling van vaklessen over een heel jaar. Door het durven loslaten van deze oude structuren zal innovatie een extra impuls krijgen.

De docenten waren weliswaar vakinhoudelijk redelijk voorbereid, maar ze waren slecht voorbereid op een verandering van hun lesaanpak. Velen van hen geven nog altijd de voorkeur aan frontaal lesgeven boven kinderen echt te leren en te leren onderzoeken. Ze hebben moeite met het creëren van meer betrokkenheid, meer activiteit en zelfwerkzaamheid en meer uitdaging voor elk kind.

Docenten weten dat er een andere manier van lesgeven wordt gevraagd, maar zij weten daar vaak geen inhoud aan te geven. Veranderen is een proces dat een flexibele en reflectieve houding vraagt van alle betrokkenen, een proces waarbij iedereen serieus genomen moet worden.

Onderwijsvernieuwing noodzaakt een omslag in denken en handelen bij docenten. Daarvoor is tijd en geld nodig. Docenten ervaren een te hoge werkdruk en beschikken over weinig ontwikkeltijd naast hun lestaak. Ze hebben niet voldoende geleerd om te gaan met kinderen in onderwijskansenbeleid, te differentiëren in hun lessen, bijvoorbeeld via ICT of aan te sluiten bij verschillende leerstijlen. Dat maakt dat docenten het werken aan onderwijsvernieuwing ervaren als iets wat erbij komt, wat extra moet gebeuren, in plaats van als een integraal onderdeel van hun werk.

Scholen werken nogal eens met volle klassen, waardoor het aansluiten bij de verschillen tussen kinderen vaak niet lukt. Bovendien volgen veel docenten, vooral in tijden van onzekerheid, nogal strikt hun methodes. Deze methodes bieden docenten weinig mogelijkheden tot differentiatie en het negatieve effect op de motivatie is daardoor groot.

De basisvorming vraagt soms ook een cultuuromslag van de leerling, ze komen nu eenmaal van zeer verschillende basisscholen en zelfstandig werken heeft niet elke leerling op de basisschool in voldoende mate geleerd.

Elk zichzelf respecterende school heeft, al is het soms in de marge, vormen van na- en bijscholing georganiseerd in deze invoeringsjaren. Toch heeft dit niet opgeleverd dat er in het leerproces veel is veranderd. Door schoolleiding opgelegde en georganiseerde nascholing

werkt ook niet altijd zomaar en is niet altijd adequaat genoeg voor wat de individuele docent moet leren.

De praktijk heeft inmiddels uitgewezen dat voor de zittende docenten scholing op maat, in de school met een groep docenten, in de praktijksituatie, intercollegiaal, met intervisie en begeleidingstrajecten of uitwisseling in netwerken beter werkt dan individuele cursussen buiten de school. Pas onder zeer specifieke voorwaarden komt daadwerkelijk leren van docenten tot stand, leren dat blokkades opheft en doorwerkt in de dagelijkse onderwijspraktijk. Docenten zullen een proces van levenslang leren moeten accepteren en deze voortgezette scholing voor zittende docenten kan allerlei vormen aannemen.

Lerarenopleidingen bereiden hun studenten niet goed genoeg voor op wat er in de scholen gevraagd wordt. Er zijn verschillende initiatieven voor een nauwere samenwerking van lerarenopleidingen en scholen waarbij scholen een grotere rol zullen krijgen dan voorheen in de opleiding van docenten. Als scholen zelf meer opleidingsinstituut zouden worden dan biedt dat in de toekomst mogelijkheden voor taakdifferentiatie van docenten.

In het kader van het (dreigend) lerarentekort is het heel belangrijk dat de gevraagde competenties van de docent duidelijk worden, vakinhoudelijk, maar vooral ook op pedagogisch en didactisch gebied.

## 5. ONDERWIJSAANBOD

Op het gebied van onderwijsaanbod kijken we naar vakken, kerndoelen, toetsen en naar de tweedeling die in het aanbod is ontstaan.

### *Vakken*

Het programma van de basisvorming in 15 vakken is overladen; het kwam vaak boven op het bestaande aanbod en niet in plaats van of geïntegreerd met de oude stof van de eerste fase voortgezet onderwijs. Zeker voor VBO- en MAVO-leerlingen is het programma te overladen, mede door de beperkte vierjarige cursusduur. Voor VBO- en IVBO-leerlingen is het programma ook inhoudelijk te zwaar.

Vakkenintegratie werd echter door de wet ontmoedigd, zeker ook door de bevoegdhedenregeling alhoewel daar meestal behoorlijk soepel mee wordt omgegaan. Onbevoegd is bepaald niet vanzelfsprekend onbekwaam. De lobby van vakdocentenorganisaties voor aparte vakken en vooral de inhoud van die aparte vakken is steeds weer groot en daar blijkt de politiek gevoelig voor te zijn. De starre indeling van vakken en dus ook van de methodes zou moeten worden doorbroken door het maken van leergebieden of vakkenclusters. Vakkenclustering werkt niet alleen overlap en versnippering tegen, maar minder docenten voor de klas bevordert ook de pedagogische binding. Een docent kan meer vakken geven en een klein kernteam kan verantwoordelijkheid nemen voor een klas. Een docent leert zijn klassen en leerlingen beter kennen en kan daardoor beter rekening houden met de verschillen tussen zijn leerlingen.

Samenhang in het aanbod en inspelen op verschillende leerstijlen is belangrijk voor de motivatie van leerlingen. Helaas bepalen methodes het aanbod en de toetsing en boeken worden leidraad van de docent. Aansluiten bij verschillen is daardoor veel moeilijker en docenten kunnen hun didactiek niet meer flexibel maken. Methodes blokkeren de vernieuwing binnen de klas, ze zijn per vak en vaak ook nog per schooltype gemaakt en ze propageren te zeer een zelfde manier van leren. Alle kinderen zijn echter gebaat bij een afwisselende, onderzoekende, aanschouwelijke, toepassingsgerichte leer- en doceerstijl. Het gaat erom dat er wordt uitgegaan van competenties van kinderen, van hun mogelijkheden. In het VMBO en de Tweede Fase zien we reeds hetzelfde probleem van versnippering, overladenheid en te weinig samenhang.



*Kerdoelen*

De kerndoelen blijken moeilijk te formuleren, diverse commissies hebben er zich mee bezig gehouden en toen ze de eerste kerndoelen er waren bleken ze onwerkbaar. Ze zijn vrij snel na de start van de basisvorming aangepast en opnieuw vastgesteld (Kerndoelen 1998-2003). Het zijn er echter nog steeds erg veel; er is sprake van overlap in kerndoelen en in de toepassing in de programma's. In de discussie over de kerndoelen en de formulering speelde ook nog het aspect van educaties en facetten; dat zijn nog eens extra leerinhouden die nu vakoverstijgende thema's worden genoemd. Het kan de docent niet kwalijk worden genomen dat hij door de bomen het bos niet meer ziet en vertrouwt op de uitgevers om deze (algemene en op het vakgerichte) kerndoelen in de methodes te verwerken.

*Toetsen*

Er is eveneens grote verwarring geweest rond de inhoud en status van de toetsen. De basisvorming was niet ver genoeg doorgedacht leek wel, want er was vanaf het begin onduidelijkheid. De basisvorming zou met een landelijke toets van de stof van de verplichte kerndoelen worden afgesloten. Van deze toets werd na 1 jaar afgestapt; hij bleek voor de hoogste niveaus te makkelijk en voor de laagste te moeilijk. Het Cito ontwikkelde daarop afsluitende toetsen voor de verschillende schooltypen.

Docenten vinden nog steeds dat de inhoud en de zwaarte van de toetsen geen recht doen aan de lessen, dus scholen zullen veel meer schooleigen toetsen moeten samenstellen, toetsen met een diagnostische functie. Ze zouden daarbij gebruik moeten kunnen maken van een toetsenbank. Afhankelijk van de leerlingen kan dan op verschillende niveaus getoetst worden en de toetsen kunnen als een van de bereikte resultaten meetellen in het keuzeproces van de leerling.

*Tweedeling*

Tweedeling in de basisvorming is onvermijdelijk. De overheid zei vorming van brede scholengemeenschappen te stimuleren. Uiteindelijk hebben veel brede scholengemeenschappen zich zo georganiseerd dat de tweedeling gehandhaafd bleef. Er zijn veel brede scholengemeenschappen met aparte locaties voor het VMBO en de HAVO/VWO.

De meeste scholen hebben homogene brugklassen of een dakpanconstructie. In het eerste leerjaar is ongeveer een derde van de leerlingen categoriaal ingedeeld en zit ruim de helft van de leerlingen in een dakpanklas. In het tweede leerjaar is bijna 60 procent van de leerlingen categoriaal ingedeeld en zit bijna 40 procent in een dakpanklas (Inspectie van het Onderwijs, 1999, pagina 122). Het gevolg is dat er weinig uitstel van studie- en beroepskeuze plaatsvindt. Hierdoor zijn de gezamenlijke doelen flink onder druk komen staan. De maatschappelijke druk op de scholen is echter te groot en scholen maken soms nu eenmaal keuzes uit pure noodzaak. Na eenmaal in een leerroute terecht te zijn gekomen nog doorstromen naar een andere leerroute, is dan ook moeilijk, op een paar scholen na met heterogene brugperiodes. Het is onwenselijk dat kinderen geen optimale kans krijgen hun mogelijkheden te ontwikkelen; met name geldt dit voor allochtone leerlingen. Zij hebben extra veel tijd nodig voor het leren van Nederlands.

Er zijn tegen de oorspronkelijke bedoeling in twee typen basisvorming ontstaan. In VBO en MAVO worden veel minder uren besteed aan het basispakket van 15 vakken dan in HAVO en VWO en de leerlingen van VBO en MAVO ronden de basisvorming al vaak na het tweede leerjaar af vanwege de beperkte verblijfsduur en de examen voorbereiding (zie het artikel 'Kwaliteit van het vakspecifieke aanbod' elders in dit themanummer).

## 6. EVALUATIE VAN DE EVALUATIE

Over de evaluatie door de Inspectie willen we slechts met betrekking tot drie aspecten een opmerking maken; het betreft de rol van de scholen zelf; de periode waarna geëvalueerd werd en het gebruikte toetsingskader.



Ten eerste is de rol van de scholen zelf in de evaluatie te beperkt gebleven. De scholen hebben materiaal aangeleverd en een lesrooster overhandigd zodat de inspecteurs de lessen konden volgen en er zijn gesprekken gevoerd over de wijze waarop men de implementatie van de basisvorming ter hand had genomen. De evaluatie was teveel van bovenaf georganiseerd en scholen hebben zich te afwachtend opgesteld. Sommige schoolleiders hebben ook nog de uitslag van deze evaluatie gebruikt als legitimering van hun handelen; de vraag is of dat een juiste wijze van handelen is en of daarmee het gewenste effect bereikt wordt. Het gaat er bij onderwijsvernieuwing toch om de mensen zelf enthousiast te maken en te motiveren daadwerkelijk aan het nieuwe leren vorm te geven. Schoolleiders en docenten dragen samen verantwoordelijkheid voor wat er in de school gebeurt, maar dan moet die verantwoordelijkheid wel geboden worden.

Ten tweede was de evaluatie na 5 jaar te vroeg; een dergelijk grote vernieuwingsoperatie kost misschien wel het dubbele in tijd (de WRR sprak over tien jaar, zie het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer). Bovendien was door de 'oude' bovenbouw de invoering van de basisvorming in het begin gehandicapt.

Ten derde deed de evaluatie niet in alle scholen recht aan wat er in de scholen gebeurde en bereikt werd. Het toetsingskader zorgde ervoor dat bijzondere onderwijspraktijken niet of niet voldoende belicht konden worden omdat de vergelijkbaarheid van de gegevens van scholen dan in het gedrang kwam, terwijl evaluatie van de kwaliteit van een organisatie rekening dient te houden met de doelen die de organisatie wil bereiken. Een beoordeling is altijd relatief aan de hand van een tevoren opgestelde norm. Is in de evaluatie van de basisvorming de juiste norm gehanteerd? Als je het Nederlandse onderwijs vergelijkt met het onderwijs in andere Europese landen dan scoort Nederland heel goed (OECD, 1998). De vraag blijft ook of de Inspectie niet teveel heeft gekeken naar de kwaliteit van wat er in de lessen gebeurde, terwijl het gaat om een veel complexer proces; ook de organisatieveranderingen van de school, de ontwikkelingsfase van de school op die plaats, in die context, in die situatie moeten betrokken worden in een evaluatie van onderwijsvernieuwing.

Het is de vraag of de innovaties van de Tweede Fase en het VMBO op zo'n termijn, op zo'n schaal, en op zo'n wijze geëvalueerd zullen moeten worden. Scholen hebben de opdracht op zeer korte termijn een eigen kwaliteitszorgsysteem te ontwikkelen waardoor de overheid op doelstellingen kan controleren en niet op uitvoering, maar dat is een ander verhaal.

## LITERATUUR

- Inspectie van het Onderwijs. (1999). *Werk aan de basis. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het Onderwijs.
- Kemenade, J.A. van (1997). Onderwijs dat werkelijk vrijmaakt. In: *Inspiratie*. Amsterdam: Open Schoolgemeenschap Bijlmer.
- OECD (1998). *Education at a Glance*. Parijs: OECD.
- Peters-Sips M., Linden, van der J. & Wald A. Inspectie van het Onderwijs. (2000). *Verder werken aan de basis. Basisvorming bundelt krachten*. Utrecht: Inspectie van het Onderwijs.
- Wetenschappelijke Raad voor het Regeringsbeleid. 1986. *Basisvorming in het onderwijs*. 's-Gravenhage: Staatsuitgeverij.

# Evaluatierapport Basisvorming uniek document

Marleen Barth, lid Tweede Kamer der Staten Generaal voor de PvdA

## 1. INLEIDING

'Basisvorming mislukt', kopte menig landelijk dagblad met een verbazingwekkende gretigheid, na het verschijnen van het evaluatierapport van de Inspectie, najaar 1999. Wie de moeite neemt om het imposante resultaat van het werk van de Inspectie in zijn geheel tot zich te nemen, kan niet anders dan vaststellen dat dit een onjuiste conclusie is. Want met een blik, gevormd door een periode waarin aan de stroom financiële meevallers op de rijksbegroting geen einde lijkt te komen, is het eigenlijk een wonder dat er van de basisvorming de afgelopen jaren nog zo *veel* terecht is gekomen. Wat waren de ambities hoog, en wat was er weinig geld. Fuseren én nieuwe vakken én nieuwe kerndoelen én een nieuwe didactische aanpak, en dat praktisch budgettair neutraal. Nogmaals, de financieel-economische omstandigheden van de rijksoverheid waren toen heel anders dan nu. Maar dat neemt niet weg dat het heel veel loyale inzet en inspiratie van leraren en schoolleiders gekost moet hebben om uit te komen bij het onderwijspeil dat de momentopname van de Inspectie weergeeft.

## 2. POSITIEVE BEVINDINGEN INSPECTIE

De evaluatie van de basisvorming door de Inspectie is een uniek document. Nooit eerder is de kwaliteit van een onderwijssector zo minutieus tegen het licht gehouden. De stukken bieden een schat aan informatie, en interessante aanknopingspunten voor de toekomst. Daarover later meer. Eerst een samenvatting van het goede nieuws dat de Inspectie te melden had. Bijvoorbeeld dat de belangrijkste doelstelling van de wet, "*verhoging van het peil van het jeugdonderwijs*", is gehaald. De prestaties liggen nu hoger dan voor de invoering van de basisvorming. Toegegeven, de oorzaak daarvan is niet alleen gelegen in de onderwijsvernieuwing; zo eendimensionaal werkt het menselijk leervermogen nu eenmaal niet. Maar wel eerder reden voor optimisme, dan cynisme.

Een tweede, positief effect hangt wel samen met de invoering van de basisvorming en de bijbehorende brede scholengemeenschappen. De doorstroming van leerlingen is verbeterd; leerlingen belanden soepeler op de plek waar ze thuis horen, hetzij door opstroom, hetzij door afstroom. De eerste is groter dan de laatste, en nog nooit zaten er zoveel leerlingen op HAVO en VWO. Leerlingen verliezen minder tijd onderweg, en kunnen overstappen zonder dat hun hele sociale omgeving op zijn kop wordt gezet.

Een derde resultaat, nog steeds volgens de Inspectie: het beoogde uitstel van studie- en beroepskeuze vindt plaats. Dat hangt nauw samen met de toegenomen mobiliteit van leerlingen in het voortgezet onderwijs. Zoals bekend, is uitstel van studie- en beroepskeuze vooral voor leerlingen uit kansarme milieus van belang. Bij hen duurt het vaak langer voor hun talenten aan de oppervlakte komen, en zij hebben daarom alle belang bij voldoende tijd om zich te kunnen ontplooiën voor de bijl van de verwijzing valt.

Ten slotte: de gedroomde modernisering en harmonisering van het lesprogramma. De evaluatie laat zien dat verreweg de meeste scholen met de nieuwe vakken en vakinhouden van de basisvorming aan de slag zijn gegaan. Wel tekent de Inspectie hierbij aan dat de verschillen in aanbod van kerndoelen tussen en binnen scholen groot zijn.

### 3. KNELPUNTEN

Uiteraard zijn de resultaten van de evaluatie niet alleen maar positief. Zo stelt de Inspectie onomwonden vast dat het kerndoelenprogramma overladen en versnipperd is. Voor veel leerlingen is vijftien vakken (op VBO en Gymnasium zelfs nog meer) te veel van het goede. Veel leraren krijgen alle kerndoelen niet weggezet in de tijd die daar voor staat. Tot op heden onderbelicht gebleven, maar wel de aandacht waard, is de rol van de educatieve uitgeverijen daarin. Lang niet alle naar aanleiding van de basisvorming ontwikkelde methoden dekken de kerndoelen op een adequate manier. Veel methoden bieden een breder aanbod van lesstof dan de kerndoelen. Dat maakt in de klas selectie in het aanbod door de leraar noodzakelijk. Voeg daarbij de klacht vanuit het (VBO-)veld dat de nieuwe methoden te academisch en theoretisch zijn voor meer praktisch ingestelde leerlingen, en het wordt alleszins de moeite waard het werk van educatieve uitgeverijen nog eens nader onder de loep te nemen.

Leraren nemen in de alledaagse praktijk te weinig afstand van die methoden, blijkt uit de evaluatie. Hier en daar is zelfs gesproken over 'methode-verslaafdheid' van leraren. Een inspecteur vertrouwde me toe dat de collega's uit het buitenland die bij de evaluatie betrokken waren, zich verbaasden over die strakke uitvoering van methoden in de klas. Daarnaast is de mate waarin leraren in staat zijn om in te spelen op verschillen tussen leerlingen onder de maat, zo constateert de Inspectie.

Ander knelpunt is dat de beoogde tempodifferentiatie niet van de grond komt. Was tijdens het debat over de wet nog het uitgangspunt: 'sneller waar het kan, langzamer waar het nodig is', in de praktijk gebeurt precies het omgekeerde. Een VWO-leerling krijgt drie jaar voor de basisvorming, tot het studiehuis aanvangt. VBO-leerlingen worden geacht het programma in twee jaar te verwerken, omdat daarna de pakketkeuze voor het eindexamen wacht. Die druk is volgens de Inspectie 'een structureel knelpunt' in het VBO.

Hoewel uitstel van studie- en beroepskeuze van de grond komt in bredere scholengemeenschappen, lijkt de trend zich te verwijderen van de brede brugklas. De Inspectie treft meer dakpanklassen aan, en minder breed samengestelde groepen dan voor de invoering van de basisvorming. Dat is jammer, omdat onderzoek steeds opnieuw toont dat het rendement van het onderwijs stijgt in een gemengde klas (zie bijvoorbeeld Westerbeek, 1999). Kinderen leren immers niet alleen van hun leraar, computer of uit hun boeken, maar ook veel van elkaar. Als kansarme kinderen de gelegenheid krijgen zich op te trekken aan kansrijkere leeftijdsgenoten, heeft dat een gunstig effect op de gemiddelde onderwijsprestaties.

Laatste kritiekpunt van de Inspectie: de toetsing van de basisvorming werkt niet.

### 4. BASISVORMING VERDER UITBOUWEN

Vanuit sommige hoeken klonk, naar aanleiding van de in het Inspectierapport geconstateerde problemen, het pleidooi om de basisvorming maar weer te verlaten, omdat de vernieuwing toch niet zou werken. Er zijn echter goede redenen om aan dat verzoek geen gehoor te geven. Zo doet het geen recht aan de vele inspanningen die scholen hebben gepleegd om de basisvorming handen en voeten te geven, noch aan de genuanceerde conclusies en aanbevelingen van de Inspectie. En hoeveel scholen zouden er, met de handen vol aan de huidige herziening van de bovenbouw, zitten wachten op een overheid die in de onderbouw de koers weer radicaal omgooit?

Dat zou ontwikkelingen in de eerste en de tweede fase van het voortgezet onderwijs bovendien onaanvaardbaar ver uit elkaar trekken. Immers: de didactische vernieuwing van studien en leerhuis komen met veel meer enthousiasme van de grond dan indertijd de basisvorming (het gesternte waaronder dat gebeurt is, alle groei pijn in de bovenbouw ten spijt, dan ook aanmerkelijk gunstiger). De Inspectie koestert de verwachting dat onder invloed van studie- en leerhuis vooral de didactische vernieuwing van de basisvorming met hernieuwd elan zal wor-

den opgepakt. Het streven voor de komende jaren moet daarom veel meer zijn om beide fasen naadloos met elkaar te integreren. Al was het maar om tegemoet te komen aan de verzuchting van een ouder, vorig jaar in een landelijke ochtendkrant: "Op de basisschool leerde mijn zoon zelfstandig werken, in de onderbouw van het voortgezet onderwijs leerden ze het hem weer af, en nu hij in het studiehuis zit, moet hij het weer aanleren".

Laatste, maar zeker niet onbelangrijkste argument om de basisvorming uit te bouwen en niet af te breken, is dat de doelstellingen van de wet nog altijd springlevend zijn. Sterker nog, wie het advies van de Wetenschappelijke Raad voor het Regeringsbeleid (WRR, 1986) dat aan de wet ten grondslag lag er nog eens bijpakt, kan niet anders vaststellen dan dat de Raad haar tijd vooruit was. Nederland bevindt zich nu midden in een omslag van een hiërarchisch georganiseerde, industriële samenleving naar een kenniseconomie, waarin scholing van en dienstverlening aan individuen een centrale plaats innemen. De investeringen in menselijk kapitaal die dat vergt, sluit aan bij de motivatie van de WRR om tot basisvorming te komen: "*een verhoging van het onderwijspeil voor alle leerlingen en een deelname van alle (toekomstige) leden van de samenleving aan het culturele erfgoed*", "*de cultuurpolitieke noodzaak van co-existentie in een pluriforme samenleving*", "*de eisen van economische en technische modernisering*", "*vermindering van ongelijkheid van kansen in het onderwijs*" en "*verbetering en herwaardering van het beroepsonderwijs*".

Dat kinderen van verschillende achtergrond én met een verschillende toekomst leren samenwerken en profiteren van elkaars talenten, en tegelijkertijd leren zelfstandig hun weg te vinden in overstelpende hoeveelheden informatie en een steeds individueler vormgegeven normen- en waardenpatroon, de komst van een nieuwe economie maakt het allemaal actueler dan ooit.

## 5. ADVIEZEN

Grote uitdaging voor de politiek moet daarom zijn hoe de omgeving van de basisvorming zo vruchtbaar kan worden gemaakt dat ze tot volle wasdom kan komen. De conclusies en aanbevelingen van de Inspectie leveren een niet te onderschatten bijdrage aan die discussie. Bij deze gelegenheid beperk ik me tot de aanbevelingen in de richting van de overheid.

### *Stel de kerndoelen bij*

Of, kort samengevat: maak een einde aan de overladenheid en versnippering van het programma. Het is een oproep die de politiek ter harte moet nemen. Maar met het simpelweg schrappen in kerndoelen doen we dit probleem geen recht. Opvallend is bijvoorbeeld de grote verschillen in de mate waarin kerndoelen haalbaar blijken tussen verschillende vakken. De conclusie dat de kerndoelen beter gehaald worden naarmate het schooltype hoger is, stemt overeen met het feit dat leerlingen meer tijd hebben voor het programma naarmate het schooltype hoger is. Interessant is wel dat de drie nieuwe vakken van de basisvorming -techniek, verzorging en informatiekunde- in positieve zin afwijken van deze trend. In deze vakken ligt het percentage behandelde kerndoelen in het VBO veel hoger dan bij andere vakken. Dat kan te maken hebben met het feit dat juist deze vakken voortkomen uit het meer praktische vakkenaanbod (verzorging en techniek waren op het VBO reeds bestaande vakken), maar even plausibel is de veronderstelling dat de kerndoelen behoorlijk te doen zijn omdat het nieuwe vakken zijn. Reeds bestaande vakken kregen door de basisvorming minder uren, dus moest er vakinhoud uit. En schrappen is altijd pijnlijker dan met een schone lei beginnen... Nader onderzoek lijkt hier op zijn plaats.

Ook een betere integratie van vakken kan de overladenheid reduceren. Niemand schiet er bijvoorbeeld iets mee op als in verschillende vakken dezelfde kerndoelen worden aangeboden. Clustering is mogelijk, en verdient het om stevig gestimuleerd te worden. Experimenten op scholen bewijzen dat met name (I)VBO-leerlingen daar baat bij hebben. Maar ook voor



andere scholieren is het volgen van zeven, of negen, vakken aanmerkelijk overzichtelijker dan 15, of zelfs meer.

Een kritische blik op door educatieve uitgeverijen geproduceerde methoden kan ook geen kwaad. Voor het basisonderwijs zijn inmiddels diverse methodengidsen ontwikkeld, en dat verdient zeker navolging in het voortgezet onderwijs. Als methoden veel meer stof aanbieden dan de kerndoelen behelzen, ontstaat overladenheid van het programma zonder dat een van de bij beleidsontwikkeling betrokken actoren (ministerie, volksvertegenwoordiging, onderwijsveld) daar invloed op kan uitoefenen.

Dit alles neemt niet weg dat het een goede zaak zou zijn om als politiek de kerndoelen nog eens door te lichten. Een samenhangend, haalbaar programma dat het mogelijk maakt bij elke leerling een brede basis te leggen, moet het doel zijn. De Onderwijsraad heeft met haar advies over leerstandaarden een solide basis gelegd voor zo'n herziening. Een vervanging van kerndoelen en tussendoelen van basisonderwijs en basisvorming door leerstandaarden (stof die redelijkerwijs door 90 procent van de leerlingen op een bepaalde leeftijd gehaald moet kunnen worden) levert een prima moment voor kabinet en Kamer het totale programma in die zin te actualiseren. Een beperking tot een programma van twee jaar, zoals de Inspectie suggereert, verdient zeker aandacht om het dilemma rond het VBO op te lossen. Maar ook het meer integreren van kerndoelen basisvorming met de eindexameneisen VMBO lijkt hiertoe een begaanbare route.

#### *Stimuleer de pedagogisch-didactische variatie in scholen*

Ook hier kan op tal van manieren uitwerking aan worden gegeven. Bijvoorbeeld (een aspect dat in het Inspectierapport ontbreekt) door de lerarenopleidingen te verbeteren en te vernieuwen. Het belang van gedifferentieerd onderwijs zou bij de jonge, aanstaande leraar vanaf de eerste dag voorop moeten staan.

Ook de vraag hoe een meer praktisch ingestelde leerling theoretische kennis onder de knie kan krijgen, verdient veel meer aandacht op lerarenopleidingen, op scholen en bij de ontwikkeling van methoden. Het is een vraag waar niemand in een kenniseconomie, die steeds hogere eisen aan het scholingsniveau van de bevolking stelt, omheen kan. Experimenten zijn er; sommige scholen combineren voor VBO-leerlingen Frans met de kookles, en dan lukt het met de tweede moderne taal opeens wel. Een docent vergeleek de methoden voor de onderbouw na de komst van de basisvorming eens met een theezakje, dat door veelvuldig gebruik een steeds slapper aftreksel levert. VWO-leerlingen krijgen een pittig brouwsel, dat van de HAVO is al slapper, en in het VMBO gaat er zoveel water bij dat getwijfeld moet worden aan de drinkbaarheid. Wil de basisvorming ooit een succes kunnen worden, dan moet het anders. Praktische leerlingen verdienen een eigen didactische aanpak om theoretische kennis te leren verwerken.

Dat vergt inspanning van genoemde partijen, en investeringen van de overheid. ICT kan en moet daarbij een cruciale rol spelen. Met de computer kan het abstracte concreet gemaakt worden, en het theoretische beeldend. Het gebruik van nieuwe technologie in de basisvorming moet daarom veel verder gaan dan het vak informatiekunde alleen!

De overheid zal bij het ontwikkelen van dit soort nieuwe didactiek een actieve, stimulerende rol moeten spelen. Voor commerciële uitgeverijen blijkt telkens weer de doelgroep niet interessant en/of winstgevend genoeg om er voortvarend mee aan de slag te gaan. Uitzetten van onderzoeksprogramma's, vernieuwing van (universitaire) lerarenopleidingen, en ontwikkelen en stimuleren van didactische informatica zijn minimum-vereisten.

De conclusies van de Inspectie onderschrijven het belang van het stimuleren van een scholingscultuur onder leraren. De overheid kan daarbij verschillende *push and pull*-factoren inzetten. Enerzijds behoud van de onderwijsbevoegdheid afhankelijk maken van het blijven voldoen aan de modernste eisen aan het leraarsvak (zoals nog vast te leggen in de Wet op het Leraarschap), anderzijds zorgen dat scholen en leraren over voldoende tijd en middelen voor meer scholing beschikken.



*Ontwikkel een nieuwe adviesurentabel*

Het is de vraag of de 32-30-maatregel, zoals de Inspectie die voorstelt, de best denkbare manier is om die ruimte te bereiken. Zeker in een tijdperk dat eerder vraagt om meer dan minder scholing, is de suggestie om leerlingen twee uur per week minder op school te doen doorbrengen op zijn zachtst gezegd omstrede -al haalt de temperatuur van de discussie hierover het niet bij de hitte rond de vierdaagse lesweek in het basisonderwijs.

Het, ondanks alle inspanningen, groeiende lerarentekort maakt het lastig om lucht in het lesprogramma te brengen door een kortere lesweek voor leraren, zonder dat leerlingen daar wat van merken. Wat wel kan, is scholen uitgebreid faciliteren met meer ondersteunend personeel. Daarbij moet niet alleen gedacht worden aan traditionele ondersteuners als conciërges amanuenses en administratief medewerkers, maar ook aan systeembeheerders, personeelsmedewerkers, beleidsmedewerkers en maatschappelijk werkers. Zij kunnen leraren veel werk uit handen nemen, zodat de leraar zich kan concentreren op (het ontwikkelen van) zijn kerntaak: lesgeven. Ook de door de Inspectie richting scholen en leraren aanbevolen intercollegiale toetsing, teamwork en schoolbeleid kunnen een forse bijdrage leveren aan het bij de tijd houden van pedagogisch-didactische competenties. Scholen de ruimte geven een en ander zelf vorm te geven, sluit aan bij het versterken van de autonomie van scholen zoals vastgelegd in het regeerakkoord.

*Zorg voor een inhoudelijke oriëntatie van de leerlingen*

Vervang verplichte toetsing, advisering en certificering door een inhoudelijke oriëntatie van leerlingen op hun toekomst. Hierover kunnen we kort zijn: het zou goed zijn dit advies op te volgen. De toetsen kunnen een nuttige rol spelen voor de scholen, om intern de voortgang in het programma en de kwaliteit te meten. Maar forceren heeft geen zin, als leraren en schoolleiding de meerwaarde van verplichte toetsing niet zien.

*Extra investering*

Bovenstaande kan niet zonder een stevige financiële injectie van de overheid in de basisvorming. Gelukkig maakt de florissante staat van 's Rijks financiën dat ook mogelijk. Stimuleren van scholing en inzet van meer ondersteunend personeel, maar bijvoorbeeld ook een hoger salaris voor leraren in de onderbouw van het voortgezet onderwijs kunnen de basisvorming de nodige impulsen geven. Dat biedt vervolgens ruimte voor verwezenlijking van de idealen van de basisvorming: didactische vernieuwing, verhoging van het peil van het onderwijs, creëren van gelijkere kansen. De WRR ging er indertijd van uit dat invoering van de basisvorming ten minste tien jaar in beslag zou nemen. De ervaring leert dat die veronderstelling realistisch is. Nu we halverwege zijn, is een extra investering op zijn plaats om van de implementatie van de basisvorming een klinkend succes te maken.

**LITERATUUR**

- Inspectie van het Onderwijs (1999). *Werk aan de basis. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het Onderwijs.
- Westerbeek, K. (1999) *The colours of my Classroom: a study into the effects of the ethnic composition of classrooms on the achievements of pupils from different ethnic backgrounds*. Rotterdam: Centrum Educatieve Dienstverlening (CED).
- Wetenschappelijke Raad voor het Regeringsbeleid (1986). *Basisvorming in het onderwijs*. 's-Gravenhage: Staatsuitgeverij.

# Ik houd van basisvorming

Wim van de Camp, lid Tweede Kamer der Staten Generaal voor het CDA

## 1. DE START VAN DE BASISVORMING

Het was medio juni 1989; ik was met een aantal collega's op werkbezoek in Rotterdam; onderwerp stadsvernieuwing. Via de buschauffeur (mobile telefoons waren nog zeldzaam) kreeg ik te horen dat ik de heer Van der Voet moest bellen, de directeur voorlichting van de heer Lubbers. Het handelde om de kabinetsformatie. Mij werd verzocht met spoed naar Den Haag te reizen om daar de paragraaf onderwijs voor het nieuwe regeerakkoord voor te bereiden. Ik deed dat met mijn collega Ad Hermes en de PvdA-collega's Frans Leijnse en John Lillipaly.

Dit moment was voor mij de start van de basisvorming. Het was geen eenvoudige klus: meer dan vijftientig jaar onderwijsstrijd moest worden overbrugd. Sinds de invoering van de Mammoetwet in 1968 waren onder andere de Middenschool en het VBaO de revue gepasseerd. Het onderwijsveld heeft jaren in talrijke commissies overlegd: de verschillen van inzicht waren onoverbrugbaar.

Na veel passen en meten, waarbij de heren Lubbers en Wallage (de latere minister Jo Ritzen was nog niet in beeld) wel degelijk over onze schouders mee keken, is het ons gelukt een tekst te ontwerpen.

Een tekst met politieke compromissen<sup>1,2</sup> daar hoeft geen onduidelijkheid over te bestaan, derhalve soms botsend met gewenste wetenschappelijke pedagogische inzichten. De (politieke) hoofdpunten van destijds waren:

- uitgangspunt is het rapport 'Basisvorming in het Onderwijs' van de Wetenschappelijke Raad voor het Regeringsbeleid<sup>3</sup>;
- de basisvorming is primair een onderwijsinhoud, geen schooltype. Er komt derhalve geen stelselwijziging in het Nederlands voortgezet onderwijs;
- het aantal wekelijkse lesuren voor de leerling wordt uitgebreid van 30 naar 32 uur;
- er komt minimaal f 200 miljoen structureel beschikbaar om de basisvorming verantwoord te kunnen implementeren<sup>4</sup>. (Kom daar nu nog eens om en denk daarbij aan de invoering van de Tweede Fase Voortgezet Onderwijs en de invoering van het VMBO.)

De rest van de geschiedenis van de basisvorming laat ik hier nu achterwege. U kent die (al of niet persoonlijk aan den lijve ondervonden), dan wel kunt U die teruglezen in het evaluatie-rapport van de Inspectie<sup>5</sup> en beknopt in het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer.

## 2. WAT DE EVALUATIE ONS LEERT

Ik heb de evaluatie primair getoetst aan de drie hoofddoelen van de basisvorming:

- verhoging van het peil van het jeugdonderwijs,
- uitstel van studie en beroepskeuze,
- modernisering van het onderwijsprogramma.

Ook in mijn verdere betoog zal ik daar aan vasthouden. In alle discussies over overladenheid, een nieuwe adviesurentabel en de moeilijke positie van VBO-leerlingen lijken deze hoofddoelen nog wel eens vergeten te worden. Daarenboven merk ik op dat schrijvers dezes een eenvoudig politicus, van opleiding jurist is. Een aantal pedagogisch-didactische fitnesses en conclusies in de eerste evaluatie van de Inspectie zullen mij ongetwijfeld ontgaan, dan wel

zijn voor mij moeilijk van enig wetenschappelijk verantwoord commentaar te voorzien. Andere deskundigen zullen daar in deze aflevering van het Tijdschrift voor Onderwijsresearch ongetwijfeld op ingaan.

De evaluatie van de Inspectie is bij eerste waarneming niet negatief: er mankeert nog van alles aan de basisvorming, maar -en dat vind ik een belangrijke conclusie- de voortgaande invoering gedurende de periode 1998 - 2003 zal waarschijnlijk veel beter en sneller kunnen verlopen dan de invoering in de periode 1993 - 1998. Daarbij mag niet vergeten worden dat we anno 2000 op een groot aantal punten uitgaan van andere inzichten, daarover later meer, terwijl ook onze samenleving opnieuw ingrijpend is veranderd in het afgelopen decennium.

Wie dieper op de resultaten van de evaluatie ingaat ziet echter dat er nog veel moet verbeteren. De kwaliteit van het onderwijsleerproces laat bepaald te wensen over: de kwaliteit van de lessen is voor veel vakken nog niet goed. Dat kwaliteit van de lessen sterk samenhangt met de kwaliteit van de leraar blijkt opnieuw uit deze evaluatie.

Het aanbod van de modernere en in sommige onderwijskringen zo geliefde algemene vaardigheden is nog sterk onvoldoende.

In het algemeen vertonen de opbrengsten van de basisvorming nog een verdeeld beeld: zo is het sterk de vraag of de leerlingprestaties bij Nederlands (begrijpend lezen) en wiskunde wel substantieel zijn verbeterd.

Daar staat tegenover dat de invoering van de basisvorming naar mijn mening vele positieve 'neven-effecten' heeft gehad. Zo heeft er de afgelopen jaren in het Nederlandse voortgezet onderwijs een grootscheepse onderwijsvernieuwing plaatsgevonden die weliswaar niet rechtstreeks uit het concept basisvorming voortvloeit, maar daar wel nauw mee samenhangt. Dat klinkt vreemd maar zo is de vernieuwing en niet te vergeten de modernisering van tal van onderwijsmethoden direct het gevolg van de invoering van de basisvorming: het momentum was daar om ook een aantal andere zaken in de eerste fase voortgezet onderwijs aan te pakken.

De diverse vakinhouden en methodieken (onder andere het gebruik van werkboeken) zijn aangepast en gemoderniseerd. Bij veel vakken moest immers ook zonder invoering van de basisvorming het een en ander gebeuren en gezien de snelle (informatie)ontwikkelingen in onze maatschappij is dat een voortgaand proces.

Scholengemeenschappen en basisvorming hangen eveneens zeer met elkaar samen. In het kader van dit artikel voert het te ver om alle voor- en nadelen van de vorming van (brede) scholengemeenschappen uiteen te rafelen, wel kan gezegd worden dat het voortgezet onderwijs in Nederland georganiseerd is op een min of meer adequate schaal.

Wat wordt nog wel eens vergeten? Zoals hierboven reeds opgemerkt is de basisvorming destijds bedoeld als een onderwijshoud, niet als een schooltype. Daarin lag ook het compromis tussen CDA en PvdA. In de uitwerking en implementatie van de basisvorming is die keuze echter niet steeds helder doorgevoerd. Dat had enerzijds met politieke opvattingen te maken, maar zeker ook met ambtelijke en praktische invalshoeken.

Een goed, maar soms ook pijnlijk voorbeeld daarvan zijn de toetsen van de basisvorming. De toetsen zijn, zeker in de beginperiode, bron van grote irritatie geweest. Werkende weg zijn de toetsen en hun betekenis geminimaliseerd. Dat is een slechte zaak. In de huidige praktijk is veel moeilijker vast te stellen of de kerndoelen van de basisvorming door de diverse categorieën leerlingen worden behaald.

Over de toetsen en het bijbehorend circus ben ik zelf overigens nog steeds zeer ambivalent: aan de ene kant leiden te veel en te zware toetsen tot een te grote nadruk op de onderwijshoud basisvorming; het zou wel eens uit kunnen groeien tot een afzonderlijke schoolsoort! En dit terwijl de klassieke diplomering van VBO, MAVO en dergelijke overeind diende te blijven.

Maar waar nu voor gekozen wordt, ook in de aanbevelingen van de inspectie<sup>6</sup>, is ook allerm minst ideaal. De leerstof van bijvoorbeeld het VMBO of de onderbouw VWO gaat in die schoolsoorten zeker prevaleren, wordt getoetst, maar bestaat tegelijkertijd ook voor een groot gedeelte uit de stof van de basisvorming. Wat blijft er op deze manier over van het gezamenlijke concept van basisvorming?

Ik heb het uiteindelijk als wezensvreemd ervaren dat de toetsen basisvorming geen rol van betekenis konden c.q. mochten spelen bij de reguliere meting van de prestaties in de eerste fase van het voortgezet onderwijs. Uiteindelijk hadden die twee grootheden (afzonderlijk schooltype en basisvorming) meer geïntegreerd kunnen worden. Ik vind dat nog steeds wenselijk.

Tot slot van deze paragraaf merk ik, bescheiden als politici zijn, nog op dat het regeerakkoord 1989 in ieder geval een einde heeft gemaakt aan de jarenlange maatschappelijke en politieke impasse waarin de discussie over de toekomstige inrichting van de eerste fase voortgezet onderwijs verkeerde!!

### 3. HOE NU VERDER?

Op het moment dat ik dit artikel afrond (eind augustus 2000) is de kabinetsreactie bij de evaluatie van de Inspectie nog niet verschenen. Wel circuleren er op de Haagse burelen allerlei concept reacties. Dat maakt mijn benadering enigszins onzeker. Veel van mijn gedachten kunnen immers binnen zeer korte tijd door de actualiteit ingehaald worden. Desalniettemin een serieuze poging.

Allereerst het begrip overladenheid van het programma. Dat is in mijn ogen een uiterst relatief begrip. Ik ben niet erg onder de indruk van de analyses ter zake. Een leerling anno 2000 moet nu een maal veel kennen en kunnen om zich ook in de toekomst staande te houden in onze steeds nog complexer wordende samenleving. Daar zijn eerder meer kennis en vaardigheden voor nodig dan minder.

Wel dienen we ons bij voortduring af te vragen of de leerstof efficiënt en herkenbaar wordt aangeboden. Dat zou in de basisvorming wel eens niet het geval kunnen zijn!

Onnodige versnippering van de leerstof lijkt mij derhalve het echte probleem. De gedachte om een aantal vakken op een wetenschappelijk doordachte manier samen te voegen wijs ik dan ook zeker niet (meer) af. De benadering "terug naar tien vakken" kan daarbij als uitgangspunt gelden<sup>7,8</sup>.

Clustering van vakken op een verantwoorde manier zou ook de problematiek in de Tweede Fase kunnen verlichten en sluit bovendien beter aan bij de ervaringen van de leerlingen op de basisschool. Clustering van vakken werd tot nu toe veelal tegen gehouden door de - op zich rechtvaardige - lobby van de leraren(vak)organisaties, waar overigens ook mijn partij niet ongevoelig voor was!

Het vak informatiekunde als klein afzonderlijk vak blijkt door de snelle informatisering van onze maatschappij achterhaald.

Ik ben er zeker niet voor om de algemene minimumverplichting qua uren voor de vakken maatschappijleer, lichamelijke opvoeding en de kunstvakken los te laten. Die gedachte staat immers principieel haaks op de zo moeizaam verworven breedte in het vakkenaanbod van de basisvorming. Die breedte moet naar mijn mening absoluut behouden blijven.

Tevens een korte opmerking over de zogenaamde 32/30 maatregel. De fractiespecialisten onderwijs van CDA en PvdA hebben daar destijds fors voor moeten knokken. Wanneer het aantal contacturen met twee wordt teruggebracht om de leraren meer tijd voor voorbereiding en verdieping te geven zou ik daarbij twee voorwaarden willen stellen:



- de bekostiging van die twee uren blijft onverbloemd - derhalve herleidbaar - behouden en
- de twee uren worden omgezet in leerling- dan wel huiswerkbegeleiding. Het programma is immers zwaar genoeg.

Vervolgens leven er sterke gedachten op opnieuw te komen tot een adviesrentabel<sup>9</sup>. Dat lijkt me echt de verkeerde benadering en zet de klok ook minstens tien jaar terug. Mijn advies is: nooit doen, staat haaks op de ontwikkeling naar meer autonome scholen. We kennen maar al te goed de 'voorspellende, uniformerende werking' van een adviesrentabel. Het is juist de professionaliteit van de school, het team, maar zeker ook van de individuele leerkrachten om zelf te bepalen op welke wijze de leerstof wordt aangeboden.

Daarnaast zal een adviesrentabel weinig ruimte laten voor nieuwe pedagogisch-didactische inzichten. Zo maakt de snelle informatisering van onze samenleving nieuwe werkvormen mogelijk, voor een aantal vakken meer of minder contacturen, kortom gedurfde differentiatie waarin ook de verschillen tussen leerlingen en schoolsoorten meer ruimte krijgen<sup>10</sup>.

In de 'nieuwe basisvorming' kan en dient er meer ruimte te komen voor verschillen tussen kinderen. Aan het begin van deze nieuwe eeuw gaan wij immers meer open en eerlijker met die verschillen om. De jaren zestig en zeventig werden - in dat tijdsbeeld wellicht zeer nodig-gekenmerkt door gelijkheid en gelijkwaardigheid. Nu weten we dat alle kinderen gelijkwaardig zijn, maar niet gelijk, we kunnen daar pedagogisch mee omgaan. Maar ieder in Nederland opgroeiend kind dient de kerndoelen van de basisvorming te behalen, in zijn of haar eigen tempo, op de voor hem of haar bestemde schoolsoort.

Deze laatste opmerking brengt mij nog tot een andere zijstap: ik heb nooit kunnen begrijpen waarom de kerndoelen voor veel VBO-leerlingen niet haalbaar zijn. Naar mijn bescheiden mening is dat niet een kwestie van inhoud of moeilijkheidsgraad van de kerndoelen, maar veeleer het achterwege laten van het ontwikkelen van de juiste pedagogisch-didactische methoden om de stof van de basisvorming aan deze leerlingen aan te bieden. Als ik na ga waartoe de gemiddelde VBO-leerling later in zijn of haar onderwijs carrière in staat is, als hij of zij eenmaal het theoretische karakter van de kerndoelen basisvorming heeft verlaten (ik denk daarbij aan het pure vakmanschap, de beroepsbegeleidende leerwegen, maar ook het met succes volgen van lange MBO-opleidingen) dan begrijp ik niet waarom de huidige kerndoelen bij voortduring een struikelblok moeten zijn. De veelal geroemde 'praktisch ingestelde leerling' zou toch via een 'praktische manier van onderwijzen' de kerndoelen moeten kunnen behalen.

Ook bij een eventuele invoering van de partiële leerplicht voor 14- en 15-jarigen (vergeleijk de discussie in Amsterdam van medio 1999) zou ik vast willen houden aan het behalen van de kerndoelen. Dan over een langere tijdperiode verdeeld en al of niet met deelcertificaten behaald in de praktijk.

Een zelfde aandacht vraag ik voor de allochtone leerling. Ik hoor het geluid minder, maar nog te veel: 'voor sommige groepen allochtone leerlingen zijn de kerndoelen basisvorming niet haalbaar'. Dat vind ik absoluut de weg van de minste weerstand. Het is juist de allochtone leerling die, om tot een volwaardige inburgering te komen de kerndoelen basisvorming dient te behalen. Wat waren de drie hoofddoelen van de basisvorming ook al weer?

Tot slot van deze paragraaf wil ik in meer algemene zin nog de vraag op werpen of het VMBO concept niet ten principale haaks staat op de grondideeën van de basisvorming. Ik snij dit onderwerp niet aan om nu het gehele concept van het VMBO ter discussie te stellen - de aanbevelingen van de Inspectie roepen de overheid juist op consistent te zijn met onderwijsvernieuwingen - maar meer als theoretische exercitie ten behoeve van toekomstige discussies. Die vraag is naar mijn mening nooit goed beantwoord, ook niet door mijn eigen partij. Mevrouw Netelenbos omarmde destijds super snel het advies van de commissie-Van Veen; we moesten tenslotte wat doen voor de praktisch ingestelde leerling. En de VVD vond het



allang best. Zij hebben destijds, overigens bepaald niet lichtvochtig, tegen de invoering van de basisvorming gestemd en vinden iedere afzwakking daarvan nog steeds een prima benadering.

Ik word nog gesterkt in deze opvatting nu ik in de concept beleidsreactie van Mevrouw Adelmund de neiging bespeur om de basisvorming te beperken tot twee leerjaren. Dat wordt opnieuw voor grote groepen leerlingen een ramp. Ik heb het altijd vreemd gevonden dat de betere theoretische leerling op het Gymnasium en in het VWO in principe drie jaar over de basisvorming mag doen en de minder theoretisch getalenteerde leerling op het VBO maximaal twee jaar omdat er zeker ruimte moet zijn voor de beroepsvoorbereidende vakken.

Het idee dat je in vier jaar de basisvorming én een zekere vorm van beroepskwalificatie kan behalen is niet weg te branden uit onze samenleving: menige vereniging van ondernemers is die mening nog steeds toegedaan. Mijn stelling is en blijft dat in het huidige Nederlandse systeem je voor beroepskwalificatie een ROC-opleiding nodig hebt, van welke inhoud (vanaf assistenten niveau) en duur (minimaal één jaar) dan ook.

Geef alstublieft de diverse categorieën leerlingen de kans de (brede) kerndoelen basisvorming te behalen in de tijd die zij daar voor nodig hebben. Met methoden, beter op de diversiteit van de leerlingen toegeschreven, moet dit lukken. Basisvorming is immers het fundament voor 'een leven lang leren'.

#### 4. CONCLUSIE

Er is naar mijn mening geen enkele reden om het concept basisvorming in het Nederlandse bestel van voortgezet onderwijs te verlaten. In een steeds meer multiculturele samenleving is het een basisvoorwaarde dat alle kinderen een gezamenlijk pakket aan kennis en ervaring aanleren. De gezamenlijke culturele ervaring van onze kinderen aan het begin van het voortgezet onderwijs is immers steeds geringer in omvang.

De hedendaagse politiek - maatschappelijke discussie staat bol van termen als sociale cohesie of sociale samenhang. Die wordt mede aangeleerd door de basisvorming. Een al of niet geïntegreerd vak verzorging speelt daarin bijvoorbeeld een grote rol.

Er zal meer rekening moeten worden gehouden met verschillen tussen leerlingen: de één verwerft de kerndoelen basisvorming nu eenmaal sneller en op een meer theoretische wijze dan de ander, laten we daar alert op zijn.

De drie hoofddoelen van de basisvorming, verhoging van het peil van het jeugdonderwijs, uitstel van studie en beroepskeuze en modernisering van het onderwijsprogramma staan in 'Nederland Kennisland' nog keihard overeind.

#### NOTEN

1. Tweede Kamer, vergaderjaar 1989 - 1990, 21 132, nr. 8.
2. Zo werden in een later stadium, bij de behandeling van de wet, de uren aardrijkskunde, een van de laatste strijdpunten, letterlijk gemiddeld. De toenmalige vice-fractie voorzitter van het CDA Drs. Frans Wolters speelde 'postiljon d'amour' tussen onze fractie en het departement (de heer Wallage).
3. Rapport nr. 27, *Basisvorming in het onderwijs* d.d. 22 januari 1986.
4. In totaal bevatte het Regeerakkoord Lubbers - Kok (1989 - 1994) een structurele verhoging van de onderwijsuitgaven met f 500 miljoen. Naast de ± f 200 miljoen voor de basisvorming was er ± f 270 miljoen opgenomen voor het ongedaan maken van de zogenaamde '4 - jarigen' maatregel en de versterking van het Speciaal Onderwijs, wat later resulteerde in de operatie 'Weer Samen Naar School'.

5. Inspectie van het Onderwijs (1999). *Werk aan de basis. Evaluatie van de basisvorming na vijf jaar*. Utrecht: Inspectie van het Onderwijs.
6. *Werk aan de basis*, bladzijde 126, voorstellen 4, 5, en 6 inzake de verplichtende afsluitende toetsen, het verplichte advies aan het einde van het tweede jaar en het uit te reiken certificaat.
7. Combinaties van natuurkunde - scheikunde en techniek, van biologie en verzorging of van verzorging, aardrijkskunde en economie behoren zeker tot de mogelijkheden.
8. Voor het verschil tussen Algemene Natuur Wetenschappen en het vak Science verwijs ik gaarne naar Mevr. drs. Agnes Legierse van SLO die mij daarover eens een verhelderende notitie stuurde.
9. *Werk aan de basis*, bladzijde 125, Aanbevelingen aan de overheid no. 3.
10. Met dank aan het KPC te 's-Hertogenbosch: *De 21<sup>e</sup> CarmelSchool, een school in de 21<sup>e</sup> eeuw. Voorstudie voor een nieuw schoolontwerp*.

## **Regeringsstandpunt stelt teleur**

Ursi Lambrechts, lid van de Tweede Kamer der Staten Generaal voor D66

### **1. ADELMUND MAAKT ONVOLDOENDE GEBRUIK VAN EVALUATIERAPPORT**

Je mag rustig stellen dat D66 zwaar teleurgesteld is in het regeringsstandpunt over de basisvorming. De echte knelpunten worden niet opgelost. De hete aardappel wordt doorgeschoven naar de volgende regeerperiode en de evaluatie van de Onderwijsinspectie is nauwelijks benut.

September vorig jaar heeft de Onderwijsinspectie een uitgebreid en overtuigend evaluatierapport over de basisvorming gepresenteerd. De neerslag van ruim een jaar onderzoek.

In dat rapport worden heel helder de knelpunten van de basisvorming aangegeven

- overladenheid van het aanbod (te veel vakken);
- versnippering van het aanbod (te veel kleine vakken);
- gebrek aan samenhang tussen de vakken;
- te weinig differentiatie waardoor het voor de ene leerling (VBO/MAVO) te moeilijk was en voor de andere (HAVO-VWO) te weinig uitdagend.

### **2. ECHE VERBETERINGEN OPGESCHORT**

Naast knelpunten geeft het inspectierapport ook nog een aantal oplossingen aan voor de problemen waaronder een kerncurriculum.

Eerlijk gezegd ben ik er steeds van uitgegaan dat het voorstel van de Staatssecretaris waar we zo lang op hebben moeten wachten, een verdere uitwerking van deze voorstellen zou brengen. Hoe de overladenheid en versnippering aan te pakken, hoe de aansluiting met het basisonderwijs te verbeteren en hoe te komen tot een kerncurriculum, zodat er echte ruimte voor scholen gaat ontstaan. Helaas is dat niet het geval.

Aan de ene kant onderschrijft staatssecretaris Adelmund wel de knelpunten en gebreken van de basisvorming, zoals die zijn geconstateerd door de Inspectie. Tegelijkertijd echter schuift zij de beslissing tot welke aanpassingen in het programma dit moet leiden voor zich uit, waardoor in deze regeerperiode geen echte verbeteringen meer te realiseren zullen zijn.

De Staatssecretaris gaat nu eerst opnieuw advies vragen aan de Onderwijsraad over vragen die een jaar geleden al helder op tafel lagen. Jammer. Als extra advies van de Onderwijsraad nodig is, had het afgelopen jaar daarvoor benut moeten worden. Dan hadden we nu daarover ons oordeel kunnen geven en nog in deze periode de verbeteringen aanbrengen waarvan iedereen inmiddels weet dat ze nodig zijn.

### **3. HUIDIGE PRAKTIJK BEVESTIGD**

Op dit moment volstaat de Staatssecretaris met een aantal wijzigingsvoorstellen die ik voor het merendeel niet anders kan typeren dan als het 'witwassen' van de huidige praktijk:

- Scholen mogen zelf keuzes maken in het overladen aanbod aan kerndoelen. Maar scholen deden dat al lang want uit de evaluatie blijkt dat scholen niet meer dan 6 van de 10 kerndoelen realiseerden.

- Scholen mogen minder lesweken geven. Daarmee wordt formeel bevestigd wat in de praktijk eveneens ook al zo functioneert.
- De verplichte basisvormingstoetsen worden afgeschaft, waartoe overigens al eerder besloten was.

#### 4. TE VERREGAANDE CLUSTERING VAN VAKKEN

Wat wel nieuw is, is de mogelijkheid voor hele verregaande clustering van vakken. Juist daarbij vraagt D66 zich af of dat nu wel zo gewenst is. Een vermindering van het totaal aantal vakken per leerling met een beperkte clustering waarbij de vakken verzorging, techniek en informatiekunde elders aansluiting vinden heeft daarbij onze -D66- voorkeur.

#### 5. GEMISTE KANS

Al met al vind ik dit een gemiste kans om op afzienbare termijn te komen tot de verbeteringen die in onze ogen broodnodig zijn:

- verminderen van de kerndoelen;
- minder vakken met meer samenhang;
- betere aansluiting op het basisonderwijs middels doorlopende leerlijnen;
- differentiatie binnen de programma's van de verschillende onderwijstypen.

Door deze gang van zaken is niet alleen een jaar verloren gegaan, maar eigenlijk een hele regeerperiode. In deze regeerperiode zullen er immers geen wezenlijke verbeteringen meer worden aangebracht. Dat betreuren we zeer.

# Hervorming van de basisvorming

Clemens Cornielje, *lid van de Tweede Kamer der Staten Generaal voor de VVD*  
Luciënne Middelhof, *beleidsmedewerker bij de VVD*

## 1. TERUGBLIK

Binnen het onderwijs is er jarenlang discussie geweest over de vormgeving en inhoud van het voortgezet onderwijs. De aanloop tot het wetsvoorstel 'Basisvorming' heeft zijn oorsprong in de Mammoetwet van Minister Cals (1968). Met de invoering van de Mammoetwet werd de samenhang in het stelsel van voortgezet onderwijs vergroot. Er ontstonden meer doorstroommogelijkheden, zowel van lagere naar hogere schooltypen als omgekeerd. Uitstel van studie- en beroepskeuze werd hierdoor mogelijk. In 1967-1971 bracht minister Veringa een beleidsnotitie uit over geïntegreerd voortgezet onderwijs. Minister van Veen (1971-1973) kondigde vervolgens experimenten aan met geïntegreerd voortgezet onderwijs. In 1973-1977 kwam minister van Kemenade met de 'Contourennota' - en de 'Vervolgnota' - met hierin een schets van de Middenschool. In 1976 gaan drie experimenten middenschool van start. De politieke strijd gaat over de vraag wat de betekenis is van de start van deze experimenten: gaat het om de vraag of de middenschool wordt ingevoerd of hoe het middenschoolonderwijs eruit gaat zien. In 1977 treedt Minister Pais aan en noemt op de eerste dag van zijn nieuwe functie de middenschool een 'eenheidsworst'. In 1979 laat hij zijn 'Ontwikkelingsplan voor het voortgezet onderwijs' (OPVO) verschijnen met daarin het voorstel van een tweejarig geïntegreerde brugperiode.

In 1981-1982 komen minister van Kemenade en staatssecretaris Deetman met de nota 'Verder na de basisschool', waarin een nieuw type school wordt voorgesteld: Voortgezet Basisonderwijs (Vbao). Met opnieuw het ideaal van de middenschool en discussies over duur van en eventuele differentiaties in de eerste jaren van het voortgezet onderwijs. Een Centrale Werkgroep (bestaande uit vertegenwoordigers van departement, vakorganisaties, besturenorganisaties en ouderorganisaties) werkt de nota op tal van onderdelen concreet uit, met als doel: het leveren van bouwstenen voor wetgeving die reeds een breed maatschappelijk draagvlak zou hebben, voordat zij in parlementaire behandeling zou worden genomen (vanwege politisering in de jaren zeventig).

Minister Deetman en staatssecretaris Ginjaar-Maas komen in 1985 met het wetsvoorstel 'Ontwikkelingswet voortgezet onderwijs', dat het nooit tot wet heeft gebracht. In 1987 wordt dan eindelijk het wetsvoorstel tot wijziging van de Wet op het Voortgezet Onderwijs (Wet op de Basisvorming) uitgebracht, dat in mei 1992 door de Eerste Kamer wordt aangenomen op voorpraak en verdediging van staatssecretaris Wallage. Aan de uiteindelijke wetgeving ligt het rapport 'Basisvorming in het Onderwijs' van de Wetenschappelijke Raad voor het Regeringsbeleid uit 1986 ten grondslag. De WRR stelde vast dat de discussie over de vorm en inhoud van het voortgezet onderwijs zodanig was gepolariseerd dat elk voorstel de één te ver ging en de andere niet ver genoeg. De WRR koos daarom voor een meer inhoudelijke benadering binnen de bestaande structuur. In de eerste leerjaren van het voortgezet onderwijs zou er een gemeenschappelijke onderwijsinhoud moeten komen voor alle leerlingen, welke afgesloten zou moeten worden op twee niveaus. Tegelijkertijd adviseerde zij de bestaande schooltypen te laten voortbestaan. De VVD was positief over het advies van de WRR en deze vorm van basisvorming. Dit veranderde toen in de politieke besluitvorming de nadruk steeds meer kwam te liggen op de structuur in plaats van de inhoud. De VVD-fractie onderschreef de wens om alle leerlingen op een zo hoog mogelijk onderwijspeil te brengen, maar met de basisvorming wer-



den er onvoldoende garanties geboden dat er rekening werd gehouden met verschillen tussen leerlingen. Voor de meer praktisch ingestelde leerlingen is de basisvorming met dezelfde algemene leerinhoud en dezelfde toetsen te moeilijk. En de leerlingen die qua niveau een zwaarder programma aankunnen vinden onvoldoende uitdaging in het 'niveauloze' programma. De VVD heeft destijds dan ook als enige fractie tegen het wetsvoorstel gestemd.

## 2. OPLOSSEN VAN KNELPUNTEN

Doelstelling van de basisvorming is uitstel van het moment van verplichte studie- en beroepskeuze; brede vorming voor alle leerlingen en een verhoging van het peil van het onderwijs (zie ook het artikel 'Beleidsmatige voorgeschiedenis en opzet' elders in dit themanummer). Als middel om dat te bereiken werd een vakkenpakket omschreven. De gehele operatie ging gepaard met een vorming van brede scholengemeenschappen. De VVD heeft zich neergelegd bij democratisch genomen besluiten. Wel hebben wij vanaf de invoering van de basisvorming meermalen in de Tweede Kamer naar voren gebracht dat regering en parlement gehouden zijn optredende knelpunten op te lossen.

De VVD heeft voorstellen gedaan om een aantal knelpunten op te lossen.

### *Afschaffing verplichte toetsing*

De verplichte toetsing basisvorming is afgeschaft (motie Cornielje 25 600, nr. 41). Toetsing was destijds ingezet als middel om de basisvorming effectief in te voeren, maar doordat bij de samenstelling van de toetsen geen onderscheid gemaakt werd naar schoolsoorten, dus dezelfde toetsen voor iedereen, ontstonden op scholen vele problemen. Allereerst het volumeprobleem, door een enorm aantal toetsen dat afgelegd diende te worden; verder bleek éénzelfde toets voor alle leerlingen niet hanteerbaar; bovendien bleek er een spanning te bestaan tussen basisvorming en VBO/MAVO-opleidingen: doordat bepaalde vakken in het derde leerjaar niet meer terugkwamen, moesten deze al in het tweede leerjaar getoetst worden. Een paradoxale situatie, omdat juist deze leerlingen, gezien hun capaciteiten de basisvorming eerder moesten afsluiten, dan HAVO/VWO-leerlingen. De VVD was daarom voor afschaffing van de verplichting tot toetsing.

### *Geen verdere fusiestimulering*

Er is een einde gekomen aan verdere fusiestimulering (motie Cornielje 24 400 VIII, nr. 36). Brede scholengemeenschappen zouden in belangrijke mate de doorstroommogelijkheden van leerlingen bepalen, maar uit het Onderwijsverslag 1998 blijkt juist dat de breedte van de school slechts van geringe invloed is op de doorstroming van leerlingen na het tweede leerjaar, behalve voor VWO-leerlingen. Door de groei van het aantal scholengemeenschappen is een doorstroming op gang gekomen van VWO naar HAVO en van MAVO naar VBO. Een ander nadeel is dat het percentage lesuitval bij grote scholengemeenschappen (meer dan 1000 leerlingen) boven het landelijk gemiddelde ligt. Verder worden grote scholen bevoordeeld door het fusiestimuleringsbeleid. Dit gaat ten koste van de positie van kleine scholen.

### *Afschaffing getuigschrift*

Het getuigschrift basisvorming hoeft niet meer uitgereikt te worden. Het getuigschrift basisvorming wordt door scholen als zinloos ervaren, het komt als mosterd na de maaltijd. Scholen mogen het getuigschrift alleen uitreiken aan leerlingen die na afronding van de basisvorming de school zonder diploma verlaten en aan leerlingen die, nadat ze voor het schoolexamen zijn geslaagd, de school gediplomeerd verlaten. Scholen reiken het in de praktijk dan ook nauwelijks uit.

### *Meer ruimte voor praktische beroepsgerichte vakken*

Er is meer ruimte gekomen voor praktische beroepsgerichte vakken in de eerste leerjaren van het VMBO. Kern van de inzet van de VVD bij de behandeling van het wetsvoorstel VBO/MAVO was het bieden van meer ruimte voor praktische beroepsgerichte vakken (in tegenstelling tot algemeen vormende vakken) in de eerste leerjaren van het VBO. De VVD wilde dat er vanaf het eerste leerjaar begonnen mag worden met praktische beroepsgerichte vakken, terwijl de basis van de inhoudelijke doelstellingen van de basisvorming overeind wordt gehouden en wordt uitgesmeerd over vier jaar. In de vierde nota van wijziging wordt de door de VVD-fractie al zo lang en vurig gewenste mogelijkheid gecreëerd om het beroepsgerichte karakter van het VBO te vergroten. Het wordt tevens mogelijk om reeds in het eerste leerjaar aan te vangen met een oriëntatie op de beroepsgerichte vakken. Toch ging dit de VVD niet ver genoeg, zij heeft daarom een wijzigingsvoorstel ingediend. De beperking om leerwegondersteunend onderwijs als voorwaarde te stellen om sneller te kunnen aanvangen met beroepsgerichte vakken is niet noodzakelijk. Door de verdere druk (dreiging met amendementen) is door de Staatssecretaris uiteindelijk de vijfde nota van wijziging ingediend. Deze biedt twee mogelijkheden. De eerste is dat 10 procent van de eerste twee leerjaren voor alle VBO/MAVO-leerlingen besteed mag worden aan oriëntatie op de leerwegen. De tweede is dat voor leerwegondersteunend geïdentificeerde leerlingen geldt dat 20 procent van de eerste leerjaren aan oriëntatie op de basisberoepsgerichte leerweg mag worden besteed.

### *Verticale leerlijnen*

Geen fusies meer, maar wel een betere inhoudelijke afstemming van VMBO- en MBO-programma's (verticale leerlijnen). Fusies bieden slechts structuuroplossingen, terwijl onderwijsinhoudelijke oplossingen door goede afstemming van onderwijskundige programma's juist zijn geboden. Het beroepsonderwijs kan verder worden versterkt door met leerlijnen te werken. Het gebruiken van een leerlijn kan bijdragen aan een positievere uitstraling van de beroepsonderwijskolom VMBO, MBO en HBO. Dan zal duidelijk worden dat een plaats in het MBO een beroepskwalificatie en een doorstroomkwalificatie oplevert.

### *Meer basisvormingvarianten*

Hans Dijkstal heeft in april van dit jaar tijdens het Integratiedebat één of zelfs meer knippen in de basisvorming bepleit. Namelijk een meer praktische beroepsgerichte variant met een concentratie op de kernvakken voor het beroepsonderwijs. Vervolgens een wat uitdagender basisvorming voor het algemeen vormend onderwijs. En tenslotte een praktisch beroepsgerichte variant, waarbij werkend-leren mogelijk wordt, met instandhouding natuurlijk van de leerplaat.

## **3. POSITIEVE BELEIDSVOORSTELLEN ADELMUND**

De VVD betreurt het dat de evaluatie van de Onderwijsinspectie in 1999 zich niet heeft beperkt tot de basisvorming, maar zich heeft uitgestrekt tot de kwaliteit van het voortgezet onderwijs in het algemeen. Het wordt daardoor niet helder wat de effecten van de invoering van de basisvorming zijn. Verder is de fusiestimulering buiten beschouwing gebleven, terwijl de vorming van brede scholengemeenschappen wel gekoppeld was aan de invoering van de basisvorming.

De beleidsvoorstellen van Karin Adelmund daarentegen zijn positief en komen neer op een dubbele knip en dus een driedeling van de basisvorming conform de suggestie van Hans Dijkstal met meer ruimte voor verschillen tussen leerlingen:

- Basisvorming van twee jaar voor praktisch ingestelde leerlingen, met extra ruimte voor praktisch beroepsgerichte vakken.
- Basisvorming van twee jaar voor HAVO/VWO-leerlingen. De voornamelijk algemeen vormende vakken worden op een hoog niveau afgesloten.

- Ten slotte een twee- of drie-jarige variant van de basisvorming voor de meeste VMBO-leerlingen, waarbij het vakkenaanbod en het niveau van de vakken beter aansluiten bij de mogelijkheden van de leerlingen.

Het zou van moed getuigen als de Staatssecretaris de scholen meer houvast zou bieden. Nu heeft haar voorstel iets vrijblijvends. De scholen mogen het zelf uitzoeken. Als dat zo blijft, dan is er helemaal niets meer over van wat je basisvorming zou kunnen noemen.

## Medewerkers

Aan de totstandkoming van dit themanummer over de basisvorming werkten de volgende personen mee.

**Drs. Marleen Barth** is lid van de Tweede Kamer der Staten Generaal voor de PvdA.

E-mail: : [M.Barth@tk.parlement.nl](mailto:M.Barth@tk.parlement.nl)

**Dr. Huub van den Bergh** is als universitair hoofddocent verbonden aan de Faculteit Letteren van Universiteit Utrecht, Trans 10, 3512 JK Utrecht.

E-mail: [huub.vandenbergh@let.ruu.nl](mailto:huub.vandenbergh@let.ruu.nl)

**Prof. dr. Kerst Th. Boersma** is als hoogleraar verbonden aan de Universiteit Utrecht, Centrum voor Didactiek van Wiskunde en Natuurwetenschappen, Princetonplein 5, 3584 CC Utrecht en tot 1 januari 2001 tevens aan de SLO, Postbus 2041, 7500 KA Enschede.

E-mail: [k.th.boersma@pys.uu.nl](mailto:k.th.boersma@pys.uu.nl)

**Prof. dr. Roel J. Bosker** is als hoogleraar werkzaam aan de Universiteit Twente, Postbus 217, 7500 AE Enschede.

E-mail: [r.j.bosker@edte.utwent.nl](mailto:r.j.bosker@edte.utwent.nl)

**Dr. Hennie Brandsma** werkte ten tijde van het gerapporteerde onderzoek aan de Faculteit Toegepaste Onderwijskunde van de Universiteit Twente en is thans werkzaam bij ECNO, Educatief Centrum Noord en Oost, Ubbo Emmiusingel 1a, Postbus 804, 9701 AV Groningen.

E-mail: [h.p.brandsma@ecno.nl](mailto:h.p.brandsma@ecno.nl)

**Mr. ing. Wim G.J.M. van de Camp** is lid van de Tweede Kamer voor het CDA. Hij is belast met de portefeuilles onderwijs en justitie. (De portefeuille voortgezet onderwijs wordt binnen de CDA-fractie inmiddels beheerd door zijn collega Dr. Aart Mosterd.)

E-mail: [W.vdCamp@tk.parlement.nl](mailto:W.vdCamp@tk.parlement.nl)

**Clemens G.A. Cornielje** is lid van de Tweede Kamer der Staten Generaal voor de VVD.

E-mail: [C.Cornielje@tk.parlement.nl](mailto:C.Cornielje@tk.parlement.nl)

**Drs. Wim J. Deetman** is thans burgemeester van Den Haag. Hij was Staatssecretaris van Onderwijs van 1981-1982 en Minister van Onderwijs van 1982-1989.

**Dr. Wim J.C.M. van de Grift** is als stafinspecteur verbonden aan de Inspectie van het Onderwijs, Rijksinspectiekantoor Utrecht, Postbus 2730, 3500 GS Utrecht.

E-mail: [w.vandegrift@owinsp.nl](mailto:w.vandegrift@owinsp.nl)

**Anke Hartevelde** is als schoolleider verbonden aan het Petrus Canisius College te Alkmaar.

E-mail: : [a.hartevelde@pcc.nu](mailto:a.hartevelde@pcc.nu)

**Piet L.A. Helmholt** is als inspecteur van het basisonderwijs werkzaam bij de Inspectie van het Onderwijs, Rijksinspectiekantoor Groningen, Postbus 706, 9700 AS Groningen.

E-mail: [P.Helmholt@owinsp.nl](mailto:P.Helmholt@owinsp.nl)

**Drs. Frans Kleintjes** is als methodoloog verbonden aan het Cito, Postbus 1034, 6801 MG Arnhem.

E-mail: [frans.kleintjes@cito.nl](mailto:frans.kleintjes@cito.nl)

**Drs. Ursi Lambrechts** is lid van de Tweede Kamer der Staten Generaal voor D66.  
E-mail: [U.Lambrechts@tk.parlement.nl](mailto:U.Lambrechts@tk.parlement.nl)

**Prof. dr. Wim J. van der Linden** is als hoogleraar verbonden aan de Universiteit Twente, Postbus 217, 7500 AE Enschede.  
E-mail: [w.j.vanderlinden@edte.utwente.nl](mailto:w.j.vanderlinden@edte.utwente.nl)

**Drs. Lucienne Middelhof** is beleidsmedewerker bij de VVD-kamerfractie.  
E-mail: [L.Middelhof@tk.parlement.nl](mailto:L.Middelhof@tk.parlement.nl)

**Drs. Marielle W.H. Peters-Sips** werkt bij de Inspectie van het Onderwijs als coördinerend inspecteur voortgezet onderwijs, Rijksinspectiekantoor Utrecht, Postbus 2730, 3500 GS Utrecht.  
E-mail: [m.peters-sips@owinsp.nl](mailto:m.peters-sips@owinsp.nl)

**Prof. dr. Gert Rijlaarsdam** is werkzaam bij de Universitaire Lerarenopleiding van de Universiteit van Amsterdam, Wibautstraat 2-4, 1091 GM Amsterdam, en de faculteit Letteren van de Universiteit Utrecht, Trans 10, 3512 JK Utrecht.  
E-mail: [rijlaars@ilo.uva.nl](mailto:rijlaars@ilo.uva.nl)

**Drs. Lisette Schuurmans** werkt als wetenschappelijk medewerker voor de Inspectie van het Onderwijs, Rijksinspectiekantoor Utrecht, Postbus 2730, 3500 GS Utrecht.  
E-mail: [l.schuurmans@owinsp.nl](mailto:l.schuurmans@owinsp.nl)

**Mia Verhagen** is als schoolleider verbonden aan de Open Schoolgemeenschap Bijlmer te Amsterdam.  
E-mail: [vansteijnverhagen@planet.nl](mailto:vansteijnverhagen@planet.nl)

**Drs. Jacques Wallage** is thans burgemeester van Groningen. Hij was Staatssecretaris van Onderwijs van 1989-1993.

**Dr. Michel Zwarts** is werkzaam als stafinspecteur voortgezet onderwijs voor de Inspectie van het Onderwijs, Rijksinspectiekantoor Utrecht, Postbus 2730, 3500 GS Utrecht.  
E-mail: [m.zwarts@owinsp.nl](mailto:m.zwarts@owinsp.nl)



## Notities en Commentaren

### HET MINIMUM AANTAL ITEMS IN EEN MULTIPLE-CHOICE- OF OPEN-ANTWOORDTOETS

Willem K.B. Hofstee en Frits E. Zegers

Rijksuniversiteit Groningen

#### Abstract

*Under defensible assumptions, and using Lord's (1959) approximation formula for the standard error of measurement, the required number of items  $m$  in a multiple-choice test (of which the open-answer test is treated as a limiting case) appears to be a function of the number of response alternatives per item  $A$ , namely:  $m \geq 20.54[A/(A-1)]^2$ . For example, a true-false test should contain four times as many items as an open-answer test.*

#### Inleiding

Deze notitie bevat een beredeneerd antwoord op de vraag 'hoeveel items moet een toets bevatten'. Om te beginnen het antwoord: onder verdedigbare assumpties hangt het aantal items alleen af van het aantal alternatieven  $A$  per vraag, en luidt het antwoord als volgt:

|                          |             |    |
|--------------------------|-------------|----|
| A:                       | Itemaantal: |    |
| 2                        | 83          |    |
| 3                        | 47          |    |
| 4                        | 37          |    |
| $\infty$ (open-antwoord) |             | 21 |

Deze aantallen zijn uiteraard indicatief; de redenering bevat ingrediënten die systematisch kunnen worden gevarieerd en is dus 'exemplarisch'.

#### Redenering

Als algemeen uitgangspunt fungeert hier de stelling dat de inrichting van het onderwijs dient te worden gebaseerd op de fictie van de *normstudent* en met name niet op de realiteit van de *modale student*. De normstudent is iemand met een (voor de opleiding in kwestie) modale intelligentie, die echter *de volle tijd die er voor staat aan de studie besteedt*, zulks in tegenstelling tot de modale student. Op de achtergrond van deze fictie staat een beeld van de normstudent als iemand die probeert zich de inhoud van het vak in kwestie eigen te maken en er zich een oordeel over te vormen, in plaats van te streven naar het behalen van studiepunten volgens het homo-economicusmodel. Dat eerste kost meer tijd en aandacht dan het laatste.

Als onderwijs en examens worden afgesteld op de *modale student*, bijvoorbeeld omwille van het numeriek rendement, kan dat alleen maar leiden tot diploma-inflatie en kwaliteitsverlies. Omgekeerd heeft de opleiding zich wel degelijk te richten naar de *normstudent*. Met name dient de opleiding redelijkerwijs te garanderen dat deze geen noemenswaardige obstakels in de studie ondervindt. In de optiek van de (nogmaals fictieve) normstudent hebben tentamens louter een marginale functie, in plaats van een centrale; de opleiding dient in die attitude haar aandeel te nemen. Men zou in dit verband kunnen spreken van *normstudent-georiënteerd* onderwijs, ter overbrugging van de valse tegenstelling tussen student- en leerstofgeoriënteerd onderwijs.

Specifiek houdt oriëntatie op de normstudent hier in dat deze niet dan bij hoge uitzondering onvoldoendes haalt. De normstudent moet op een tentamen een ruime voldoende kunnen verwachten; het betreft hier enerzijds geen cum-laudestudent, maar anderzijds ook niet iemand die op een 6 werkt of niet intelligent genoeg is om met serieus werken meer dan een 6

te halen. Vervolgens is, behalve de verwachte score, de meetfout van belang waaraan die score onderhevig is, uitmondend in de kans om bij toeval toch een onvoldoende te behalen.

Ten eerste mogen dus tentamens niet te moeilijk of te gemakkelijk zijn. Het feit dat de modale student op vrijwel alle tentamens ver beneden de 7 scoort, is hier niet maatgevend, aangezien de modale student aanzienlijk minder dan de volle tijd aan de studie besteedt. De moeilijkheidsgraad van het tentamen komt ter discussie als normstudenten - dat wil zeggen: personen die zich in ernst als zodanig presenteren - systematisch lager dan een 7 blijken te scoren, in zodanige frequentie dat dit tot een 'collectieve klacht' zou kunnen leiden. In dat geval is de opleiding gehouden het niveau van het tentamen aan te passen. Dat geldt overigens ook, maar dan in omgekeerde richting, als het tentamen voor de normstudent te gemakkelijk is. In de praktijk fluctueert uiteraard de moeilijkheid van vakken en van tentamens, binnen zekere tolerantiegrenzen.

Uit het hier ingenomen gezichtspunt vloeit vervolgens de eis voort dat de normstudent (met een verwachte score  $T$ , bijvoorbeeld  $T = 7$ ) slechts bij hoge uitzondering door meetfouten een onvoldoende ( $<5.5$ ) haalt. Zekerheid kan de student niet worden geboden, maar de kans op een onvoldoende mag niet hoger zijn dan  $p$  (bijvoorbeeld  $p = .05$ ). De argumentatie voor  $p = .05$  luidt:

- eens in de twintig keer een herhaaltentamen te moeten doen, is acceptabel;
- de kans dat de normstudent ook voor de herhaling zakt, valt onder de toevalsassumptie te schatten als  $p^2 = .0025$ , en is dus zeer klein;
- de kans dat zo'n calamiteit zich in de hele studie *niet* voordoet, dus bij geen van de  $n$  tentamens die de studie telt, valt te schatten als  $(1 - p^2)^n$ ; bij  $n = 40$  is dat ongeveer 90%, dus slechts een op de tien normstudenten loopt een keer tegen die calamiteit aan;
- de enkele normstudent die dat overkomt, kan zonder bezwaar op individuele basis een extra kans krijgen.

Het bovenstaande houdt meteen in dat er geen goede reden is om meer dan twee reguliere tentamenkansen te bieden. Die service dient louter diegenen tegenover wie de opleiding onderwijskundig gesproken geen verplichtingen heeft; bovendien gaan er verkeerde signalen van uit.

Onder de assumptie van normaal verdeelde meetfouten rond  $T = 7$  mag op basis van het voorafgaande de standaardmeetfout van de cijfers maximaal  $(T - 5.5)/z$  zijn, waarbij  $z$  de standaardscore is, die overeenkomt met de overschrijdingskans  $p$ . Bij  $p = .05$  is  $z = 1.65$  en de standaardmeetfout van de cijfers .909.

De standaardmeetfout  $SE$  van de ruwe tentamenscore is, volgens een klassieke benaderingsformule van Lord (1959), luidende  $SE = c/m$ , louter een functie van het aantal items  $m$  in de toets; de empirische constante  $c = .412$  die hier wordt gebruikt, is afkomstig uit ongepubliceerd onderzoek van de tweede auteur dezes met een groot aantal plaatselijke psycholobietentamens. De gehanteerde waarde is exemplarisch: het materiaal waarop de waarde is gebaseerd, is wellicht relevanter voor de gebruiker dan Lord's materiaal, maar het verdient aanbeveling de waarde naar bevinden aan te passen aan de eigen situatie (zie ook verderop).

De standaardmeetfout van de cijfers is mede een functie van het aantal antwoordalternatieven  $A$ ; dat hangt samen met de rol die de correctie voor gissen speelt bij de berekening van de cijfers. In het algemeen wordt voor  $m$  goede antwoorden het cijfer 10 toegekend, en voor de kansscore  $m/A$  het cijfer 0. Bij overbrenging van de ruwe scores, met een bereik van  $m - m/A$ , naar de schoolcijferschaal, wordt de standaardmeetfout van de cijfers derhalve verkregen door de standaardmeetfout van de ruwe score te vermenigvuldigen met  $10/(m - m/A)$ .

Uit het bovenstaande valt de algemene formule af te leiden voor het vereiste aantal items in een toets:

$$m \geq [10cz/(T - 5.5)]^2 [A/(A - 1)]^2;$$

bij  $c = .412$ ,  $z = 1.65$  en  $T = 7$  resulteert:

$$m \geq 20.54[A/(A - 1)]^2.$$

Bij een binair gescoorde open-antwoordtoets, waarbij  $A$  op oneindig kan worden gesteld, zijn er dus minstens 21 vragen nodig. Voor een toets bestaande uit tweekeuzevragen ligt het benodigde aantal items 4 keer zo hoog (83 items). Voor een driekeuzetoets is die vermenigvuldigingsfactor 2.25 (47 items), voor een vierkeuzetoets 1.78 (37 items).

### Kanttekeningen

In de bovenstaande redenering zijn enkele factoren buiten beschouwing gelaten. In de eerste plaats zijn open vragen onderhevig aan beoordelaarsfouten. Daar staat tegenover dat ze vaak op een fijnere schaal worden gescoord dan alleen goed-fout, hetgeen een bescheiden winst zal opleveren. Bij meerkeuze-items doet zich het probleem voor dat sommige antwoordalternatieven zo verwerpelijk zijn dat het item feitelijk minder dan  $A$  antwoordalternatieven heeft. Daar staat weer tegenover dat het bij voorbaat verwerpen van zulke alternatieven getuigt van elementaire ('partiële') kennis. Er zijn al met al geen doorslaggevende redenen bij voorbaat om de algemene formule te specificeren voor elke  $A$  afzonderlijk; empirisch onderzoek kan echter uitwijzen dat die noodzaak er wel is.

De besliskundig geïoriënteerde lezer zal wellicht geneigd zijn de uitkomsten te variëren naar de omvang van het tentamen: immers, een onterechte (on)voldoende op een 6-puntsvak is ernstiger dan eentje op een 1-puntsvak. Inderdaad hebben veel docenten ook spontaan de neiging een toets langer te maken naarmate er meer studiepunten mee gemoed zijn. Die neiging kan in het model expliciet worden verdisconteerd door  $p$ , en daarmee  $z$ , in functie te stellen van de omvang van het vak. Men moet er dan op bedacht zijn dat de vereiste toetslengte een kwadratische functie is van  $z$ . Onderwijskundig gesproken echter verdient het wellicht voorkeur de omvang van het vak binnen bepaalde marges te standaardiseren, zodat het probleem zich niet voordoet.

Uit de algemene formule voor de vereiste toetslengte valt af te lezen dat het vrijwel ondoenlijk is studeerbaarheidsgaranties te geven aan de marginale student (gedefinieerd aan een verwachte score  $T = 6$ ) in de zin dat pech bij een tentamen binnen de perken kan worden gehouden: daartoe zouden de vereiste toetslengtes met 9 moeten worden vermenigvuldigd. Uitgaande van de hier aanbevolen toetslengtes is de kans op een onvoldoende bij  $T = 6$  ruim 29% (corresponderend met  $z = .55$ ). Als we er verder van uitgaan dat in de praktijk de doorsnee student een marginale student is, dan geeft de analyse aanleiding tot gedachten zoals:

- als men - wat hier niet wordt bepleit - met relatieve cesuren zou willen werken, dan zou onder de omstandigheden het percentage onvoldoenden op 30% moeten worden ingesteld, in plaats van de veel lagere waarden die de ronde doen, tenzij er aanwijzingen zouden zijn dat er duidelijk meer norm- en cum-laudestudenten ( $T \geq 7$ ) dan submarginale ( $T \leq 5$ ) studenten zouden voorkomen;
- de kans gegeven  $T = 6$  om ook bij herhaling te zakken is minder dan 9%, zodat de uitnodiging tot marginaal studeren blijft bestaan, zulks tot geruststelling van velen;
- de kans gegeven  $T = 5$  om te *slagen* is natuurlijk ook ruim 29%; bij herhaling is dat  $1 - (.71)^2 = 50\%$ , en zelfs met twee herhaalgelegenheden is de slaagkans nog maar 64%; dat is weliswaar niet niks, maar redelijkerwijs niet genoeg om tot submarginaal studeren uit te nodigen.

Merk tot slot op dat de betrouwbaarheid van de toets, in de vorm van coëfficiënt alpha (c.q. Kuder-Richardson 20), *als zodanig geen enkele rol speelt* in de beoordeling van de kwaliteit van de toets, in de verklaring van het tentamengedrag van studenten, en in de formulering van beleidsoverwegingen zoals hier kort aangeduid. De toetsbetrouwbaarheid fungeert integendeel louter als afleider in het denken over tentamens. In plaats van alpha-coëfficiënten routinematig mee te leveren, zouden de betreffende analyseprogramma's de feitelijke standaardmeetfout van het cijfer moeten berekenen en afdrukken. Op basis van die feedback kan de  $c$ -parameter in het model, en daarmee de vereiste toetslengte, lokaal worden aangepast.

## Verwijzing

Lord, F.M. (1959). Tests of the same length do have the same standard error of measurement. *Educational and Psychological Measurement*, 19, 233-239.

Correspondentie: Willem K.B. Hofstee, Heymans Instituut,  
Grote Kruisstraat 2/I, 9712 TS Groningen  
Tel. 050 363 6340, w.k.b.hofstee@ppsw.rug.nl

## Overpeinzingen bij een fusie

Van de redactie

Dit is het laatste nummer van het Tijdschrift voor Onderwijsresearch, beter bekend onder het acroniem TOR, als zelfstandig tijdschrift. Fusiebesprekingen met het tijdschrift Pedagogische Studiën zijn afgerond en met gezamenlijke inspanning wordt de taak voortgezet. Het opgeven van de zelfstandigheid betekent desalniettemin het afsluiten van een markante periode en is aanleiding voor de redactie tot enkele overpeinzingen.

De institutionalisering van een wetenschappelijke discipline gaat gepaard met het instellen van leerstoelen, het oprichten van tijdschriften, het organiseren van conferenties en, in Nederland, door het Ministerie van Onderwijs erkend worden als een universitaire opleiding. In de jaren zeventig stond het onderwijs in Nederland in het brandpunt van de publieke belangstelling. Als gevolg van de grote toestroom van leerlingen en studenten groeide de sector explosief. Het ministerschap van Onderwijs en het voorzitterschap van de vaste Kamercommissie voor Onderwijs behoorden tot de meest geliefde politieke functies in Nederland. De verbeelding raakte aan de macht en het onderwijs was in veler ogen het voertuig om de gedroomde toekomst vorm te geven. Binnen diverse universiteiten ontstonden interdisciplinaire opleidingen Onderwijskunde, meestal in een soort haat-liefde verhouding levend met disciplinaire benaderingen zoals de onderwijspsychologie en de onderwijssociologie binnen dezelfde universitaire instelling. Jaarlijks werden vanaf die tijd ook de Onderwijsresearch-dagen georganiseerd en de Stichting voor Onderzoek van het Onderwijs (SVO) kwam tot wasdom.

Het Tijdschrift voor Onderwijsresearch, opgericht in 1975, is het prototype van een tijdschrift dat past binnen een dergelijk institutionaliseringsproces. Het kwam voort uit twee bronnen: de Memo-reeks Onderwijsresearch gefinancierd door de SVO, en de didactometrische nummers van het Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden. De betrokken onderzoekers bundelden hun krachten en besloten een volwaardig tijdschrift voor onderwijsonderzoek op te richten. Het zou bovendien vanaf het begin het huisorgaan van de Vereniging voor Onderwijsresearch (VOR) moeten worden, een vereniging die ook in die tijd het levenslicht aanschouwde.

De emancipatie van de onderwijskunde als wetenschappelijke discipline ging tevens gepaard met een duidelijke positiebepaling in wetenschapstheoretische zin: empirisch-analytisch. Daarmee zich distantiërend van de meer beschouwelijke aanpak die tot die tijd de analyses van het onderwijs domineerde, dan wel daar een voorname plaats innam. Een tijdschrift dat bij die emancipatie paste, diende deze visie ook uit te stralen. In de eerste tien jaar van haar bestaan kenmerkten de opgenomen artikelen zich dan ook door een sterk psychometrisch gehalte waarbij de criteria van deskundigen uit de wereld van de toets- en testconstructie een belangrijke rol speelden. Van de 40 geplaatste artikelen in de eerste twee jaargangen was dan ook 57% gewijd aan methodologische en statistische probleemstellingen, terwijl slechts 28% betrekking had op empirisch onderzoek van het onderwijs.

Het volume aan eerste en derde geldstroomformatie voor onderwijsonderzoek nam in de loop van de jaren zeventig en tachtig zeer sterk toe en tal van probleemstellingen uit nieuwe aandachtsgebieden werden voorwerp van onderzoek. De Vereniging voor Onderwijsresearch breidde haar divisies uit en ook deze nieuwe takken van sport wilden zich uiteraard graag herkennen in bijdragen aan TOR. De inhoud van de geplaatste artikelen correspondeerde daar niet mee en het tijdschrift begon daardoor in het midden van de jaren tachtig aan legitimering binnen de gemeenschap van onderwijsonderzoekers in te boeten. De redactie besloot haar beleid bij te stellen. Het palet van inhoudelijke onderwerpen moest worden verbreed door



onder andere het uitbrengen van themanummers en ook reviewartikelen zouden meer aandacht moeten krijgen. Voorts werd voor de rubriek Notities en Commentaren een prominente plek ingeruimd. Het redactiebeleid met betrekking tot de strenge criteria omtrent bewijsvoering bleef echter overeind. Weliswaar hoefde niet meer in elk artikel ieder gebruikt instrument aan de meest geavanceerde psychometrische eisen te voldoen maar methodisch-technische maatstaven bleven in de beoordeling een dominante rol spelen. Dat heeft ongetwijfeld geleid tot zelfselectie bij potentiële inzenders van artikelen: de toevoer van kopij uit bepaalde divisies van de VOR bleef gering. Ook het schrijven van reviewartikelen voor TOR bleek geen geliefde bezigheid van beoogde auteurs. Daarentegen sloeg het beleid tot het uitbrengen van meer themanummers aan.

Midden jaren tachtig begon het streven naar de internationalisering van het Nederlandse onderwijsonderzoek voorzichtig op gang te komen om vervolgens in de jaren negentig fors door te zetten. De onderwerpen die van oudsher het hart uitmaakten van de kopijstroom van TOR vonden ook het eerst hun weg naar de internationale tijdschriften waardoor een belangrijke bron van kopij begon op te drogen. Toen vervolgens het ICO, de onderzoeksschool voor onderwijskundigen, halverwege de jaren negentig de toelatingscriteria tot de school (bijna) uitsluitend wenste te baseren op bijdragen aan internationale tijdschriften, daalde de aantrekkelijkheid van TOR voor een aantal vooraanstaande auteurs aanzienlijk. Dat had uiteraard consequenties voor de hoeveelheid ingestuurde artikelen.

Tal van pogingen zijn in het werk gesteld om daarin verandering aan te brengen. Onderzoekers die op de Onderwijsresearchdagen interessante bijdragen hadden gepresenteerd, werden benaderd met een verzoek om deze te bewerken tot een artikel, redactieleden werden aangespoord om hun netwerk te activeren, onderzoekers werden uitgenodigd om een themanummer samen te stellen, enz. Dat alles leverde niet het beoogde resultaat op. De redactie bezon zich vervolgens op andere formules: van 'going international' dan wel prijsgeven van het predikaat 'wetenschappelijk' tot doorgaan als verenigingsblad met voor de beroepsgroep interessante bijdragen. Deze alternatieven zijn door middel van een schriftelijke enquête voorgesteld aan de VOR-leden. Voor geen der alternatieven was een uitgesproken voorkeur waardoor impliciet een besluit viel over het zelfstandig voortbestaan van TOR. Fusering met het tijdschrift *Pedagogische Studiën* (PS) bleek toen een aantrekkelijk alternatief. Dit tijdschrift vertoont namelijk een sterke overlap in type bijdragen, zeker na de 'verbreding' in de jaren tachtig van TOR. Zo was in de jaargangen 1997 en 1998 van TOR nog slechts 11% van de artikelen gewijd aan statistische en methodologische problemen en maar liefst 68% aan empirische studies over het onderwijs of de maatschappelijke invloeden daarop. Een percentage dat vergelijkbaar zal zijn met dat van PS.

Een belangrijke rol speelde ook het gegeven dat de uitgever van PS het tijdschrift niet langer koste wat kost in haar fonds wilde handhaven. Daarmee was de weg vrij voor het openen van fusiegesprekken. Dat resulteert in een samengaan per 1 januari 2001 waarbij het nieuwe PS tevens zal functioneren als verenigingsorgaan van de VOR. De naam *Pedagogische Studiën* heeft gezien traditie (78 jaar!), lezersbestand en strategische redenen de voorkeur gekregen boven andere. Delen van de redacties van beide voormalige tijdschriften zetten gezamenlijk het werk voort. Door het overnemen van de functie van verenigingsorgaan en het nieuwe redactiebeleid, dat overigens nauwelijks verschilt van dat van TOR, wordt de erfenis van TOR goed beheerd. Dat neemt niet weg dat de oude redactie van TOR met spijt afscheid neemt van het haar zo geliefde tijdschrift. En wel om een aantal redenen. Zo leefden de opeenvolgende redacties steeds in de veronderstelling dat ze met haar beleid een stevige bijdrage leverde aan het handhaven van de normen van goed wetenschappelijk onderwijsonderzoek, in het bijzonder in het kader van de ontwikkeling van de onderwijskunde als een empirisch-analytische discipline. Voorts zijn er in TOR tal van artikelen verschenen die een interessante bijdrage hebben geleverd aan de inhoudelijke ontwikkeling van het vak en, ten slotte, de inhoudelijke discussies binnen de redactie waren voor veel redactieleden zeer vormend van aard: verbreding en verdieping van reeds aanwezige kennis en inzicht. Het rekruteren van gekwalificeerde redactieleden leverde dan ook nooit een probleem op.

De inhoud van het laatste nummer van een tijdschrift dient een geheel eigen karakter te hebben. De redactie heeft daarom gemeend dit te moeten vullen met de herdrukken van de beste artikelen uit haar bestaan. Daarbij past een aantekening. Aan de divisiecoördinatoren van de VOR is gevraagd het beste artikel uit hun divisie te nomineren. De redactie heeft gemeend de criteria voor selectie aan de coördinatoren te moeten overlaten. Een aantal divisies heeft schriftelijk verslag uitgebracht over de gevolgde selectieprocedure. Daaruit bleek dat deze zeer zorgvuldig te werk zijn gegaan.

Opgemerkt is al dat niet voor alle divisies TOR het meest representatieve platform was en is. Met als gevolg dat sommige coördinatoren aanvankelijk aarzelden op het verzoek in te gaan omdat volgens hen in de afgelopen 25 jaar in andere tijdschriften kwalitatief veel betere artikelen over hun onderzoeksveld zijn verschenen. Ten onrechte zou de indruk kunnen ontstaan dat het gekozen artikel een goede indicatie zou zijn voor de wetenschappelijke prestaties op dit terrein. Na toezegging tot opname van deze relativering in het voorwoord bij dit nummer hebben de coördinatoren hun keus gemaakt. Bijgevolg kan ook niet worden gezegd dat het totaal aan geselecteerde artikelen in absolute zin het beste is dat in dit tijdschrift is verschenen: er is geselecteerd op basis van een gestratificeerde steekproef.

Tot slot wenst de redactie enkele personen en instellingen in het bijzonder te bedanken. Allereerst Bert Creemers en Rinus Voeten. Zij waren tot het eind van de jaren tachtig TOR. Zij hebben het blad opgericht en zijn jarenlang als secretaris en voorzitter de spil van de onderneming geweest. Dat Bert mede door het oprichten van tal van internationale tijdschriften mede heeft bijgedragen aan de uiteindelijke opheffing van TOR valt hem niet kwalijk te nemen. Vervolgens bedanken we Swets en Zeitlinger en Giethoorn ten Brink die al die jaren gezorgd hebben voor de uitgave en verspreiding van het tijdschrift. In ruimere zin gaat onze dank voorts uit naar de tientallen onderzoekers die in de loop der jaren deel hebben uitgemaakt van de redactie, zij worden bedankt voor hun inzet. De meeste dank gaat echter uit naar het legioen van beoordelaars. Zoals bekend, werd elke inzending anoniem door twee beoordelaars gerevied alvorens het ter bespreking in de redactievergadering kwam. In veel gevallen dienden de herschreven stukken door dezelfde beoordelaars opnieuw bekeken te worden. De uren werk die daarmee gemoeid zijn en die, vallend onder de anonimiteitsclausule, voor betrokkenen prestige noch materiële vergoeding opleveren, dwingen respect af. Dat zoveel onderzoekers het tot hun taak rekenen bij te dragen aan een goed functioneren van het wetenschappelijk forum, ondanks de geringe incentives, heeft de redactie van TOR altijd zeer op prijs gesteld.

Tot slot. Het einde van TOR als zelfstandig tijdschrift kan worden betreurd. Vanuit het perspectief van een zich ontwikkelende discipline kan het ook als een teken van vooruitgang worden gezien. Is een startende discipline vooral gericht op het forum van lokale wetenschappers en vinden publicaties plaats in de lokale taal, op den duur zal bij succes, het internationale forum de doelgroep zijn waar de resultaten van onderzoek beoordeeld worden. Deze ontwikkelingsgang hebben zeer veel disciplines in het verleden doorgemaakt met het gevolg dat publiceren daar in internationale media de norm is geworden. Hetgeen onverlet laat dat met name de verspreiding van kennis onder beoepsbeoefenaren, beleidsmakers en politici voor met name de 'kundes' van levensbelang is. Daarom een fusie en geen opheffing.

# Ontwikkelingen in het Nederlandse Onderwijsonderzoek

Bert P.M. Creemers

## ABSTRACT

*After a short review of descriptions of the development of educational research in the past, the period of the Dutch Tijdschrift voor Onderwijs Research (Journal for Educational Research) running from 1975-2000 is brought into focus. During this time, educational research in the Netherlands came to full development, supported by extra funding of educational research. Gradually from 1985, while the financial resources declined, due to the pressure on productivity within universities, the output in terms of Ph.D.s and especially international publications increased. This brings about the present situation in which the Dutch Journal will cease to exist as a separate publication. Important areas of research during this period were learning and instruction and the design and organization of education. It is argued that the future programming of research should emphasize the scientific input and output in terms of validated knowledge. The local level of the research institute is an adequate place to start the (scientific) programming. The external relevance however is initiated at a more global- (international-) level and can be intertwined with the scientific programming at a local level according to local needs and translation of more global and national priorities.*

## INLEIDING

Het lijkt een kenmerk te zijn van jonge wetenschap dat met grote regelmaat een bestandsopname plaatsvindt waarbij dan, uiteraard, veel aandacht wordt besteed aan de problemen en de gewenste ontwikkelingen. Over de onderwijskunde in het algemeen, en onderwijsonderzoek in het bijzonder, zijn in het verleden herhaaldelijk dergelijke bijdragen verschenen (Van Trotsenburg, 1972; De Groot, 1976; Creemers, 1982). In een eerdere publicatie heb ik dan ook de opmerking gemaakt dat dergelijke overzichten niet zozeer een bewijs vormen dat de betreffende wetenschap 'erbij hoort' maar juist van het feit dat de wetenschap nog in ontwikkeling is en nog een plaats te midden van andere wetenschappen moet zien te verwerven. Wanneer het aantal overzichten de norm is voor de kwaliteit ervan zou het wel goed zitten met het onderwijsonderzoek, want behalve de regelmatig verschijnende programmeringsstudies op dit terrein zijn er in het afgelopen decennium geen overzichten van de ontwikkeling van onderwijsonderzoek meer gepubliceerd. Ook zou het hier volgende overzicht niet verschenen zijn ware het niet dat er actuele aanleiding is om terug te blikken, namelijk het beëindigen van het *Tijdschrift voor Onderwijsresearch* (TOR) als zelfstandige uitgave. De vraag die daarbij opkomt is of dat een teken van bloeiende ontwikkeling is van het onderwijsonderzoek of juist een teken van verval. Dit was wel niet de vraag van de redactie van het TOR bij dit laatste nummer, maar het is wel een vraag die mij van tijd tot tijd bezighield toen ik met het bericht werd geconfronteerd. De initiatiefnemers, die ook betrokken waren bij de start van de Onderwijsresearch Dagen en de oprichting van de Vereniging voor Onderwijsresearch, hadden indertijd de bedoeling bij te dragen aan de inhoudelijke ontwikkeling en de

organisatie van het Nederlandse onderwijsonderzoek. Ze kwamen voort uit instituten die onderwijsonderzoek uitvoerden voornamelijk gefinancierd door de (voormalige) Stichting voor Onderzoek van het Onderwijs (SVO) en verder uit universitaire eenheden binnen vooral Pedagogiek en Psychologie die zich met onderzoek bezighielden. De doelstelling bij de oprichting van de Vereniging voor Onderwijs Research, het tijdschrift en de onderwijsresearchdagen was de bevordering van het onderwijsonderzoek. Onderwijsonderzoek zou daardoor bijdragen aan het valideren van theorieën en de ontwikkeling van een kennisbestand en het oplossen van problemen in het onderwijs. In concreto ging het daarbij om onderwerpen die betrekking hadden op de kwaliteit van het onderzoek, zoals de methoden en technieken, het interdisciplinaire karakter en de ontwikkeling van programma's en onderzoeksvoorstellen. Daarnaast waren de activiteiten erop gericht het onderwijsonderzoek en de programmering ervan te plaatsen in de maatschappelijk politieke context, zoals de bijdrage aan het onderwijsbeleid en de onderwijspraktijk, dit laatste onder meer in relatie tot de bijdrage van de andere instanties in de onderwijsverzorging, zoals de begeleiding en advisering, curriculumontwikkeling en toetsing. Ten slotte waren ook het belang van de beoefenaars zoals de bevordering van de professionele ontwikkeling en de condities voor onderzoek dat door SVO en het ministerie werd aanbesteed, een belangrijk aandachtspunt.

Vijfentwintig jaar later kan de vraag gesteld worden of dit geluk is en wat het *Tijdschrift voor Onderwijsresearch* in dit verband heeft betekend. Om deze vraag te beantwoorden zal in het volgende eerst aandacht worden besteed aan een aantal momentopnamen die in het verleden zijn gemaakt van het onderwijsonderzoek en de ontwikkeling van de onderwijskunde. Vervolgens zal een schets worden gegeven van de ontwikkeling van de afgelopen 25 jaar. Ten slotte zal in een aantal paragrafen worden nagegaan welke problemen er eventueel nog resteren voor het post-TOR tijdperk.

## MOMENTOPNAMES IN HET VERLEDEN

In het verleden zijn er reeds diverse beschrijvingen van de ontwikkelingen in de onderwijskunde en het onderwijsonderzoek verschenen. Creemers en Hoebe (1984) geven een overzicht en analyse van de beschrijvingen die op dat moment waren verschenen en voegen daar een eigen perspectief op de ontwikkelingsgeschiedenis van onderwijsonderzoek aan toe.

Voor Deen (1969) is de ontwikkelingsgeschiedenis nauw verbonden met de geschiedenis van een enkel instituut, namelijk het Nuts-Seminarium. Van Trotsenburg (1972, pp. 103-113) beschrijft de geschiedenis van de onderwijskunde en het onderwijsonderzoek als de ontwikkeling van methoden en technieken binnen de onderwijskunde en het losmakingproces van de pedagogiek en didactiek. Zijn beschrijving is niet uitsluitend op de Nederlandse situatie gericht, maar hij schetst in grote lijnen ook de ontwikkelingen in andere landen zoals Amerika, Engeland, België en Zweden.

De Groot (1976) heeft de ontwikkeling van het Nederlandse onderwijsonderzoek geschetst bij de oprichting van de Vereniging voor Onderwijsresearch. Hij neemt als invalshoek de theorievorming en het onderzoek met betrekking tot het onderwijzen en de organisatorische verbanden die daarvoor werden opgericht. Daarom zijn in zijn perceptie belangrijke markeringen in de ontwikkeling de geschiedenis van het R.I.T.P., de oprichting van het CITO, van de SVO en uiteindelijk, nog niet door hem genoemd, van de SLO (Stichting voor Leerplan Ontwikkeling). Volgens De Groot liggen de wortels van het onderwijsonderzoek in de constructie van toetsen voor het onderwijs en in de evaluatie van het onderwijs. Deze aandachtspunten waren ook gebaseerd op psychologische ervaring en kennis op het terrein van testconstructies. Hierdoor ontstond psychometrische belangstelling voor andere aspecten van het onderwijs zoals de doelstellingenproblematiek en uiteindelijk ook de leerplanevaluatie. Van Gelder (1970) heeft zich niet aan geschiedschrijving van de onderwijskunde en het onderwijsonderzoek als zodanig gewaagd. Maar uit de continue stroom van publicaties tot en met



zijn afscheidsrede in Groningen toe, wordt duidelijk dat hij de ontwikkeling ziet vanuit de bijdrage die onderwijskunde heeft geleverd, zou kunnen leveren of zou moeten leveren aan een vernieuwing van het onderwijs die vooral tot uitdrukking zou moeten komen in het scheppen van meer kansen voor de zwakkeren. Zo hield hij wel een pleidooi voor de ontwikkeling van empirische onderwijskunde en van empirisch onderzoek binnen de onderwijskunde, maar hij bleef dit vooral beoordelen vanuit de bijdrage die werd geleverd aan de vernieuwing van het onderwijs in de vorm van ontwikkeling, begeleiding en het verschaffen van inzichten waarvan bij de vernieuwing gebruik zou kunnen worden gemaakt. Het hield onder meer in dat hij de theorievorming en de beproeving van theorieën zelf als een bijproduct van onderwijsontwikkeling zag. Zoals Postma en Wardekker (1981) hebben aangegeven is zijn visie op het onderwijsexperiment een pleidooi voor de combinatie van onderwijsvernieuwing en de ontwikkeling van onderwijs door middel van theorievorming en beproeving van theorieën. De beproeving van theorieën is evenwel vaak een bijproduct.

Zijn de meeste van deze beschrijvingen gekleurd door een bepaalde disciplinaire of een methodologisch-technische voorkeur, zij benadrukken de empirische benaderingswijze en het interdisciplinaire karakter van de onderwijskunde. Men komt ook beschrijvingen en waarden van de ontwikkeling tegen vanuit een heel ander perspectief zoals van Crombag (in Vrijenhoef, 1982) die binnen de onderwijskunde een meer pedagogische en op normen en waarden georiënteerde benaderingswijze voorstaat.

Het is duidelijk dat dergelijke beschrijvingen gegeven worden vanuit een bepaald perspectief. De Groot neemt een leerspsychologische, in elk geval een psychologische invalshoek. Van Trotsenburg is voorstander van een empirische benadering binnen de pedagogiek en een sterk voorstander van de ontwikkeling van methoden en technieken. Van Gelder is voorstander van een maatschappelijke bijdrage van de wetenschap, Crombag is iemand die een aantal meer uit de geesteswetenschappelijke denkwereld stammende methoden en theoretische inzichten in het moderne wetenschapsbedrijf wil behouden. Een dergelijk perspectief leidt er dan toe dat bepaalde ontwikkelingen betreffende methoden en technieken, het belang van de wetenschap voor de onderwijspraktijk, de theorievorming en het beproeven van theorieën op een bepaalde specifieke wijze worden beoordeeld. Zo ziet De Groot (1982) een bijdrage aan de oplossing van problemen in het onderwijs vooral via verdergaande theorievorming en onderzoek. Voor Van Trotsenburg is de ontwikkeling van technieken een aanzet tot verdere ontwikkeling van het wetenschapsgebied en dan met name op het gebied van theorievorming. Voor Van Gelder is theorievorming een bijproduct van onderwijsontwikkeling. De consequentie hiervan is dat ook de thematisering, dat wil zeggen de aandachtsvelden in theorievorming en onderzoek, de objecten van de onderwijstheorie, vanuit diverse perspectieven heel anders beoordeeld worden. Voor sommigen zijn de leerprocessen en de vormgeving van de leerprocessen de voornaamste aandachtsvelden. Voor anderen is de essentie van onderwijstheorieën gelegen in de vormgeving van het onderwijs; zij stellen curriculumtheorieën, innovatietheorieën en daarbinnen weer de curriculumontwikkelingspraktijk (ontwikkeling, implementatie, evaluatie) en de innovatie van het onderwijs centraal.

Zo kan bijna iedereen zijn eigen geschiedenis van de onderwijskunde en het onderwijsonderzoek reconstrueren. Enige uitzonderingen daargelaten kan plausibiliteit aan deze reconstructies niet worden ontzegd; deze is in belangrijke mate afhankelijk van ieders perspectief op onderwijsonderzoek.

Creemers en Hoeben (1984) stellen dat er geleidelijk overeenstemming is ontstaan over een opvatting dat onderwijskunde zowel gericht is op theorievorming over als op bijdragen aan probleemoplossing in de onderwijspraktijk. De bijdragen aan een dergelijke onderwijskunde blijken steeds meer afkomstig te zijn uit verschillende disciplines zoals pedagogiek, psychologie en sociologie. De autonomie en de authenticiteit van de disciplines wordt in theorievorming over, in onderzoek van en in bijdragen aan de oplossing van onderwijsproblemen hoe langer hoe minder van belang. Dit komt met name voort uit de extra prikkel op interdisciplinaire samenwerking die is gelegen in de probleemgerichte vragen die vanuit onderwijs-



praktijk en onderwijsbeleid worden gesteld, en uit de bonus die op deze interdisciplinaire samenwerking wordt gezet in de vorm van extra middelen ten behoeve van onderwijsonderzoek, onderwijsontwikkeling en onderwijsbegeleiding. Meer duurzame vormen van samenwerking worden bereikt omdat de maatschappij vragen aan wetenschappers stelt die ons noodzaken, maar het ons tevens mogelijk maken, om langer met elkaar samen te werken. Echter, daarvoor is ook nodig dat deze meer langdurige vormen van samenwerking gepaard gaan met een grotere gerichtheid op het, weliswaar mede ten behoeve van de praktijk, bedrijven van wetenschap. In 1984 constateren Creemers en Hoeben dat de effecten in termen van een bijdrage aan theorievorming én aan de verbetering van de onderwijspraktijk nog niet zo groot zijn. Zij wijten dat aan de beoordelingscriteria en procedures van de Stichting voor Onderzoek van het Onderwijs (SVO) die meer recht zouden doen aan de wetenschappelijke eisen dan aan de verbetering van het onderwijs dat gelijktijdig in beschouwing zou moeten worden genomen.

Anderzijds constateren ze dat in de jaren zeventig een sterke druk op onderwijsonderzoek op gang is gekomen om een bijdrage te leveren aan de praktijk en vooral aan de onderwijsvernieuwing (en het beleid ten aanzien van de onderwijsvernieuwing). Dit heeft de aandacht voor theorievorming en beproeven van theorieën doen afnemen. De Verkenningcommissie Onderwijsresearch was in haar rapport in 1976 erg pessimistisch over de bruikbaarheid van de opbrengsten van onderwijsonderzoek voor de onderwijspraktijk en het onderwijsbeleid. Men kan zich afvragen of dit het gewenste effect – een grotere bijdrage aan het onderwijs – wel heeft gehad.

Uiteindelijk is een bijdrage aan de praktijkverbetering ook niet gediend met een onderaccentuering van een dergelijke kennisproductie. In de jaren tachtig zijn maatregelen om hierin verbetering te brengen ter hand genomen. Een verdergaande organisatie van programmeringswerkzaamheden en een doelmatige organisatie van onderzoeksuitvoering zouden de dubbele output van onderwijsonderzoek bevorderen, verwachtten Creemers en Hoeben (1984). Zij concentreren zich vervolgens op het samenspel tussen de interne op theorieën gerichte wetenschapontwikkeling en de externe beïnvloeding daarvan in de vorm van onderwijsontwikkeling, onderwijsverbetering en onderwijsinnovatie. De resultaten van dit samenspel zouden dan tot uitdrukking komen in de progressie in de probleemformuleringen voor onderzoek en onderzoeksprogramma's. De publicatie kan beschouwd worden als een hartstochtelijk pleidooi voor de verbetering van het samenspel tussen onderwijsontwikkeling, dat wil zeggen de maatschappelijke opbrengst van onderwijsonderzoek en de theoretische opbrengsten die tot uitdrukking komt in wetenschapontwikkeling.

Vanuit dit perspectief heeft Creemers in diverse publicaties (1982, 1983, 1985) een aantal thema's voor onderwijsonderzoek geformuleerd en de opbrengsten van het onderzoek daarbinnen geëvalueerd. Deze publicaties waren gebaseerd op de publicaties onder andere in het *Tijdschrift voor Onderwijsresearch* en gesprekken met collega's, hoogleraren en senior onderzoekers, en voorzien van een forse dosis eigenwijsheid. In deze publicaties komen een aantal gemeenschappelijke punten naar voren:

1. Terreinen van onderzoek die leiden tot wetenschaps- en onderzoeksontwikkeling of ertoe zouden kunnen leiden waren:

- het leren van de leerlingen
- de instructieprocessen
- de inhoudelijke en organisatorische vormgeving van onderwijs
- ontwikkeling van testen en evaluatiesystemen

Stagnatie werd geconstateerd op het terrein van onderwijsinnovatie, zeker na de grote aandacht ervoor tussen 1970 en 1980. Als nieuwe terreinen werden aangeduid de onderwijstechnologie, de onderwijs economie (inclusief onderwijsplanning gerelateerd aan arbeidsmarkt vraagstukken) en het nieuwe toegankelijkheidsprobleem van het onderwijs, de etnische minderheden.

2. Geconstateerd werd dat het onderwijsonderzoek – in eerste instantie beperkt tot dat op het terrein van de statistiek en methodeleer maar later ook op de andere hierboven genoemde terreinen – internationale erkenning kreeg of begon te krijgen. Dit ondanks de in de jaren tachtig krimpende universitaire formatie. Verdere internationalisering via publicaties en samenwerkingsverbanden werd voorzien en bepleit.
3. Het effect van pleidooien voor een samenspel tussen wetenschapsontwikkeling en onderwijsontwikkeling was betrekkelijk gering waarvoor organisatorische en inhoudelijke redenen te bedenken zijn. De interne op wetenschapsontwikkeling gerichte dynamiek zou nog weinig aansluiten op de externe dynamiek in onderwijspraktijk en -beleid. Daarnaast echter kan ook de versnippering in het onderwijsonderzoek – en daarbinnen de theorievorming – en het gefragmenteerde en globale karakter van onderwijsonderzoek en vooral evaluatieonderzoek ten behoeve van onderwijsontwikkeling hieraan debet zijn. Daarom wordt (bijvoorbeeld in Creemers, 1985) gepleit voor de ontwikkeling van een conceptueel omvattend raamwerk voor theorieontwikkeling en praktijkevaluatie. Als elementen voor dat raamwerk werden genoemd – natuurlijk – het leren en de leeropbrengsten gerelateerd aan de condities en processen.

### ONTWIKKELINGEN IN HET ONDERWIJSONDERZOEK IN DE PERIODE 1975-2000

De Groot (1976) verdeelde voor zijn beschrijving van de ontwikkeling van de onderwijsresearch de periode van 1950 tot 1975 in eenheden van vijf jaar, en ordende de beschrijving in de rubrieken klimaat, activiteiten en problemen. Deze beschrijving was een van de eerste publicaties in het *Tijdschrift voor Onderwijsresearch*. Weer 25 jaar later is het doek gevallen voor TOR. In deze paragraaf wordt aan de hand van De Groot's rubrieken een schets van deze periode gegeven.

De tien jaar na 1975 zijn in de vorige paragraaf al gedeeltelijk meegenomen aan de hand van de eerdere beschrijvingen van de ontwikkeling van het onderwijsonderzoek. Ik zal het accent leggen op 1985-2000.

#### **Klimaat**

In de periode 1975-1985 was er een positief klimaat voor het onderwijsonderzoek. In onderwijsbeleid en -praktijk – zij het daar met meer scepsis – werd verwacht dat onderwijsonderzoek een bijdrage zou kunnen leveren aan het oplossen van problemen in het onderwijs, en met name aan de vernieuwingen en veranderingen en ontwikkelingen die in het onderwijs plaatsvonden. Kleine strubbelingen tussen SVO en het onderwijsbeleid en tussen SVO en de instituten voor onderwijsonderzoek daargelaten, was er een ruimhartige financiering van het onderwijsonderzoek, vanuit de overtuiging dat hiermee in combinatie met de andere instellingen binnen de onderwijsverzorging op het terrein van begeleiding, curriculumontwikkeling en toetsconstructie, een geconcentreerde ondersteuning van vernieuwing en verzorging van het onderwijs gerealiseerd zou kunnen worden. In het onderwijsonderzoek en de onderwijskunde zelf waren vragen aan de orde rondom het wetenschappelijke gehalte en de relevantie van onderzoeksvraagstellingen die vanuit de onderwijspraktijk werden aangereikt.

In de tweede periode van 1985-1995 concentreert het onderwijsonderzoek zich, mogelijk in relatie tot ontwikkelingen binnen de universiteiten, op de meer wetenschappelijke ontwikkelingen en op de versterking van de onderzoeksoutput op het wetenschappelijke terrein. Dit was ook gedeeltelijk een reactie op de ervaren beperkingen van het onderwijsonderzoek, niet alleen binnen de praktijk maar ook in het onderwijsonderzoek zelf, om aan praktische problemen in de onderwijsinnovatie direct een bijdrage te kunnen leveren. De wisselwerking tussen de meer wetenschappelijke oriëntatie van het onderwijsonderzoek en de bijdrage aan de ont-

wikkeling van de onderwijspraktijk had minder aandacht dan voorheen en dit resulteerde in een geleidelijke verwijdering van onderwijsonderzoek enerzijds en onderwijspraktijk (en beleid) anderzijds.

Dit vertaalde zich ook in een geleidelijke vermindering van het budget voor onderwijsonderzoek en de oprichting met de steun van SVO van SARDES, een organisatie waarin meer op praktijk gerichte (onderzoeks)vraagstellingen centraal stonden. In feite creëerde SVO daarmee misschien niet zozeer zijn eigen concurrent maar wel een voor die onderzoeksinstituten en wetenschappers die de wisselwerking tussen theorieontwikkeling en onderwijsontwikkeling nastreefden. Binnen het wetenschapsbeleid en het onderwijsbeleid kwam bovendien de vraag naar voren of het voor de concretisering van de bijdrage aan onderwijsontwikkeling niet beter zou zijn dat het onderwijsonderzoek zich zou richten op maatschappelijk-politieke vraagstellingen in het onderwijs die langdurig in de aandacht zouden blijven staan of in de toekomst misschien van belang zouden worden, dan zoals voorheen op een bijdrage aan het oplossen van actuele problemen in de onderwijspraktijk en onderwijsbeleid. Misschien door de vermindering van het volume via SVO gefinancierd onderzoek maar ook door de accentuering van de onderzoekstaken binnen de universitaire onderwijskunde-opleidingen kwam er meer aandacht voor het theoriegerichte onderzoek binnen de eerste en tweede geldstroom. Dit werd nog verstrekt door de formulering van outputcriteria, de onderzoeksprogrammering en evaluatie van onderzoeksprogramma's die verderop in deze paragraaf aan de orde komen.

In de derde periode, vanaf 1995, kregen deze ontwikkelingen hun beslag in de onderbrenging van het overgrote deel van het onderwijsonderzoek binnen NWO waarbij in feite het accent op wetenschapontwikkeling gelegd werd. De directe lijn met onderwijsontwikkeling werd, als uitdaging of uiteindelijk ideaal, losgelaten, evenals een verbinding tussen programmering vanuit de wetenschap met die vanuit de onderwijspraktijk. Verwacht werd dat een sterkere koppeling aan de scholingsgebieden zoals wiskunde en aan onderwijsonderzoek uit andere disciplines tot een vergroting van het empirisch kennisbestand zou kunnen leiden. De dynamiek tussen maatschappelijke en theoretische gezichtspunten in opzet en uitvoering van een programma voor onderwijsonderzoek is daarmee verdwenen. Het directe aan het op onderwijspraktijk en -beleid gerichte onderzoek zou op een andere wijze gefinancierd, geprogrammeerd of georganiseerd worden. Een klein deel van het oorspronkelijke SVO-budget werd voor actuele onderwijsvraagstellingen afgezonderd. De LPC's werd verzocht om dit te programmeren en zij kozen voor een model dat tot programmering voor meerdere jaren en tot afspraken met onderzoeksinstituten zou kunnen leiden. Het onderwijsbeleidsonderzoek werd of vanuit het ministerie zelf georganiseerd of vanuit het ministerie via NWO wel geprogrammeerd zoals in het BOPO voor het beleidsonderzoek in het primaire onderwijs maar niet gerelateerd aan het meer theoretisch gerichte programma voor onderwijsonderzoek dat in de PROO plaatsvindt.

Ondanks dat er een bijdrage van onderwijsonderzoek aan theorievorming én aan onderwijsontwikkeling mogelijk wordt geacht en in onderzoeksprojecten nog mogelijk is, is de hoofdstroom van het onderwijsonderzoek door deze organisatie en financiering meer wetenschappelijk georiënteerd. De consequenties voor het imago van onderwijsonderzoek in onderwijspraktijk en onderwijsbeleid laten zich raden. Onderwijsonderzoek wordt van weinig belang geacht voor de oplossing van onderwijsbeleids- en praktijkproblemen. Merkwaardig is wel dat terzelfder tijd individuele onderwijsonderzoekers met grote regelmaat worden ingeschakeld bij onderwijsbeleidsproblemen en nadere concretisering geven aan voorstellen op het terrein van het onderwijs of het nu gaat over klassenverkleining, integraal schooltoezicht, grote innovatieprojecten zoals de basisvorming en het studiehuis of 'kleine' zoals de leerlinggebonden financiering. Individuele onderwijskundigen en onderwijsonderzoekers spelen een 'rol' van betekenis. Het is echter de vraag of het onderwijsonderzoek dat ook doet.

Van een wetenschappelijk tijdschrift, wat TOR beoogde te zijn, ook al was het tevens het blad van de VOR, kan niet een sturende of zelfs actieve bijdrage aan de wijzigingen in het onderwijsklimaat – en de activiteiten – verwacht worden. TOR geeft vooral een weerslag

ervan in de vorm van de opgenomen artikelen. In de inhoudsopgave is de verschuiving van psychometrische naar inhoudelijke bijdragen waar te nemen en van meer praktische naar meer theoretische vraagstellingen. In de rubriek 'Notities en Commentaar' klinkt sporadisch het straatruoer door over meer actuele thema's in onderwijsonderzoek, tussen theorie, praktijk en beleid.

### Activiteiten

Een van de meest in het oog springende activiteiten gedurende de afgelopen 25 jaar zijn diverse pogingen tot het programmeren van onderzoek. De programmering was in het onderwijsonderzoek van extra belang vanwege de diverse cliënten voor onderwijsonderzoek, het beleid, de onderwijspraktijk en de wetenschap. De stand van zaken en de wensen van de klanten moesten aan elkaar gerelateerd worden in een onderling samenhangend groter verband; het programma dat voor meerdere jaren zou kunnen gelden. Aanvankelijk lag het accent van de programmering nog bij de autonome onderzoeksinstituten, zoals blijkt uit het feit dat SVO aan het eind van de jaren zeventig en tachtig vroeg om programmering en meerjarenplannen op te stellen door de onderzoeksinstituten. Naderhand echter verplaatste het initiatief zich naar de programmering van onderzoek op landelijk niveau als resultaat van de verkenningen en de beleidsbeslissingen die daaromtrent werden genomen. Op deze wijze zou het mogelijk zijn de vragen voor onderzoek vanuit het beleid en de overheid te koppelen aan het aanbod van onderzoek in de vorm van onderzoeksdeskundigheid en de aanwezige kennis. Dit programmeringsideaal van selectieve ontwikkeling van bepaalde thema's voor onderzoek van onderzoeksprojecten binnen instituten en binnen vakgroepen lag binnen de eerste geldstroom ook ten grondslag aan de voorwaardelijke financiering, en later aan de ontwikkeling van onderzoeksprogramma's en daarmee gepaard gaande onderzoeksbeoordelingen. De achtergrond voor deze onderzoeksprogrammering is onder meer dat het voor de kennisgroei gewenst is, dat onderzoekers werken binnen grote programma's waarin telkens nieuwe vraagstellingen tot ontwikkeling komen en de resultaten van verschillende onderzoeken gecombineerd kunnen worden. Dit zou ook de publicatie, verspreiding en de benutting van onderzoeksresultaten ten goede komen. Behalve onderzoeksprogrammering binnen vakgroepen, onderzoeksinstituten, de SVO en later NWO (PROO en BOPO) en de LPC's vond ook een programmering op landelijk niveau plaats, al zij het minder sturend. Dat gebeurde voornamelijk op basis van de resultaten van de verkenningscommissie (De Moor, 1996).

Een tweede reeks van activiteiten die in de afgelopen 25 jaren veel aandacht heeft gekregen en tot resultaten heeft geleid is de benadrukking van onderzoekoutput en het karakter van de onderzoekoutput. Het gaat hierbij om minstens drie soorten resultaten: 1) promoties; 2) wetenschappelijke publicaties; en 3) praktijkpublicaties.

Ad 1). Ook al bestrijkt De Groot (1976) een lange periode (1950-1975), onderwijskunde en het daaraan gerelateerde onderwijsonderzoek kwamen pas na 1965 echt tot ontwikkeling. De toename van het aantal studenten in die tijd leidde tot een snelle uitbreiding van de staf in universiteiten om de onderwijsverplichtingen aan de nieuwe generatie studenten te kunnen vervullen. Gepromoveerde kandidaten waren nauwelijks aanwezig en de prikkel om te promoveren was gering.

In het onderwijsonderzoek lag in de periode tot 1975 zeker niet de nadruk op wetenschapsontwikkeling. Integendeel, van onderwijsonderzoek werd verwacht dat het een bijdrage leverde aan de oplossing van problemen in onderwijsbeleid of -praktijk en de verbetering van het onderwijs. Dit zou moeillijk te verenigen zijn, was de opvatting, met een meer wetenschappelijke gerichtheid en met promotieactiviteiten. In de periode na 1975 werd hieraan meer aandacht besteed en het beleid van de onderzoeksinstituten en vakgroepen was om, op basis van de onderzoeksprojecten (veelal gefinancierd door SVO), promoties te realiseren. Het aantal promoties is in de jaren daarna dan ook met sprongen omhoog gegaan. Nadat de medewerkers van onderzoeksinstituten en vakgroepen gepromoveerd waren, is dit beleid voortgezet in de promotietrajecten voor AiO's. In het onderwijsonderzoek zijn veel onderzoekers gepro-



moveerd op de resultaten van extern gefinancierd opdrachtonderzoek, daarmee illustrerend dat een combinatie van wetenschappelijke en praktijkopbrengst mogelijk is.

Ad 2). Door de toename van het aantal promoties werd ook een klimaat geschapen waarin gepromoveerden meer dan in het verleden publiceerden in wetenschappelijke tijdschriften over de resultaten van hun onderzoek. In eerste instantie gebeurde dat in de Nederlandstalige wetenschappelijke tijdschriften zoals het *Tijdschrift voor Onderwijsresearch*. Naderhand werd meer in de internationale wetenschappelijke tijdschriften gepubliceerd omdat wetenschapsontwikkeling internationaal is. Onderzoek en publicaties moeten in die opvatting dan ook internationaal de competitie met andere onderzoeken en publicaties aangaan. Zeker in de periode na 1985 valt een toename van de onderzoeksproductie in termen van publicaties nationaal en internationaal te constateren, zelfs in de periode waarin door reorganisatiemaatregelen binnen de universiteiten de eerste geldstroomformatie voor onderwijsonderzoek afnam. Nadruk op verantwoording van onderzoeksgelden via onderzoeksoutput leidde in elk geval tot een toename van publicaties.

Door de oprichting van de landelijke onderzoeksschool op het terrein van de onderwijskunde, ICO, met toelatingscriteria, die internationale publicaties benadrukken en de regelmatige onderzoeksvisitatie, is gerichtheid van onderzoek op internationale publicaties nog versterkt. Nederlandstalige wetenschappelijke publicaties krijgen daardoor minder interessant aanbod van manuscripten, een negatief maar onvermijdelijk effect van een juiste internationalisering van het theoriegerichte onderzoek.

Ad 3). Binnen onderzoeksinstituten was het publicatiebeleid, zeker in de SVO-periode, gericht op wetenschappelijke publicaties, maar, teneinde de maatschappelijke legitimering van onderwijsonderzoek en aparte financiering te verwerven en te behouden, was het ook noodzakelijk voor beleid en praktijk te publiceren en de onderzoeksresultaten voor de praktijk te vertalen. Het reeds genoemde blad *Didaktief* werd voornamelijk door onderzoekers volgeschreven op basis van de uitgevoerde onderzoeksprojecten. Dergelijke projecten zijn sinds de opheffing van SVO in aantal afgenomen. Dit kan ertoe leiden dat er minder praktijkpublicaties verschijnen die direct aan uitgevoerd onderzoek zijn gerelateerd. Dit is dan een gevolg van de vermindering van de omvang van het praktijkgerichte onderzoek en de gerichtheid op wetenschapsontwikkeling in de periode 1995-2000.

Zoals ook bij het ontstaan van onderzoeksprogramma's het geval is – het samenvoegen van projecten binnen thema's – zijn ook voor de praktijk overzichtstudies verschenen waarin resultaten van meerdere studies worden gecombineerd. SVO richtte daarvoor de reeks *Balans van Onderwijsonderzoek* op (o.a. Scheerens 1990; Creemers, 1991).

De wetenschapstheoretische en methodologische discussie in de onderwijskunde en het onderwijsonderzoek was vóór 1975 en ook in de eerste enkele jaren na 1975 bijzonder hevig. Het ging dan bijvoorbeeld over de theorie- en praktijkgerichtheid van het onderzoek en de bijdrage aan de vernieuwing van het onderwijs, en in dat verband over de vraag in hoeverre normen en waarden – wat 'goed onderwijs' was – een rol speelden in onderwijsonderzoek. De wetenschapstheoretische belangstelling voor dit soort vragen nam in de jaren na 1980 snel af. De nadruk kwam te liggen op de formulering van theorieën binnen het onderwijs en de toetsing ervan. Het beperkte zich tot de meer methodologische, technische aspecten, zoals de voor- en nadelen van kwantitatieve en kwalitatieve methodes. Waarschijnlijk is deze verminderde belangstelling voor het debat gedeeltelijk te verklaren uit de negatieve ervaringen die onderwijsonderzoekers hebben opgedaan bij hun deelname aan het debat over de inhoud en de richting van de onderwijsvernieuwingen in de jaren zeventig. Dit resulteerde in distantie ten aanzien van het debat over de inhoud, de waarden en de doelstellingen van het onderwijs. De nadruk van het onderzoek lag op de effecten van onderwijs en de wijze waarop de effecten verklaard zouden kunnen worden door middelen, processen, activiteiten, condities, en input in het onderwijs (Creemers, 1991). De door Leune (1985) verwachte opleving van het debat in het laatste decennium van de vorige eeuw heeft nauwelijks plaatsgevonden, gezien de flauwe deelname van onderwijsonderzoekers aan het debat over de pedagogische functie van de school.



## Problemen

Er zijn in de afgelopen 25 jaar diverse problemen in het onderwijsonderzoek te constateren; sommige zo ernstig dat het verwonderlijk is dat het onderzoek nog steeds bestaat. De financiering van het onderzoek was, tenminste in de perceptie van iemand die lange tijd directeur van een onderzoeksinstituut is geweest, een voortdurend probleem. In de periode 1970-1985 was het SVO-budget groter dan nodig was voor de goed beoordeelde voorstellen. Onderzoeksinstituten namen in omvang toe en er werd een groep onderwijsonderzoekers gevormd. Toen dit budget later sterk terugliep en verder gereduceerd werd na de opheffing van SVO, creëerde dit voor onderzoeksinstituten en universiteiten, faculteiten en vakgroepen forse problemen voor de personeelsorganisatie maar ook voor de wetenschapsontwikkeling. Er moest veel tijd en energie gestoken worden in de (re)organisaties en dit ging ten koste van de inhoud – aandacht voor theorievorming en ontwikkeling en uitvoering van onderzoek. Reductie van middelen en kwalitatieve vooruitgang zijn moeilijk te combineren.

Een probleem dat de afgelopen jaren bij voortduring heeft gespeeld was het geloof in de zegeningen van de programmering. Er is de afgelopen jaren ontzettend veel tijd en energie gestoken in het programmeren van onderzoek in de verwachting dat de kwaliteit en de maatschappelijke opbrengsten van onderzoek door een dergelijke programmering vergroot zouden worden. Daarvoor is nodig dat de intern wetenschappelijke argumenten en de maatschappelijk en beleidsmatige aandachtspunten goed op elkaar afgestemd worden en in een juiste verhouding staan. Dit is echter niet vaak het geval. Binnen de programmering zouden buitenwetenschappelijke argumenten minder invloed moeten hebben dan de binnenwetenschappelijke argumenten, te weten: het 'interessante' van de probleemstelling, de kennis die aanwezig is, het uitdagende van de nieuwe ideeën en dergelijke. Vaak is de programmering er echter juist op gericht om deze buitenwetenschappelijke argumenten binnen het wetenschapsbedrijf te halen (financieringskaders, maatschappelijk belang en dergelijke). Een gevolg daarvan is dat de programmering en de daaraan gekoppelde financiering tot modieus, uniform onderzoek heeft geleid vanuit een min of meer vergelijkbaar perspectief op het betreffende terrein.

Een minder helder probleem is de bijdrage van onderwijsonderzoek aan het maatschappelijke debat. Enerzijds kan geconstateerd worden dat onderzoekers wat huiverig stonden tegenover een actieve participatie aan en plaatsbepaling in het maatschappelijke debat rondom het onderwijs (zie hierboven), anderzijds werden de argumenten van onderwijsonderzoekers in het debat over het onderwijs nauwelijks gehoord, of het nu over de grootte van de klas gaat of de opvang van allochtonen, de plaats en functie van onderwijs, de basisvorming of het studiehuis. De empirische argumenten die hiervoor of vooral hiertegen werden ingebracht hebben in het maatschappelijk-politieke debat nauwelijks een rol gespeeld, waarna vervolgens geconstateerd wordt dat onderwijsonderzoek en onderwijskundigen weinig aan het debat hadden bijgedragen.

## THEMA'S VAN ONDERZOEK

In het voorgaande is er al op gewezen dat bij onderzoeksprogrammering op het terrein van het onderwijs twee gezichtspunten in beschouwing worden genomen en soms met elkaar concurreren. Dit zijn theoretische integratie, het kennisbestand op bepaald terrein en de maatschappelijke kennisvraag op korte, middellange en lange termijn. Programmering op landelijk niveau kan een eigen input leveren ten aanzien van de maatschappelijke kennisvraag, ervan uitgaande dat op dat niveau duidelijk en helder is wat de maatschappelijke ontwikkelingen zijn, wat de wensen van beleidsmakers en onderwijspraktijk zijn en dergelijke. Dit is echter veel minder het geval ten aanzien van theoretische integratie. Dat wordt vooral bepaald door de voortgang van het onderzoek en kennisontwikkeling binnen onderzoekseenheden zoals universitaire eenheden, onderzoeksinstituten en landelijke onderzoeksscholen. Op lan-

delijk niveau past een zekere terughoudendheid ten aanzien van deze inhoudelijk theoretische integratie hoewel het beleid natuurlijk vrij staat op basis van maatschappelijke kennisvragen bepaalde prioriteiten toe te kennen aan onderzoeksgebieden.

Voor het ontwikkelen van een kennisbestand omtrent het onderwijs is het van belang dat grote onderzoeksprogramma's voor de langere termijn bestaan mits ze voldoende progressie maken. Met progressie wordt hier bedoeld de mate waarin programma's in staat zijn om de eigen doelstellingen te realiseren. Het is opvallend dat een programma op het terrein van onderwijs-effectiviteit, waar het gaat om verschillen in de prestaties van de leerlingen te verklaren door middel van onderwijsvariabelen, uit de PROO-programmering verdwenen is terwijl het programma behoort tot de kern van onderwijs-theorie, -praktijk en -beleid. In de afgelopen jaren is juist op dit terrein progressie geboekt, wat blijkt uit de opbrengsten en de toenemende specificeringen van probleemstellingen.

Bij een dergelijk langer bestaand programma op landelijk niveau met een uitwerking ervan in lokale programma's in instituten is het mogelijk om een organisatorische integratie van onderzoek te realiseren. Onderzoek vanuit diverse geldstromen kan in onderling verband worden geprogrammeerd zowel nationaal als lokaal zodat opbrengsten van onderzoek zowel theoretisch als praktisch relevant gemaakt kunnen worden en daarmee als uitgangspunt voor nieuw praktijk- of theoriegericht onderzoek kunnen gelden.

De programmeringsinstanties zouden zich sterker bewust moeten worden van hun positie bij de programmering, bijvoorbeeld dat zij als beleidsinstantie zich meer met de maatschappelijke vraagstellingen dienen bezig te houden dan met de inhoudelijke integratie. De inhoudelijke integratie moet meer gericht zijn op de vooruitgang die binnen onderzoeksinstituten en binnen onderzoeksscholen geboekt wordt op bepaalde terreinen. Mutatis mutandis geldt hetzelfde voor onderzoeksinstituten; de maatschappelijke en beleidsmatige relevantie van onderzoeksvraagstellingen en terreinen voor onderzoek wordt niet door hen uitgemaakt maar door maatschappelijke en beleidsinstanties in het kader van hun programmering. Op het lokale niveau, het niveau van onderzoeksprogramma's binnen instituten, zal duidelijker gemaakt moeten worden via welke deskundigheid en kennisbestand en via welke onderzoeken een verdere bijdrage aan de ontwikkeling van kennis op een bepaald terrein kan worden geleverd. Uiteindelijk zal ook op die wijze verantwoording worden afgelegd, nl. aantonen welke nieuwe bijdrage dit onderzoek aan de ontwikkeling van het kennisbestand levert.

Het programma van de PROO 2001 probeert de indruk te vestigen dat het theoretische, maatschappelijke, organisatorische gezichtspunten integreert, terwijl de keuze van thema's meer bepaald lijkt te zijn door de maatschappelijke probleemstellingen dan door theoretisch interessante, productieve probleemformuleringen. Thema's als leerkrachten, ICT, de pedagogische functie en de invulling van het thema 'beleid' zijn meer het resultaat van beleidsmatige aandacht en wenselijkheden, bijvoorbeeld ten aanzien van de status van het beroep en de implementatie van ICT, dan dat ze bepaalde theoretisch rijke gezichtspunten reflecteren. Ook de nadere invulling ervan weerspiegelt meer beleidsmatige aandacht dan een grondige theoretische analyse met daaraan gekoppeld de formulering van de te verwachten wetenschappelijke bijdrage van het onderzoeksprogramma.

De programmering van de onderzoeksschool ICO op het terrein van onderwijskunde vindt duidelijk vanuit het gezichtspunt van theorieontwikkeling en beproevingsplaats. Er worden in het ICO-programma 10 thema's onderscheiden. Op deze thema's werken onderzoekers (senioren en PhD-studenten) van diverse instellingen en onderzoeksinstituten samen. Bij nadere beschouwing kunnen de thema's gerubriceerd worden in twee domeinen:

- 1) (Onderwijs)-leerprocessen
- 2) Vormgevingsvraagstukken in het onderwijs.

In deze tweedeling sluimert het gevaar van een te geringe wederzijdse benutting van theorie en de beproeving ervan; leren en instructie en de vormgeving ervan moet in een conceptueel framework geïntegreerd worden waarna onderzoeksthema's of programma's voor deeltherreinen worden ontwikkeld. Onderwijs is individueel – in termen van het leren en leeropbreng-

sten van individuele studenten - maar ook lokaal - leergroep (klas) en schoolgebonden - en contextueel - landelijk en internationaal.

Regelmatig maar met grote hardnekkigheid duiken bij de programmering vragen op over het eigene en de (inter)disciplinaire bijdrage in het onderwijsonderzoek aan theorievorming en onderwijsontwikkeling. Onderzoek dat een koppeling tussen onderwijsopbrengsten, leerprocessen, de vormgeving en condities voor leren en onderwijs omvat, zoals het onderwijs-effectiviteits- en verbeteringsonderzoek, is per definitie interdisciplinair omdat het gebruik maakt van theorieën en gezichtspunten vanuit sociologie, psychologie en pedagogiek en deze integreert. Het eigen gezichtspunt van interdisciplinair onderwijsonderzoek zit erin dat vanuit deze uiteenlopende disciplines de vormgeving van het onderwijs- en leerproces in de school als institutie en het onderwijsstelsel benaderd wordt. Mocht het zo zijn dat daar weinig nieuws over te vertellen valt in aanvulling op datgene wat uit de afzonderlijke disciplines aan kennisbasis is ontwikkeld en dat het niet te verwachten is dat er een dergelijke 'body of knowledge' over de verklaring van de verschillen in onderwijs-effecten verder is te ontwikkelen, dan is de vraag aan de orde of onderwijskunde en onderwijsonderzoek recht hebben op verder voortbestaan. Voorshands - ik kom hierop in de volgende paragrafen terug - kan het antwoord zijn dat de afgelopen jaren het interdisciplinaire onderwijsonderzoek middels thematische programma's aan een kennisbestand heeft bijgedragen maar dat er tevens veel, langzamerhand nauwkeuriger te formuleren, probleemstellingen voor toekomstig onderzoek (en theorievorming) overblijven. Die concentreren zich alle op de relatie tussen vormgevingsvariabelen, leerprocessen en leeropbrengsten.

## NOGMAALS: DE OPBRENGSTEN VAN ONDERWIJSONDERZOEK

In het voorgaande is er al op gewezen dat in een periode dat er minder middelen beschikbaar waren - en zijn - en daardoor minder onderzoek en minder onderzoekers, toch het aantal publicaties is toegenomen. Dat betekent dat de gemiddelde opbrengst per onderzoeker is gestegen. Dit is waarschijnlijk mede een effect van de formulering van outputcriteria en de onderzoeksbeoordeling. Vooral de verhouding tussen de omvang van het onderwijsonderzoek (aantal onderzoekers, budget, etc.) en het aantal internationale publicaties kan als een kwaliteitsindicatie worden opgevat. Uit het grote aantal publicaties dat internationaal geaccepteerd wordt, kan worden afgeleid dat het Nederlandse onderwijsonderzoek zowel kwantitatief als kwalitatief goed is. Internationale publicaties zijn niet alleen van belang vanuit het oogpunt van strengere selectie en grotere competitie. Het forum voor wetenschapontwikkeling is internationaal. In de (internationale) gemeenschap van onderwijsonderzoek wordt de standaard voor het vakgebied vastgesteld en wordt de bijdrage aan de wetenschapontwikkeling gewaardeerd. De progressie in de wetenschapontwikkeling kan alleen op dat niveau worden vastgesteld ook al zal benutting van het kennisbestand in veel gevallen contextualisering vereisen.

In een overzichtsartikel over de stand van zaken in de pedagogiek en onderwijskunde (Creemers, 1991) verwees ik naar een collega in de pedagogiek die voor zijn vakgebied het einde voorzag binnen een tiental jaren. Hij heeft ongelijk gekregen. Naar verwachting zal dit ook gebeuren met die onderwijsonderzoekers die nu op basis van de gerichte proporties verklaarde variantie het einde van het onderwijsonderzoek zien naderen.

Voor degenen die pessimistisch zijn over de verdere theorieontwikkeling op het terrein van onderwijs, is de bijdrage voor het voortbestaan van het onderwijsonderzoek of de individuele professie aan de onderwijsontwikkeling van nog groter belang dan voor degenen die in aansluiting op Brus (en anderen, 1981) de verantwoordelijkheid voor de praktijk van het onderwijs en de benutting van de kennisbasis bij die praktijk (inclusief het beleid) leggen.

De nadruk binnen onderwijsonderzoek heeft in de afgelopen jaren, onder andere door de herpositionering van het onderwijsonderzoek, sterk gelegen op wetenschappelijke theorievorming en dat zou wel eens ten koste van deze maatschappelijke bijdragen kunnen zijn

gegaan. Echter wanneer theoretisch geen verdere vooruitgang wordt geboekt, worden bestaande inzichten herhaald c.q. opnieuw vertaald naar het actuele onderwijsprobleem. Zo wordt het debat tussen voor- en tegenstanders over de effectiviteit van het onderwijs in Engeland niet zozeer gevoerd op basis van de ontwikkeling van nieuwe inzichten maar op basis van het 'recyclen' van bestaande inzichten en nog vaker 'opinions' (Sammons & Reynolds, 1999). Deze stagnatie kan worden voorkomen door binnen een onderzoeksprogramma het accent te leggen op theoretisch interessante progressieve probleemformuleringen in een maatschappelijk relevant gebied van problemen. Maar die zijn er in het onderwijs te over.

## CONCLUSIES: DE TOEKOMST VAN HET ONDERWIJSONDERZOEK

Het *Tijdschrift voor Onderwijsresearch* heeft 25 jaar bestaan. In die periode is een groot aantal publicaties over de meest uiteenlopende thema's gepubliceerd. In het begin waren dat vaak zogenaamde 'didactometrische' onderwerpen, zoals toetsen en examens, maar hoe langer hoe meer waren dit onderwijsinhoudelijke problemen met betrekking tot leerprocessen en de vormgeving ervan. Het onderwijsonderzoek heeft in die periode een sterke ontwikkeling doorgemaakt van een beginnende professie in 1975 met een eigen tijdschrift tot een internationaal erkende bijdrage van het Nederlandse onderwijsonderzoek aan de 'educational research'. Geconstateerd is dat de opbrengst en de internationale bijdrage de afgelopen jaren enorm gegroeid zijn en min of meer als gevolg van dit succes legt het *Tijdschrift voor Onderwijsresearch* het loodje. Dit is, hoe jammer het feit op zich ook is, niet zozeer een teken van armoede maar juist van de sterke internationale positie van het onderwijsonderzoek. Vanuit een onderzoeksprogrammering en de theoretische en maatschappelijke gezichtspunten hierbij hebben wij ons de vraag gesteld of ons kennisbestand op het gebied van onderwijs in die jaren ook vergroot is. Het antwoord is positief maar er is een kanttekening gemaakt bij de onvoldoende persistentie en stabiliteit in de huidige onderzoeksprogrammering, veroorzaakt door het hanteren van de verkeerde criteria bij het vaststellen en vasthouden van onderzoeksprogramma's. Het lijkt erop dat wetenschappelijk succes een negatieve indicatie is voor de handhaving van bepaalde programma's. Dit zou echter ook het resultaat kunnen zijn van het feit dat binnen een bepaalde beleids- en praktijkcontext geen verdere vooruitgang in maatschappelijke probleemformuleringen ten aanzien van het onderwijs wordt geboekt. De discussie over de kwaliteit van het onderwijs kan dan verzanden. Dit leidt ook niet tot verdere ontwikkelingen binnen theorievorming en onderzoek; die komen voort uit de interne dynamiek van de wetenschap. Wanneer er aan getwijfeld wordt of er theoretisch nog winst te boeken is zou dit een contra-argument voor het voortbestaan van onderwijsonderzoek kunnen betekenen, zeker wanneer dit versterkt wordt door onvoldoende maatschappelijke participatie van onderwijsonderzoek. Ook al hebben individuele onderzoekers zeker een bijdrage geleverd aan beleid en praktijk van het Nederlandse onderwijs, de internationale oriëntatie van het Nederlandse onderwijsonderzoek is gepaard gegaan, mede door de wijzigingen in de organisatorische en financieringsstructuur in het onderzoek, met verminderde aandacht voor de problemstellingen in de praktijk en het beleid van het Nederlandse onderwijs. De twijfels over theoretische groei en de bijdrage aan onderwijsontwikkeling zouden twee argumenten kunnen vormen tegen het voortbestaan van het onderwijsonderzoek c.q. de onderwijskunde.

Er zijn echter heel duidelijk contra-argumenten. Die worden gevonden in het succes van grotere programma's van onderwijsonderzoek ten aanzien van leerprocessen, en de daaraan gekoppelde onderwijsprocessen. In een 'mastertheorie' op het terrein van het onderwijs kunnen diverse gezichtspunten, thema's en opbrengsten geïntegreerd worden en weer aanzetten tot verdere theorievorming, progressieve probleemformuleringen en onderzoek. In 1982 werd hiervoor al gepleit voor pedagogiek en onderwijskunde. In latere publicaties is daarop verder ingegaan – waarbij in 1991 geconstateerd werd dat een dergelijk frame aanwezig is. In onderzoeksprogramma's is de afgelopen 10 jaar aan onderdelen ervan voldoende besteed en dit



heeft tot verdere kennisgroei geleid die – dat voor de pessimist – zeker niet voltooid is. Vanuit een dergelijk framework zal het antwoord moeten worden gegeven op maatschappelijke vraagstellingen over het onderwijs.

Dit vereist een programmering van onderwijsonderzoek die meer vanuit de theorie en de onderzoeksstrategie plaatsvindt dan vanuit actuele probleemstellingen, zeker wanneer die praktijk niet tot progressieve probleemformuleringen komt omdat deze keus in universiteiten en instituten aanwezig is, is dat het niveau van programmering. Een dergelijk, meer theoretisch geïnitieerd programma kan vervolgens gekoppeld worden aan de programmering van praktijk- en beleidsprobleemstellingen. Deze twee onafhankelijke inputs kunnen dan resulteren in een financieringsprogramma ten aanzien van onderzoek voor instituten die bereid zijn theorie- en praktijkgericht onderzoek te verrichten. In ieder geval in de landen waar onderwijsonderzoek overeind blijft, is onderzoek betrokken bij onderwijs- c.q. schoolontwikkelingen en -verbetering (zie het onderzoek in de VS, UK en Australië). Ook al staat de theoretisch-empirische oriëntatie daarbij voorop, die opbrengst dient te 'werken' in de onderwijspraktijk en in onderwijsbeleid. Dit zou een argument kunnen zijn om niet alleen de theoretische oriëntatie in de programmering vanuit het basisniveau te doen plaatsvinden maar dit uit te breiden ten aanzien van de maatschappelijke oriëntatie en maatschappelijke opbrengsten. Onderzoekers en instituten bewijzen in de praktijk dat het werkt. De programmering op hogere niveaus is een mechanisme voor financiering dat het succes in theoretisch en maatschappelijk opzicht van het onderzoek op basisniveau zou moeten volgen. Op deze wijze, ook al blijft in mijn visie het accent voor onderzoeksprogrammering liggen op de intern wetenschappelijke dynamiek (kennisbestand en progressieve probleemformuleringen) kan deze programmering zeker op het basisniveau van het programma en de uitvoering ervan verbonden worden met de onderwijsontwikkeling in praktijk en beleid. Programmering op hogere niveaus zou dit kunnen aggregeren en zich vooral richten op het scheppen en verbeteren van de condities voor succesvol onderzoek. Dat is onderzoek dat blijkens publicaties leidt tot uitbreiding van het kennisbestand en nieuwe onderzoeksvragen. Daaraan heeft TOR de afgelopen 25 jaar een bijdrage geleverd.

## LITERATUUR

- Bosker, R.J. en Roede, E. (1996). *Sturing en versnippering: een programmeringstudie voor onderwijsonderzoek in het primair onderwijs gebaseerd op een toekomstanalyse*. Amsterdam: Kohstamm Instituut.
- Brus, B. Th. (1981). Kleren voor een keizer. In *Tijdschrift voor Onderwijsresearch*, 6: 1, 3-15.
- Creemers, B.P.M. (1982). Pedagogiek in de jaren tachtig. In A.D. Wolff-Albers en H.F.M. Crombag (red.), *Visies op onderzoek in enkele sociale wetenschappen: preadviezen ten behoeve van de beleidsnota maatschappij- en gedragswetenschappen*. 's-Gravenhage: Staatsuitgeverij.
- Creemers, B.P.M. (1983). De bijdrage van onderwijsonderzoek aan de verhoging van de kwaliteit van het onderwijs. In: B. Creemers, W. Hoeben en K. Kooops (Red.) *De kwaliteit van het onderwijs*. Haren: RION.
- Creemers, B.P.M. & W.Th.J.G. Hoeben. (1984). *Onderzoek tussen onderwijsontwikkeling en wetenschapsontwikkeling*. 's-Gravenhage: SVO.
- Creemers, B.P.M. (1985). De verdere ontwikkeling van de onderwijskunde. *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs*, 1985, 1, 30-40.
- Creemers, B.P.M. (1991). Ouder en wijzer? Reflecties over onderzoek in de pedagogiek, andragogieek en onderwijskunde. In A. Pennings et al. *Bijdragen aan pedagogisch onderzoek 1990*. Amersfoort: ACCO.
- Creemers, B.P.M. (1991b). Effectieve Instructie: een empirische bijdrage aan de verbetering van het onderwijs in de klas. 's-Gravenhage: SVO.
- Deen, N. *Een halve eeuw onderwijsresearch in Nederland, Net Nutsseminarium voor Pedagogiek 1919-1969*. Groningen: Wolters Noordhof, 1969.
- Gelder, L. van. *De middenschool*. Amsterdam: Stichting JVJO, 1970
- Groot, A.D., de. (1976) "Ontwikkeling in de Nederlandse onderwijsresearch". *Tijdschrift van Onderwijsresearch*, 1 (1976) nr. 4, p. 145-160.
- Groot, A.D., de. (1982). *Academie en forum: over hoger onderwijs en wetenschap*. Meppel: Boom.



- Leune, J.M.G. (1985). Enkele opmerkingen over waardeoverdracht door de school. *Pedagogisch Tijdschrift* 10, 3, 114-121.
- Moor, R.A. de (1996). *Nederlands Onderwijsonderzoek verkend 1996*. Amsterdam: Overlegcommissie Verkenningen.
- Postma, L. en W.L. Wardekker, (1981). Experimenterend veranderen van onderwijs. In: B.P.M. Creemers (Red.): *Onderwijskunde als opdracht*. Groningen: Wolters Noordhoff.
- Sammons, P. en Reynolds, D. (1999). A Partisan Evaluation: John Elliott on School Effectiveness. In *Cambridge Journal of Education* 27, 1: 123-136
- Scheerens, J. (1989). *Wat maakt scholen effectief?: samenvatting en analyses van onderzoeksresultaten*. 's-Gravenhage: SVO.
- Trotsenburg, E.A. van. *Ontwikkelingslijnen in het empirisch onderzoek van pedagogische en didactische vraagstukken*, Scholars' Press, 1972
- Vrijenhoef, H.: "Prof. Crombag's uitspraken op de ontleedtafel van de discussie, Vijf zorgen van een Utrechtse forum" in *Didactief* 12 (1982): 10.

# Metacognitie van lerenden in onderwijsleerprocessen

Marianne Elshout-Mohr

SCO-Kohnstamm Instituut\* en Instituut voor de Lerarenopleiding van de UvA

## ABSTRACT

The topic of this article is metacognition. What is it and how is it affected by training? First, the concepts metacognitive experience, metacognitive knowledge, metacognitive strategies and cognitive monitoring are presented. Individual differences in metacognitive knowledge about learning and studying are briefly discussed as well as individual differences in cognitive monitoring. In the second part of the article we argue that study skill trainers tend to be too optimistic about the potential benefits of training in metacognitive strategies. A more realistic view on training in metacognition is presented.

## INLEIDING

Lerenden kunnen in principe veel te weten kunnen komen over de optimale sturing van eigen leerprocessen. Afhankelijk van intelligentie, scholing en ervaring beschikken mensen over metacognitieve kennis en maken zij er gebruik van. Diegenen die meer en beter gebruik maken van metacognitieve kennis leveren in de regel de betere prestaties. Deze bevinding is aanleiding geweest om lerenden te stimuleren tot het verwerven van metacognitieve kennis, het ontwikkelen van metacognitieve strategieën en het zelfstandig reguleren van hun studeergedrag. Het effect van stimulerings- en trainingsprogramma's blijkt echter in de praktijk beperkt. In deze bijdrage wordt gesteld dat de resultaten tegenvallen omdat men te hoge verwachtingen koestert. Beargumenteerd wordt dat men tot meer bescheiden verwachtingen komt wanneer expliciet aandacht besteed wordt aan de samenhang tussen metacognitie en andere aspecten van het cognitieve functioneren die voor de taakuitvoering van belang zijn. Men krijgt daarmee tevens beter zicht op de reële mogelijkheden om door middel van training een goed gebruik van metacognitieve kennis en strategieën te bevorderen.

Voordat in het tweede deel van het artikel wordt ingegaan op de problematiek rondom training in metacognitie, wordt eerst toegelicht wat onder de diverse aspecten van metacognitie wordt verstaan. Daartoe behandelen we de begrippen metacognitieve gewaarwording en metacognitieve kennis en vervolgens het begrip metacognitieve strategie. Daarna gaan we inhoudelijk in op individuele verschillen in metacognitieve kennis in het algemeen en in metacognitieve kennis over studeerprocessen in het bijzonder. Tot slot van het eerste gedeelte van het artikel wordt het begrip 'cognitive monitoring' (bewaken van eigen cognitief functioneren) onder de loep genomen.

## METACOGNITIEVE GEWAARWORDINGEN EN METACOGNITIEVE KENNIS

Mensen die leren en studeren zijn zich bewust van een deel van de processen die hierbij een rol spelen en hebben opvattingen over factoren die het behalen van goede studieresultaten vereenvoudigen dan wel bemoeilijken. Dit geldt zowel voor beginnende leerlingen als voor lerenden met meer ervaring. Aan de basis van kennis over cognitieve processen (in dit geval leer- en studeerprocessen) ligt de *metacognitieve gewaarwording of ervaring*. Flavell (1979) stelt dat er van een 'metacognitive experience' sprake is zodra iemand zichzelf of een ander beleeft als een

\* SCO-Kohnstamm Instituut, Grote Bickersstraat 72, 1013 KS Amsterdam.

cognitief functionerend persoon die zich in een bepaalde cognitieve toestand bevindt. Die toestand kan er bijvoorbeeld een zijn van begrip of onbegrip, van zekerheid of onzekerheid of van weten of niet weten. In de metacognitieve gewaarwording kan de aandacht gericht zijn op de cognitieve toestand als zodanig, maar ook op de persoons-, taak- en strategie-kenmerken die van invloed zijn (of zijn geweest) op *het bereiken van* deze toestand. Een meisje dat signaleert dat het goed onthouden heeft welke boodschappen er gedaan moesten worden bij de groenteman, kan opmerken dat zij dit beter onthouden heeft dan haar jongere broertje; zij legt dan de nadruk op de persoonskenmerken. Zij kan daarnaast opmerken dat ze het beter onthoudt als het om dagelijkse dingen gaat waarvan ze weet hoe die eruit zien dan wanneer het gaat om dingen die zij niet kent; zij let dan op de taakkenmerken. Tot slot kan ze opmerken dat het onderweg repeteren van de opdracht helpt om niets te vergeten; vooral als je alles stuk voor stuk op je vingers aftelt. In dit laatste geval legt zij het accent op de invloed van de aanpak, de strategie.

Metacognitieve ervaringen worden soms onder woorden gebracht en uitgewisseld met anderen. Maar ook als dat niet gebeurt stapelen de observaties zich gaandeweg op. Ouders en opvoeders hebben een inbreng en zo ontstaat een fonds aan metacognitieve kennis. Gezien het onsystematische karakter van het verwervingsproces is de kwaliteit van de kennis niet gegarandeerd. In navolging van Flavell omschrijven we *metacognitieve kennis* als dat segment van kennis dat betrekking heeft op de cognitieve taken, doelen, acties en ervaringen. Flavell legt er de nadruk op dat metacognitieve kennis alle kenmerken heeft van 'gewone' kennis, kennis van andere segmenten van de wereld. Metacognitieve kennis kan correct of incorrect zijn; aangevuld, gecorrigeerd, onthouden en vergeten worden; spontaan opduiken of via een zoekprocedure in het geheugen worden opgespoord. De *relatie tussen metacognitieve kennis en cognitieve processen* is echter, gezien binnen het kader van de opvattingen van Flavell (1979), toch een beetje bijzonder omdat de beschikbaarheid van metacognitieve kennis er op wijst dat de persoon zich (op enig moment) bewust is geweest van eigen cognitief functioneren en van variabelen die hij of zij voor dit functioneren van belang achtte. In situaties waarin iemand zich afvraagt hoe, in cognitieve taken, successen uit het verleden herhaald kunnen worden en mislukkingen vermeden, kan metacognitieve kennis daarom op verschillende manieren een rol spelen, zoals blijkt in de volgende paragraaf.

### METACOGNITIEVE STRATEGIEËN

Flavell beschrijft dat metacognitieve kennis spontaan kan opduiken tijdens het werken aan een cognitieve taak. Iemand die vastraakt in een probleem kan zich bijvoorbeeld herinneren dat hij of zij in het verleden een soortgelijke impasse met succes heeft opgelost door het probleem even te laten rusten. In zo'n geval gaat de metacognitieve ervaring samen met het beschikbaar komen van informatie ten aanzien van de te volgen gedragslijn of strategie. Als dit herhaaldelijk gebeurt kan zich een gewoonte vormen. Het 'even laten rusten van een probleem dat tot een impasse leidt' kan bijvoorbeeld een manier van doen worden waar de persoon zich niet eens meer van bewust is. Hij of zij *doet* het vervolgens gewoon zo, zonder de bewuste tussenkomst van metacognitieve kennis. Voor deze vorm van vanzelfsprekend strategisch gebruik van metacognitieve kennis is geen aparte benaming; het valt onder de ruimere categorie 'zelfregulatie'.

Metacognitieve kennis kan door de persoon ook bewust worden uitgewerkt tot een *metacognitieve strategie* die vervolgens doelbewust wordt gehanteerd. Het meisje dat ontdekt heeft waarom zij beter is in het onthouden van de boodschappen dan haar broertje, kan die kennis bewust gaan gebruiken en daar haar gedrag op gaan afstemmen. Zij zal bijvoorbeeld zorgen dat ze haar handen vrij heeft (voor het aftellen) als ze op iets belangrijks wordt uitgestuurd. Flavell vestigt er de aandacht op dat het gebruik van metacognitieve strategieën niet per definitie gunstig uitwerkt. Soms worden prestaties ongunstig beïnvloed door bewuste begeleiding van eigen cognitieve activiteiten. Het aftellen op de vingers kan, als de ontdekking net is gedaan, de aandacht zozeer vasthouden dat het onthouden eronder lijdt. Ook kan het kind, ten onrechte gaan denken dat het bijzonder moeilijk, zo niet onmogelijk, is om iets goed te onthouden als je

er, om welke reden dan ook, je handen niet bij kunt gebruiken. Een opmerkelijke bijkomstigheid is dat er geen consistent verband bestaat tussen de graad van correctheid van de metacognitieve kennis en het nut dat de persoon ondervindt van het gebruiken van die kennis. Strategieën die zijn gebaseerd op correcte metacognitieve kennis leiden (op een gegeven moment, bij één gegeven persoon) dikwijls niet tot betere resultaten dan strategieën die zijn gebaseerd op incorrecte aannamen. Dit maakt het voor mensen heel moeilijk om een consistent en correct metacognitief model te construeren op basis van een privé-analyse van 'wat werkt'. Bij gebrek aan formele scholing in leren en studeren, bepalen daarom zelfs studenten in het hoger onderwijs hun studeergedrag dikwijls op idiosyncratische opvattingen (Elshout-Mohr & Van Daalen-Kapteijns, 1985).

Flavell merkt op dat veel van de strategieën die mensen gebruiken zowel metacognitief als cognitief genoemd kunnen worden. Als voorbeeld noemt hij de strategie die inhoudt dat de persoon tijdens het studeren regelmatig nagaat of hij in staat is samen te vatten wat hij heeft gelezen en geleerd. De persoon die de strategie gebruikt *begeleidt* het eigen studeerproces door te controleren of er kennis wordt opgebouwd. Men kan daarin een goede reden zien om te spreken van een metacognitieve, procesbegeleidende, strategie. Tijdens het controleproces is de persoon echter ook *uitvoerend* bezig. Hij herhaalt informatie, herformuleert en elaboreert. Deze cognitieve activiteiten vormen als zodanig een bijdrage aan het leerproces. Men kan de strategie daarom ook beschouwen als een cognitieve, informatie verwerkende, strategie. Belangrijke categorieën procesbegeleidende activiteiten zijn 'planning' (wat zal ik gaan doen en in welke volgorde?), 'monitoring' (met welke cognitieve activiteit ben ik bezig en in welke cognitieve toestand bevind ik me?), 'checking' (laat ik eens controleren of mijn subjectieve indruk van proces- of produkt-kwaliteit wel klopt) en het testen, reviseren en evalueren van eigen werkmethoden (Baker & Brown, 1984, Vermunt, 1989).

## DE ONTWIKKELING VAN METACOGNITIEVE KENNIS

Wat kinderen in eerste instantie ontdekken over het verwerven van kennis en vaardigheden wordt bepaald door toevallige observaties die gemakkelijk leiden tot overhaaste en onvruchtbare conclusies. Een voorbeeld vormt een jongen van vijf jaar die opmerkt dat de volwassenen in zijn omgeving in verschillende vakantielanden in voor hem vreemde talen communiceren. Zijn conclusie is: "Dat kan ik ook als ik groter ben. Net zoals mijn voeten even groot zullen worden als die van hen, zo zal ook mijn talenkennis uitgroeien". Aangezien hij beide ontwikkelingen met vertrouwen tegemoet ziet, vindt hij geen aanleiding om erop vooruit te lopen door zich in te spannen. Als het zover is dat er een taal 'bijgekomen' is, zal hij het vanzelf wel merken. Geheel willekeurig zullen de denkbeelden die bij kinderen opkomen overigens niet zijn. Elders in dit nummer wordt ingegaan op de voorlopers van verwoordbare metacognitieve kennis die kunnen worden aangetoond bij jonge kinderen (Samarapungavan & Milikowski).

Het 'naar school gaan en formele scholing ontvangen' is een belangrijke factor in het ontwikkelen van bruikbare constructieve opvattingen over wat leren is. Een kenmerk van het leren op school is dat er veel wordt geïnvesteerd in het behandelen van algemene regels en principes die van toepassing zijn op *verschillende* vraagstukken binnen *verschillende* kennisdomeinen. Dit accent is aantoonbaar van invloed op de opvattingen van mensen over de wijze waarop verworven kennis toegepast kan worden in nieuwe situaties. Schliemann & Acioy (1989) rapporteren dat (Braziliaanse) volwassenen die geen formeel onderwijs hebben gevolgd, inderdaad anders denken over het verwerven en gebruiken van kennis dan mensen die wel enkele jaren op school hebben gezeten. Wanneer men de niet formeel geschoolden vraagstukken voorlegt die *in oppervlakte-kenmerken* verschillen van vraagstukken die zij goed kunnen oplossen, maar die er in onderliggende probleemstructuur volledig mee overeenkomen, vinden zij het vanzelfsprekend dat het 'nieuwe' vraagstuk voor hen *niet* oplosbaar is. Als er een oplossing voor moet komen, dan is het zaak om door middel van concrete proefnemingen te achterhalen wat de uitkomst zou kunnen zijn; of beter nog: men kan deskundigen raadplegen. Diegenen die wel enkele jaren



formeel onderwijs hebben genoten, hebben ervaren dat kennisverwerven ertoe kan leiden (en zelfs tot doel kan hebben) dat men middelen in handen krijgt om zelfstandig oplossingen te vinden voor vraagstukken die 'als zodanig' nooit eerder zijn opgelost.

Aan kennis over leren en studeren en de manier waarop leerlingen dat zouden moeten doen, kan op school niet alleen impliciet een bijdrage worden geleverd, maar ook expliciet. Binnen alle fasen van het reguliere onderwijs is groeiende aandacht voor 'leren leren' en de ontwikkeling van studievaardigheden.

In dit verband zijn twee zaken van belang. In de eerste plaats is veel kennis omtrent de variabelen en principes waarvan lerenden eigenlijk weet zouden moeten hebben, op dit moment wel beschikbaar, maar voor de onderwijspraktijk nog te weinig toegankelijk (Van Hout-Wolters, 1992). Zo dringt bijvoorbeeld pas geleidelijk door dat lerenden verschillen in leerconceptie en voorkennis en dat die verschillen doorwerken in de doelen die zij zichzelf stellen, de lesmaterialen die zij benutten en de cognitieve en metacognitieve strategieën die zij hanteren (Marton, Hounsell & Entwistle, 1984; Boekaerts, 1991; Duffy & Jonassen, 1991; Ng & Bereiter, 1992). In de tweede plaats moet rekening gehouden worden met de mogelijkheid dat een deel van de metacognitieve kennis die door 'deskundigen' wordt doorgegeven aan docenten en leerlingen, misvattingen bevat. In een vaktijdschrift voor (onderwijs)psychologen wijst Elshout (1987) op het bestaan van de mythe van het 'heilzame leren' en Snow (1990) bespreekt, eveneens in een blad voor vakgenoten, de mythe van de 'harmonieuze ontwikkeling van vaardigheden'. De essentie van beide mythen is dat het nooit kwaad zou kunnen om te leren of om vaardigheden te verwerven, omdat men bij verder leren nooit last en alleen maar plezier zou kunnen hebben van wat er reeds verworven is. Dit 'baat het niet dan schaadt het niet'-principe is onjuist. Vaardigheden die lerenden op het ene moment aanleren (bijvoorbeeld vraagstukken oplossen met behulp van een trucje), moeten ze op een later moment weer *afleren* en dat maakt het leerproces er bepaald niet gemakkelijker op (Greeno, 1988).

Op grond van ontwikkeling, ervaring en scholing ontstaan bij mensen verschillen in opvattingen over wat het sturen van eigen leren kan opleveren, wat het inhoudt en wat de belangrijke punten zijn om op te letten. We noemen daar drie voorbeelden van.

*a. Met de leeftijd veranderen leerlingen in hun opvattingen over de basis van goed presteren op school*

Bij veel leerlingen breekt, ergens omstreeks het twaalfde jaar, het besef door dat het leerlingkenmerk 'intelligentie' van wezenlijk belang is voor schoolprestaties. De oorspronkelijke opvatting dat 'goed je best doen' op school gelijk te stellen zou zijn met 'goed kunnen leren' en 'goed presteren', wordt daardoor herzien. Sommige leerlingen doen dit zo drastisch dat 'testen van de eigen begaafdheid' en 'vermijden van situaties waarin de beperkingen ervan aan het licht kunnen treden' belangrijker worden dan het leveren van goede (of zelfs maar acceptabele) schoolprestaties (Nicholls, 1978; Dweck & Leggett, 1988). Het is in die ontwikkelingsfase van belang dat in het metacognitieve model van de leerlingen plaats wordt ingeruimd voor zowel de 'intelligentie' als de 'strategie' (Nolen, 1988). Over de samenhang tussen beide moet de leerling een wat genuanceerder idee krijgen dan: "wie zijn best moet doen voor het behalen van goede leerresultaten en belang hecht aan een goede aanpak, is dommer dan wie 'zonder' kan".

*b. Ook 'volwassen' opvattingen over goed studeren en goed doceren zijn persoons- en situatiegebonden*

Er is veel onderzoek gedaan naar de complexe opvattingen over leren en studeren die bestaan bij 'gevoerde' lerenden onder verschillende leer-omstandigheden (Weinstein & Mayer, 1986; Schmeck, 1988; Vermunt, 1991). Bij studenten in het Hoger Onderwijs is er een duidelijke *samenhang* te constateren tussen opvattingen over goed studeren en opvattingen over goed onderwijzen (Janssen en de Neve, 1988). Studerenden die sterk hechten aan het zo snel en goed mogelijk bereiken van de leerdoelen van een cursus, stellen er bijvoorbeeld meer prijs op dat de docent de leerstof goed structureert dan studerenden die het van primair belang vinden dat wat ze leren praktische toepassingswaarde heeft of tot zelfontplooiing bijdraagt. Niet alleen per-



soonskenmerken, maar ook de leerinhoud en de leeromgeving spelen hierbij een rol (Elshout-Mohr & Van Daalen-Kapteijns, 1989).

Ook nadat de periode waarin formeel onderwijs wordt genoten, is afgesloten, staat de ontwikkeling van metacognitieve kennis bij sommigen niet stil. Het accent ligt dan echter vooral op de opvattingen die een rol spelen bij het 'zelfstandig verwerven van complexe kennis'. Van die opvattingen maakt bijvoorbeeld deel uit dat lerenden zich persoonlijk (moreel) verplicht gaan voelen tot het vermijden van oppervlakkig 'kennis nemen van' en tot volharding in hun streven naar het verkrijgen van inzicht in hoe de dingen werkelijk zijn (Bereiter, 1990). Lerenden die op dit niveau van 'gevoerdheid in leren en studeren' functioneren, doen bijpassende metacognitieve kennis op. Deze heeft onder andere betrekking op het soort moeilijkheden dat eigen is aan situaties waarin een persoon zich zelfstandig complexe kennis tracht eigen te maken. Voorbeelden daarvan zijn de altijd weer grotere hoeveelheid te verwerken informatie dan werd verwacht en de voortdurende moeite die gedaan moet worden om niet te zwichten voor te eenvoudige oplossingen of halve waarheden. Deze metacognitieve kennis kan op zijn beurt weer de basis vormen voor metacognitieve strategieën die bewust kunnen worden ingezet om de moeilijkheden te omzeilen of te overwinnen. Een in de praktijk door academici veel gevolgde werkwijze bestaat bijvoorbeeld uit het vormen van een studiegroep rondom een gezamenlijk thema. De leden van de groep weten uit ervaring dat ze als lid van de groep meer kans hebben de bovenstaande moeilijkheden het hoofd te bieden dan als eenling.

c. *Lerenden met dezelfde metacognitieve kennis en vaardigheden verschillen in de mate waarin ze het nuttig en nodig vinden er gebruik van te maken*

Het blijkt van aanzienlijk praktisch belang of lerenden zelf over de nodige kennis en vaardigheden beschikken om medeverantwoordelijkheid te dragen voor het eigen studeerproces (Corno, 1986). Dit neemt niet weg dat opmerkelijk frequent valt te constateren dat lerenden weten wat ze zouden moeten doen om het eigen leren te optimaliseren, maar bewust weinig gebruik maken van effectieve cognitieve en metacognitieve strategieën. In de recente literatuur wordt veelvuldig gewezen op de verwevenheid van motivatie en cognitief functioneren (Prawat, 1989; Pintrich, 1989). Experimenteel onderzoek naar de preciese aard en gevolgen van die verwevenheid is echter schaars. Rothkopf (1988) vergelijkt lerenden die 'wel weten wat ze zouden moeten, maar het niet doen' met mensen die onverstandig eten of zich onfatsoenlijk gedragen terwijl ze beter weten. Veel van wat door lerenden zelf optimaal geacht wordt, komt *doelbewust* te vervallen als een kosten-baten-analyse uitwijst dat het realiseren van de strategie veel werk met zich meebrengt en niet echt nodig is. De lerende geeft op zo'n moment een lage prioriteit aan het bevorderen van het eigen leren en heeft weinig behoefte aan het benutten (of aanvullen) van de metacognitieve kennis die in dat verband relevant zou zijn. Eerder is zo iemand wellicht geïnteresseerd in (eveneens metacognitieve!) kennis omtrent de preciese relatie, per vak, tussen verschillende gradaties van oppervlakkig werken en de bijbehorende studieresultaat-verwachtingen. De situatie is in dit geval heel anders dat die van lerenden die graag de door hen optimaal geachte cognitieve en metacognitieve *leerstrategieën* zouden benutten, maar daar geen kans toe zien (Biggs, 1985). Studenten die uit onvermogen geen 'diepte-strategieën' gebruiken terwijl ze dat wel zouden willen, steken van training in diepe verwerking van leerstof relatief veel op.

#### INDIVIDUELE VERSCHILLEN IN 'COGNITIVE MONITORING'

Het begrip 'cognitive monitoring' neemt in het werk van Flavell een belangrijke plaats in. Het 'in de gaten hebben en houden' van eigen cognitief functioneren en de cognitieve toestanden die er het resultaat van zijn, is voor lerenden belangrijk. Iemand die opmerkt (gewaar wordt) dat hij iets niet snapt, kan stappen ondernemen; hij kan zich een ander *doel* stellen ('misschien moet ik dit niet begrijpen, maar gewoon onthouden'), ander *materiaal* in het proces betrekken ('eens kijken wat er over dit onderwerp nog meer verteld wordt in de volgende alinea') of een andere *methode* bedenken om het beoogde resultaat alsnog te realiseren ('ik heb zeker te snel gelezen; laat ik even teruggaan naar wat ik nog wel begreep en daar opnieuw beginnen, maar dan

langzamer'). Iemand die in een objectief vergelijkbare toestand van onbegrip niets opmerkt, mist de aanleiding om bij te sturen.

Er zijn sterke indicaties dat mensen verschillen in de mate waarin ze in staat zijn om te diagnostiseren in welk cognitieve toestand ze zich bevinden. Brown, Campione en Barclay (1978) en Flavell en Wellman, (1977) beschrijven jonge leerlingen die 'verrassend slecht' in staat bleken te discrimineren tussen ver uiteenliggende cognitieve toestanden, zoals het 'wel en niet kennen van een rijtje woorden' of het 'wel en niet begrijpen van een instructie'. Vergelijkbare verschijnselen zijn beschreven in verband met meer gevorderde lerenden. Het gaat dan echter niet om onderscheidingen tussen ver uiteenliggende cognitieve toestanden, maar om meer subtiele onderscheidingen (Van Oostendorp, 1988; Van Daalen-Kapteijns & Elshout-Mohr, 1989; Entwistle & Entwistle, 1992).

Of een lerende in staat is om de ene cognitieve toestand te onderscheiden van de andere is niet alleen een kwestie van leeftijd en ervaring, maar ook van gebruik van adequate metacognitieve strategieën. Lerenden kunnen doelbewust hun leerproces zo inrichten dat ze bepaalde aspecten van het eigen functioneren en de resultaten ervan goed in de gaten kunnen houden. Nadat ze bijvoorbeeld een aantal malen ervaren hebben dat het begrijpen van een tekst geen alles of niets kwestie is, kunnen ze maatregelen treffen en extra gaan letten op de kwaliteit van het tekstbegrip, bijvoorbeeld door zichzelf vragen te gaan stellen. Men spreekt in dat geval van 'comprehension-monitoring'. Gaat de lerende op grond van ervaringen vooral letten op wat 'onthouden' wordt of op wat er met de 'motivatie' gebeurt, dan kan men spreken van 'memory-monitoring' en van 'motivation-monitoring'. De overkoepelende term is 'cognitive monitoring'. Als Nederlandse term zullen we, zo af en toe, de term 'procesbewaking' gebruiken.

Vooruitlopend op het tweede deel van deze bijdrage waarin het gaat om training in metacognitieve strategieën, stellen we ten aanzien van 'cognitive monitoring' drie dingen vast. In de eerste plaats kunnen de *aspecten* van het cognitieve functioneren die door een persoon bewaakt worden, verschillen. Het hangt bijvoorbeeld af van de omstandigheden of een lerende erop let of hij informatie 'begrepen' heeft of 'onthouden heeft' of beide. In de tweede plaats zijn er individuele verschillen in het vermogen tot diagnostiseren van eigen cognitieve toestanden en in het vaststellen van verschillen tussen toestanden. Deze vermogens zijn van belang voor het effectief kunnen bepalen en bijsturen van eigen activiteiten, leerdoelen en middelen. In de derde plaats voeren lerenden de kwaliteit van de procesbewaking op door gebruik te maken van metacognitieve strategieën. De mate waarin ze dat doen verschilt van persoon tot persoon, maar ook per situatie. Garner (1990) behandelt een aantal omstandigheden waaronder mensen er (meer in het algemeen) van afzien om strategieën in te zetten waar ze wel over beschikken. De aanleiding kan zijn dat ze menen 'hetzelfde' ook te kunnen bereiken met eenvoudiger middelen of dat ze de meerwaarde die strategiegebruik zal opleveren niet de moeite waard vinden. Ook komt het dikwijls voor dat mensen er eenvoudig niet aan denken om een strategie in te zetten, terwijl ze achteraf constateren dat ze dit wel hadden moeten doen.

Met deze bespreking van 'cognitive monitoring' en de factoren die bijdragen tot verschillen in de kwaliteit van de bewaking van eigen cognitieve processen, is het eerste deel van deze bijdrage afgerond. Besproken zijn de begrippen metacognitieve gewaarwording, metacognitieve kennis, metacognitieve strategie en 'cognitive monitoring' en de verschillen die er op deze punten tussen lerenden bestaan. In het tweede deel van deze bijdrage gaat het om training. Kan men studeergedrag van studenten beïnvloeden door hen metacognitieve kennis en strategieën bij te brengen en op welke gronden mag men verwachten dat de interventies tot verbetering van de studieprestaties zullen leiden?

#### ACHTERGROND VAN DE VRAAGSTELLING OMTRENT DE TE VERWACHTEN EFFECTEN VAN INTERVENTIES

Het metacognitieve functioneren van lerenden is een gewild aanknopingspunt voor interventies. We willen de opzet en uitgangspunten van deze interventies die dikwijls plaatsvinden onder de

(meeromvattende) noemer 'studievaardigheidstraining', bespreken en nagaan waarop men de positieve verwachtingen omtrent de mogelijke opbrengst ervan baseert. Er is aanleiding om deze kwestie eens kritisch te analyseren, omdat er tegenstrijdigheden zijn. Aan de ene kant constateren we dat metacognitie de laatste jaren steeds hoger komt te staan op de lijst factoren die bepalend zouden zijn voor de hoeveelheid en kwaliteit van leerresultaten (Wang, Haertel & Walberg, 1990; Van Hout-Wolters, 1992). Aan de andere kant komt er weinig tot geen verandering in de praktijkbevindingen omtrent 'studievaardigheidstrainingen' en de effecten ervan. Het succes van de trainingen blijft sterk wisselend en het inzicht in de factoren die het succes verklaren blijft oppervlakkig. Men is het eigenlijk slechts over twee dingen eens: de training moet *niet kort* duren, maar 'lang genoeg'; en de training moet *niet* plaatsvinden *buiten* de onderwijs-leeromgeving waarbinnen de studievaardigheden moeten worden toegepast, maar daarin 'geïntegreerd' zijn (Derry & Murphy, 1986, Pressley, 1986). Wat 'lang genoeg' is en wat 'integratie' in de onderwijs-leeromgeving inhoudt, blijft echter in het vage, evenals talloze andere kwesties ten aanzien van de trainingsdoelen, de doelgroepen en de trainingsprocedures.

Gegeven de huidige stand van zaken, heeft men de keuze uit twee stellingen. De eerste houdt in dat men van trainingen met een metacognitief karakter terecht veel verwacht en dat de gesignaleerde stagnatie in effectverbetering optreedt omdat nog zoveel studievaardigheidstrainingen dit karakter missen. De tweede houdt in dat het inmiddels weliswaar tamelijk gebruikelijk is om in studievaardigheidslessen aandacht te besteden aan metacognitieve kennis en strategieën, maar dat men te vanzelfsprekend verwacht dat langs deze weg succes kan worden geboekt. In wat nu volgt bekijken we de aannemelijkheid van het tweede deel van de tweede stelling: de vanzelfsprekendheid waarmee men succes verwacht. Daartoe schetsen we eerst de algemene gang van zaken in veel studievaardigheidstrainingen waarin aandacht wordt gegeven aan metacognitie en zelfregulatie. Vervolgens gaan we kritisch in op de evidentie waarop het optimisme omtrent de mogelijke effecten van zo'n interventie is gebaseerd.

#### STUDIEVAARDIGHEIDSTRAINING MET EEN ACCENT OP METACOGNITIE EN ZELFREGULATIE

Het denken over studievaardigheidstrainingen is voortdurend in beweging. Dit blijkt uit de bundels die het Landelijke Overleg Studievaardigheden uitgeeft na afloop van de jaarlijkse Landelijke Dag Studievaardigheden (Slobbe, Blankenstijn, Zeeuwen & Willemsen, 1990). Een van de typen training waaraan aandacht wordt besteed, heeft tot doel de cursisten metacognitieve kennis bij te brengen en te oefenen in het gebruik van zelfregulatie-strategieën. Processen die door de cursisten zelfstandig gereguleerd en uitgevoerd moeten worden, zijn bijvoorbeeld 'tekstbestuderen' en 'de studie plannen'. Om de regulatie te bevorderen, worden aan lerenden vooral die methoden voor het bestuderen van teksten en het maken van plannen aangeboden die als kenmerk hebben dat de lerenden volgens een bepaalde *systematiek* werken en dat zij tussendoor *controleren* hoe het proces verloopt en wat het oplevert. Dergelijke strategieën hebben, zoals eerder besproken, een dubbelfunctie. Zij bevorderen de werkuitvoering *en* versterken de controle op het proces en de resultaten. Veel nadruk legt men tijdens dit type studievaardigheidstraining op het feit dat persoons-, taak- en leeromgeving-kenmerken van belang zijn voor de strategieën die door de lerende gekozen en gehanteerd moeten worden. Zo wordt duidelijk een beroep gedaan op de metacognitieve kennis van de deelnemers en hierover wordt ook van gedachten gewisseld, bijvoorbeeld naar aanleiding van een door de deelnemers ingevulde studeergedragvragenlijst. (Oosterhuis-Geers, 1991; Elshout-Mohr, van Daalen-Kapteijns & Stawski, 1992).

De vraag die ons in wat nu volgt bezig houdt, is niet of het hier geschetste type training op veel plaatsen in Nederland of daarbuiten is ingeburgerd, maar wat voor mensen aanleiding zou kunnen zijn om er positieve effecten van te verwachten. Bij het bespreken van de voor deze vraag relevante literatuur, gebruiken we zowel de term zelfregulatiestrategie als de term metacognitieve strategie. De term 'metacognitieve strategie' verwijst van dit moment af niet meer

exclusief naar het begrip dat we eerder behandelden op basis van het werk van Flavell. De betekenis is ofwel even algemeen als die van het begrip zelfregulatiestrategie (= 'strategie met een procesbegeleidend karakter') of de meer precieze betekenis wordt ter plekke op een of andere wijze nader bepaald.

## DE SUCCESVERWACHTING

Drie typen onderzoek kunnen ten grondslag liggen aan de verwachting dat training in zelfregulatie-strategieën een gunstig effect zal hebben op leerprestatie. Het eerste type onderzoek laat zien dat zelfregulatie een belangrijke *component* is van het effectief en doelmatig uitvoeren van complexe leer- en probleemoplostaken. Het tweede type onderzoek toont aan dat er een *correlatieve verband* is tussen de kwaliteit van zelfregulerende activiteiten (inclusief het gebruik van metacognitieve strategieën) en de hoeveelheid en kwaliteit van leerprestaties. Het derde type onderzoek tenslotte laat zien dat *training* in zelfregulatiestrategieën in de praktijk mogelijk is en dat er goede resultaten mee zijn geboekt.

### 1. Metacognitie als component van effectief en doelmatig uitvoeren van complexe leer-en probleemoplostaken

Het belang van metacognitie als component van effectief en doelmatig handelen komt naar voren in *beschrijvend* onderzoek waarin mensen worden geconfronteerd met complexe taken die tot op zekere hoogte nieuw voor hen zijn. In het onderzoek laat men de 'proefpersonen' de taak naar eigen goeddunken uitvoeren. Voor zover er hulp beschikbaar wordt gesteld, is het gebruik daarvan aan strikte regels gebonden. De onderzoeker voert eerst een grondige taakanalyse uit en observeert, analyseert en beschrijft het gedrag van verschillende mensen die de taak uitvoeren. Door de proefpersonen 'hardopdenkend' te laten werken, verkrijgt de onderzoeker aanvullende gegevens over de inhoud van het werkgeheugen, metacognitieve gewaarwordingen, gebruik van metacognitieve kennis en strategieën.

Tot de 'metacognitieve' gedragscomponenten worden in dit soort onderzoek alle activiteiten gerekend die het gedrag sturen, reguleren en begeleiden. Ook *gewaarwordingen, gevoelens en gedachten* die een diagnostische of regulerende functie hebben en door de onderzoeker geconstateerd worden, kunnen als metacognities worden aangemerkt. Of de lerende zelf bewuste aandacht aan metacognities besteedt, is voor de onderzoeker *niet* van doorslaggevend gewicht als het gaat om de vraag of metacognitie een rol speelt. Het is namelijk bekend dat de bewuste aandacht die een lerende besteedt aan de besturings-activiteiten afneemt naarmate de persoon meer ervaring heeft met het besturen van een bepaalde complexe taak (Anderson 1985; Elshout, 1990). Dat de besturing meer autonoom plaatsvindt, maakt het proces niet minder doelgericht, beheerst en gecontroleerd en behoeft voor de onderzoeker geen aanleiding te zijn om niet meer te spreken van metacognitieve sturing door de lerende zelf.

De procesbegeleidende activiteiten worden onder andere aangeduid met de termen 'metacognitieve regulatie-activiteiten' (Vermunt, 1989) en 'hogere-orde-processen' (Ippel & Elshout, 1990). *Metacognitieve regulatie* wordt omschreven als "exerting control over one's own cognitive and affective processing of learning contents" (Vermunt, 1989). De term *hogere-orde processen* geeft aan dat men denkt in termen van een functionele hiërarchie. Daarin zijn processen van lagere orde functioneel afhankelijk van begeleidende (hogere-orde-) processen in die zin dat lagere-orde-processen slechts goed kunnen verlopen als de hogere-orde-processen ook goed verlopen (Elshout, 1990). Dat metacognitie een belangrijke component is van leren en probleemoplossen komt tot uiting in verschillen in de gedragsregulatie van gevorderden en beginners en van goed en zwak presterende 'proefpersonen'. Elshout (1983) beschrijft bijvoorbeeld dat 'goede beginners' zich van 'minder goede' onderscheiden in de manier waarop ze het 'beginner'-zijn verdisconteren in de manier waarop ze het werk aanpakken. 'Goede beginners' zullen het werk eerder zo inrichten dat ze misschien niet zo snel klaar zijn maar wel veel kans hebben om van hun ervaringen te leren, terwijl minder goede beginners hun onervarenheid eerder zien als een alibi om 'op goed geluk' of 'stuurloos' te handelen.



## 2. Vragenlijst-onderzoek naar het verband tussen metacognitie en prestatie

Een tweede type onderzoek waaruit kan worden afgeleid dat metacognitie belangrijk is voor de kwaliteit van prestaties maakt gebruik van het feit dat men aan mensen (mondeling of schriftelijk) vragen kan stellen over hun metacognitieve kennis en over de wijze waarop ze hun gedrag plannen, reguleren en bewaken. De gegevens die zo worden verkregen, brengt men vervolgens in verband met de prestaties.

Wat leerlingen of studenten over de beste inrichting van studeerprocessen weten, toetst men in dit type onderzoek meestal niet door middel van kennis- en toepassingsvragen, zoals "Welke drie planningsmethoden ken je?" of "Maak op drie manieren een planning voor...". In de regel ontlokt men ondervraagden uitspraken over het eigen studeergedrag om vervolgens de antwoorden te compileren tot scores die indicatief zijn voor de kwaliteit van de zelfregulatie. Een bevinding die in dit type onderzoek steevast terugkomt, is dat de kwaliteit van de gerapporteerde zelfregulatie-activiteiten positief gerelateerd is aan het bereiken van goede studieresultaten (Zimmerman & Pons, 1986; Wang, Haertel & Walberg, 1990). Om een indruk te geven van wat Zimmerman en Pons onder zelfregulatie-activiteiten verstaan, volgen hier enkele voorbeelden: stellen van subdoelen; inwinnen van informatie; herhalen en inprenten; hulp vragen bij de docent, leeftijdgenoten en anderen; aantekeningen, teksten en leerboeken achteraf of vooraf in hun geheel doorkijken.

## 3. Het trainingsonderzoek van Brown

Verscheidene trainingsonderzoeken werden uitgevoerd door Brown en haar medewerkers. De onderzoeken hebben als gemeenschappelijk kenmerk dat lerenden die tekortschoten in een bepaald facet van de sturing van hun (studeer)gedrag door middel van training werden bijgespijkerd, met als gevolg dat hun prestaties verbeterden. Zo werden (moeilijk lerende) leerlingen geoefend in het zelfstandig stellen en beantwoorden van de vraag "Ken ik dit rijtje of ken ik het niet?" (Brown, Campione & Barclay, 1978). Daarmee werd het probleem opgelost dat zij op volstrekt willekeurige momenten meenden dat ze een hen opgedragen inprenttaak volbracht hadden en overhoord konden worden. Ook bij lezers die zwak presteerden op het punt van begrijpend lezen werd succes geboekt met training in metacognitieve strategieën (Campione, Brown & Ferrara, 1985; Palincsar & Brown, 1984).

Het meest succesvol zijn de interventies wanneer aan drie voorwaarden wordt voldaan. In de eerste plaats dient de training strategieën te betreffen die zowel een cognitieve (procesuitvoerende) als een metacognitieve (procesbegeleidende) functie vervullen. Een voorbeeld is de strategie van 'repereren en hardop herhalen van subgroepen van woordjes' als middel om de woordjes te leren en te controleren of je ze kent en kunt opnoemen. In de tweede plaats is wenselijk dat aan de lerenden inzicht wordt gegeven in wat zij moeten leren en in de reden waarom de strategie hen kan helpen om het gesignaleerde probleem te overwinnen. Er wordt gesproken van 'informed training'. Ten derde is belangrijk dat lerenden oefening krijgen in het zelfstandig sturen en reguleren van het gedrag tijdens het gebruik van de strategie. Dit aspect van de training noemt men 'self-control-training'. De kans dat de deelnemers aan het interventieprogramma van de geleerde strategie gebruik blijven maken, blijkt in hoge mate afhankelijk te zijn van de mate waarin ze geleerd hebben om zelfstandig de talloze kleine beslissingen te nemen die bij effectief gebruik van een strategie horen. Te nemen beslissingen zijn bijvoorbeeld: welke 'groepjes woorden' neem je samen; wat doe je met twee woordjes die overblijven als je steeds groepjes van drie genomen hebt; wat doe je als je een fout in één van de drie woordjes uit een groepje maakt, enzovoort. De instructievorm die is ontwikkeld om dit type training te geven wordt 'reciprocal teaching' genoemd. De benaming geeft aan dat de leerlingen om de beurt de rol van 'onderwijzer' vervullen, hetgeen inhoudt dat zij, volgens een vast stramien, vragen stellen en instructies geven.

Mag men op grond van het welslagen van de training-experimenten van Brown c.s. concluderen dat metacognitie ook bij niet zwakke leerlingen van belang is voor het leren? Brown meent van wel. Zij ziet het als een specifiek kenmerk van moeilijk lerende leerlingen dat men voor hen *expliciet* moet maken wat bij anderen *impliciet* kan blijven. Niet alleen de moeilijk



lerenden, maar ook de overige leerlingen moeten metacognitieve strategieën ontwikkelen om te kunnen weten wanneer 'ze aan een overhoring toe zijn' of wanneer ze 'begrepen hebben waar een stukje tekst over gaat'. Zonder adequate metacognitieve kennis en oefening in zelfregulatievaardigheden zouden ook snel lerende leerlingen in de problemen raken. Van de betreffende leer- en oefenprocessen zie je bij deze kinderen echter minder omdat zij voldoende hebben aan instructies die in veel opzichten onvolledig zijn en omdat ze veel oppikken uit incidentele ervaringen. Bij de zwakkere leerlingen heeft de onderzoeker meer kans om te ontdekken hoe ingewikkeld het leerproces is en hoe belangrijk het is dat leerlingen er desondanks in slagen om de metacognitieve kennis en vaardigheden te ontwikkelen.

### HET GEVAAR VAN OVERSCHATTING VAN WAT MEN VERWACHTEN MAG VAN TRAINING IN METACOGNITIE

Afgaand op wat de drie bovengenoemde lijnen van onderzoek opleveren, kan men gemakkelijk tot de conclusie komen dat men heel wat verwachten mag van training in metacognitie. Uit het beschrijvende onderzoek naar het cognitieve functioneren van 'gevoorderden en beginners' en 'goede en minder goede beginners' valt af te leiden dat er veel ruimte is voor verbetering van de aanpak van mensen die men (zoals in onderwijsleersituaties voortdurend gebeurt) stelt voor complexe, nieuwe taken. Voorts suggereren de correlatieve gegevens uit vragenlijst-onderzoek dat verbetering van de zelfregulatie zal samengaan met verbetering van de prestaties. En tenslotte toont het succes van de training-experimenten dat het aanleren van effectieve metacognitieve strategieën zelfs bij zwakke leerlingen mogelijk is. Daaraan kan men gemakkelijk de conclusie verbinden dat wat tot succes leidt bij zwakke leerlingen het zeker niet minder goed zal doen bij de sterkere leerlingen, zodat men in deze vorm van training een instrument heeft met een breed bereik.

Omdat we echter bevreesd zijn voor een te groot optimisme, beschouwen we in de komende paragrafen de onderzoeksliteratuur nog iets nader op kritische wijze. Deze beschouwing zal leiden tot de conclusie dat de bevindingen uit de besproken drie lijnen van onderzoek om twee redenen niet *zonder meer* kunnen worden gezien als indicatie dat training in metacognitieve strategieën in het algemeen tot positieve resultaten zal leiden. In de eerste plaats blijkt dat men op problemen stuit als men ervan uitgaat dat de trainingsvorm die Brown ontwikkelde (en varianten daarop) ook geschikt zouden zijn voor leerlingen die *niet* vanuit een achterstandssituatie aan een studievaardigheidstraining beginnen. In de tweede plaats blijkt het feit dat 'gevoorderden' en 'goede beginners' metacognitieve strategieën van goede kwaliteit gebruiken, niet *zonder meer* te impliceren dat het aanbrengen van vergelijkbare metacognitieve kennis en zelfregulatievaardigheden tot iets goeds zou leiden bij minder gevorderden en 'onverstandige' beginners.

Inzicht in de noodzaak om de succesverwachting te temperen, gaat samen met het scherper onderscheiden van de doelen die men met training mogelijk wel zou kunnen kan realiseren. Een korte bespreking daarvan vormt de afsluiting van het artikel.

### VOORWAARDEN VOOR SUCCES VAN GEÏNTEGREERD OEFENEN VAN METACOGNITIEVE EN COGNITIEVE STRATEGIEËN

Uit het onderzoek van Brown c.s. blijkt de effectiviteit van training in zelfregulatiestrategieën met betrekking tot algemene basisvaardigheden, zoals (begrijpend) lezen, informatie inprenten en luisteren. Een van de kenmerken van de training is dat er gedurende een periode van enkele weken regelmatig geoefend wordt met 'echte opdrachten' op het niveau waarop de lerende moet leren werken. De lerende moet ervaring opdoen in het zelfstandig sturen van de handelingen die met telkens kleine varianten moeten worden uitgevoerd ('self-control-training'). Strikt en onmiddellijk terugkoppelen op zowel de *sturing* als de *uitvoering* van activiteiten is in deze fase

essentieel. De afstemming van sturing en uitvoering moet goed zijn en de kwaliteitscontrole moet streng zijn. In eerste instantie lijkt slechts één probleem in de weg te staan van een succesvol gebruik van deze methode bij zowel zwakke als goede leerlingen. De methode is arbeidsintensief en dus kostbaar. Er is echter een tweede probleem dat pas naar voren komt als men de methode wil gebruiken bij lerenden die niet vanuit een duidelijke achterstandssituatie beginnen. Het blijkt dan namelijk merkwaardig moeilijk te zijn trainers te vinden die worden geaccepteerd als 'kenner en voorbeeld van studievast gedrag' en het blijkt, mede daardoor, moeilijk om omstandigheden te creëren waarin de beoogde cursisten bereid zijn actief deel te nemen aan een intensieve training met strikte kwaliteitscontrole. Collins, Brown en Newman (1989) schetsen een leeromgeving waarbinnen men het probleem omzeilt; onder andere door toegangseisen te stellen en te zorgen dat vorderingen maken een eerste vereiste is voor elke lerende. Het 'cognitive apprenticeship'-systeem, dat zij voorstellen heeft namelijk als uitgangspunt dat alleen mensen die 'als gezet' opgeleid *willen* worden door een 'meester' (die zij ook inderdaad zien als 'expert' en competent tot het leveren van kwaliteitsoordelen) aan het onderwijs deelnemen; en verder is het onderwijs zo georganiseerd dat de (sociale) status en de werkzaamheden die lerenden opgelegd krijgen direct afhankelijk zijn van hun vorderingen. Een ander voorbeeld van een situatie waarin studenten optimaal meewerkten, werd beschreven door Biggs (1985). Het betrof studenten die vanuit de studeeromgeving waarin zij zich bevonden, zelf de behoefte ontwikkeld hadden aan diepe verwerkingsstrategieën. Zij namen deel aan een studievastheidstraining die op de ontwikkeling van deze (zelfregulatie-) strategie was gericht, hadden er veel voor over zich deze eigen te maken en staken beduidend meer op dan de 'gemiddelde' cursist. De (voorlopige) conclusie is duidelijk. De methode die Brown ontwikkelde is belangrijk en werkt onder omstandigheden die veelvuldig voorkomen in speciaal onderwijs en in 'remedial teaching'; binnen andere vormen van onderwijs zijn de vereiste omstandigheden echter moeilijk te realiseren. Het is daarom niet goed mogelijk om de successen van Brown te zien als een aanwijzing dat men (ook) met gemiddelde studenten succes zal boeken door hen te stimuleren tot gebruik van goede zelfregulatiestrategieën.

Er is echter een tweede punt dat we naar aanleiding van het trainingsmodel van Brown aan de orde willen stellen en dat is de aard van het beoogde resultaat. Brown zelf gaat ervan uit dat de strategieën die de deelnemers aan de trainingen leren gebruiken, deel uit gaan maken van hun repertoire. Dezelfde zelfregulatie-strategieën die zij tijdens de training oefenen, zullen zij ook na de training gebruiken. Wat op den duur verandert, is de mate van bewuste controle die nodig is om het proces in goede banen te leiden, maar het regulatieproces als zodanig verandert niet. Heeft de lerende bijvoorbeeld, als onderdeel van de training in begrijpend lezen geleerd om voor zichzelf aan het eind van elke alinea kort samen te vatten wat de belangrijkste informatie was, dan wordt die activiteit een vast onderdeel van wat de lezer *ook na de training* doet tijdens 'begrijpend lezen'. Een iets andere visie volgt uit het werk van Resnick (1984). Zij maakt een onderscheid tussen strategieën, die tijdens de training een belangrijke rol spelen en strategieën die lerenden na afloop van de training opnemen in hun repertoire. De strategieën die tijdens de training worden gebruikt stellen lerenden in staat om in bepaalde 'cognitieve toestanden' terecht te komen die nieuw voor hen zijn en aanleiding vormen tot 'nieuwe' metacognitieve gewaarwordingen en opbouwen van 'nieuwe' metacognitieve kennis. Wat lerenden vervolgens met de opgedane ervaringen en kennis doen, is in de visie van Resnick vers twee. Een lerende die tijdens de training ervaart dat er een subtiel, maar 'waarneembaar' onderscheid is tussen het soort tekstbegrip dat voortkomt uit 'informatie gelezen hebben' en het soort tekstbegrip dat ontstaat als de informatie na het lezen ook nog wordt samengevat, heeft iets geleerd dat hij of zij op diverse manieren kan gebruiken. De ontdekking van de twee gradaties van tekstbegrip en hun samenhang met de twee manieren van 'studeren', kan het begin zijn van een reeks ontdekkingen omtrent gradaties in tekstbegrip die tot stand gebracht kunnen worden door verschillende activiteiten uit te voeren. Na afloop van de training zullen er lerenden zijn die de tijdens de training geoeffende werkwijze (of iets wat daar veel op lijkt) blijven gebruiken, maar er zullen ook lerenden zijn die daarnaast of in plaats daarvan de nieuwe ervaringen *op hun eigen wijze* benutten en tot onderdeel maken van een eigen metacognitieve strategie.

In deze alternatieve opvatting omtrent het doel van de training, ligt misschien een mogelijkheid om in het kader van studievaardigheidstrainingen toch op het succes van Brown voort te bouwen. Stel dat men zich houdt aan de aanbeveling om lerenden te informeren over het nut van de activiteiten die men hen laat uitvoeren, dan legt men in het 'klassieke' geval de nadruk op het aanleren en inslijpen van goede zelfregulatiegewoontes. Wat goed is staat niet ter discussie en wie mee wil doen accepteert de aanwijzingen. Wordt gekozen voor het 'alternatief' dan legt men er het accent op dat de oefeningen hulpmiddelen zijn en dat het gaat om exploratie van het domein van cognitieve toestanden (van begrip en onbegrip, onthouden en vergeten, gemotiveerd en ongemotiveerd zijn) en op de activiteiten die tot het ontstaan van de diverse toestanden bijdragen. Wat ieder voor zich zal ontdekken, hangt in dit geval niet alleen af van de aard van de strategieën waarmee tijdens de training wordt geoefend, maar ook van het vermogen van lerenden tot 'cognitive monitoring' en reflectie. Op de implicaties die de keuze voor het alternatief heeft voor de opzet, inrichting en duur van de zelfregulatie-training gaan we hier niet in. We zien echter wel reden om de eerder getrokken conclusie bij te stellen. De succes-stories van Brown zullen, in de gemiddelde studievaardigheidscursus, niet *zonder meer* kunnen worden gecontinueerd, maar er kan wel op worden voortgebouwd. Het werken met zelfregulatie-strategieën tijdens de cursus kan verschillende functies vervullen, die voor verschillende doelgroepen aantrekkelijk kunnen zijn. Door diverse cursusvormen te hanteren en aanstaande cursisten te informeren over het beoogde leerproces en leerresultaat, zou het succes kunnen worden vergroot; met name bij gemiddelde studenten, die onder gewone omstandigheden deelnemen aan de training.

#### VERWEVENHEID EN DISSOCIATIE VAN COGNITIE EN METACOGNITIE

Leiden aanvullingen in metacognitieve kennis en gebruik van metacognitieve strategieën altijd tot verbetering in leerprestaties of hangt dit af van de voorkennis en cognitieve vaardigheden waarover de persoon verder nog beschikt, als gevorderde of al of niet verstandige beginner? Dat is de tweede kwestie die we nader wilden bezien. Daartoe bekijken we eerst de relatie tussen metacognitieve en cognitieve componenten van het studeergedrag en de bijdragen van beide componenten aan de leerprestaties. Daarbij moeten we rekening houden met het feit dat metacognitie en cognitie niet onafhankelijk van elkaar ontwikkeld worden. Cognitieve en metacognitieve ontwikkelingen gaan in belangrijke mate gelijk op, maar niet voor honderd procent. Er is naast verwevenheid ook een zeker mate van 'metacognitieve dissociatie'; het komt voor dat metacognitieve ontwikkelingen achterblijven bij wat er op cognitief vlak gebeurt of er juist op vooruitlopen.

Doordat metacognitie met cognitie verweven is, kan dikwijls moeilijk worden nagegaan of bepaalde gedragingen en de daaruit voortvloeiende (leer)prestaties moeten worden toegeschreven aan cognitieve factoren of aan metacognitieve. Het (ons uit het eerste deel van deze bijdrage bekende) meisje dat bewust een onthoud-strategie hanteert, kan weer dienen als voorbeeld. Zij voert cognitieve activiteiten uit die resulteren in competent gedrag en zij begeleidt, bewaakt en stuurt het eigen gedrag. In zowel de taakuitvoering als de procesbegeleiding verschilt dit meisje van kinderen die zich van geen onthoudprobleem bewust zijn. Haar prestatie is de resultante van een ontwikkelingsproces waarin zij cognitieve activiteiten heeft leren uitvoeren en waarin zij de kans heeft aangegrepen om op basis van haar ervaringen metacognitieve kennis op te bouwen en te verwerken in een metacognitieve strategie. Men kan wel stellen dat voor de kwaliteit van de prestatie de procesbegeleiding van wezenlijk belang is, maar men moet zich tegelijkertijd realiseren dat de metacognitieve kennis en strategieën niet de plaats innemen van de cognitieve processen, maar er als 'hogere orde'- processen aan toegevoegd zijn.

Als men een indruk wil krijgen van *de specifieke bijdrage* van metacognitie aan competent gedrag moet eigenlijk gekeken worden naar mensen die verschillen in metacognitief opzicht terwijl zij in cognitief opzicht met elkaar vergelijkbaar zijn. Die situatie kan bijvoorbeeld ontstaan als sommige lerenden, om welke reden dan ook, nalaten om aandacht te besteden aan

cognitieve toestanden en de verschillen die daarin optreden tijdens leerprocessen, met als gevolg dat er ook weinig metacognitieve kennis wordt opgedaan. Een andere mogelijkheid is dat sommige lerenden, om welke reden dan ook, sterk geneigd zijn om beschikbare metacognitieve kennis in te bedden in metacognitieve strategieën, met als gevolg dat deze kennis een rol gaat spelen bij het strategisch aanpakken van 'nieuwe' taken. In beide gevallen spreken we van 'metacognitieve dissociatie': in het eerste geval is de metacognitieve ontwikkeling tot op zekere hoogte van de cognitieve ontwikkeling losgemaakt door wat men kan zien als 'achterblijven' van de metacognitieve ontwikkeling; in het tweede geval is er juist sprake van 'voortuitlopen'.

Om te verhelderen waar de 'dissociatie' toe leidt, geven we twee voorbeelden van mensen die op cognitief (procesuitvoerend) vergelijkbaar zijn, terwijl ze op metacognitief niveau verschillen. Deze verschillen betreffen in het eerste geval de mate van bewuste procesbegeleiding ('monitoring') en in het tweede geval de aanpak (de metacognitieve strategie). Het eerste voorbeeld is ontleend aan August, Flavell en Clift (1984). Zij constateerden dat veel lezers gewend zijn om al lezende inconsistenties in de tekst op te lossen door zelf een ontbrekende schakel te bedenken, bijvoorbeeld in de vorm van een oorzaak of een verklaring. Binnen een groep lezers die dit doen, kan men lezers onderscheiden naar de mate waarin ze zich bewust zijn (en blijven) van het hypothetische karakter van de zelf verzonnen aanvulling. Wie in dit procesbegeleidende opzicht beter functioneert, presteert beter bij allerlei taken waarbij het er toe doet wat er *precies* in de gelezen tekst stond en wat eraan werd toegevoegd op basis van een eigen interpretatie. Onder deze omstandigheden blijkt dat het monitor-proces een aparte, zelfstandige, bijdrage levert aan de prestaties die de persoon levert.

Het tweede voorbeeld is ontleend aan Chi, Bassock, Lewis, Reimann en Glaser (1989). Zij gingen in detail na hoe beginnende en meer gevorderde natuurkunde-studenten probeerden wat te leren van het bestuderen van een uitgewerkt (voorbeeld-) vraagstuk. Hoewel mag worden aangenomen dat de meer gevorderde studenten niet minder begrepen van het voorbeeld dan de beginners, viel op dat zij vaker 'onbegrip' bij zichzelf constateerden. Zij hanteerden blijkbaar een (meer algemene) metacognitieve strategie waarin 'onbegrip' functioneerde als een belangrijk gegeven, een middel tot sturing van het eigen gedrag. Bijgevolg was er voor de gevorderde studenten vaker *aanleiding* om 'goed na te gaan of ze begrepen hoe een bepaalde stap in het voorbeeld genomen werd' en om de gegeven uitleg nog eens te lezen en voor zichzelf nader uit te werken. Dit kwam de toename van kennis en inzicht ten goede.

In studievvaardigheidstrainingen kan men inspelen op de beide vormen van 'metacognitieve dissociatie'. Men kan de training richten op lerenden die uit zichzelf weinig oog hebben voor de verschillende cognitieve toestanden die zich tijdens leerprocessen voordoen en die daardoor weinig metacognitieve kennis opbouwen. Men kan de training ook gebruiken om lerenden op het spoor te zetten van mogelijkheden om beschikbare metacognitieve kennis uit te werken tot metacognitieve strategieën met een meer algemene toepassingswaarde en om deze strategieën vervolgens ook te benutten. Beide mogelijkheden leiden tot een heel verschillende invulling van de training, zoals blijkt uit de onderstaande summierere uitwerkingen.

### **De aandacht richten op de verschillende cognitieve toestanden die zich tijdens leerprocessen voordoen**

Men kan mensen laten reflecteren op verschijnselen die zich voordoen tijdens leerprocessen en zorgen dat er voor hen aanleiding is zich bewust te worden van opvattingen over leren en studeren, waarvan ze vaak 'zelf niet weten waar ze vandaan komen en of ze eigenlijk wel kloppen'. De verwachting dat dit type training effect kan hebben, is onder meer gebaseerd op onderzoek van Van Someren en Elshout (1985) waaruit blijkt dat lerenden meer opsteken van eigen ervaringen tijdens het oplossen van problemen wanneer zij werken onder condities die bevorderen dat ze zich voortdurend bewust zijn van wat ze doen en wat het oplevert (zie ook Stoutjesdijk & Beishuizen, elders in dit nummer). Om de horizon van lerenden te verbreden kan men hen tijdens de training ook cognitieve activiteiten laten uitvoeren die leiden tot cognitieve toestanden die *nieuw* voor hen zijn. Voorbeelden van materiaal dat aanleiding geeft tot een grote verscheidenheid aan metacognitieve ervaringen en dat daardoor zeer geschikt is voor dit type



training, is onder andere ontworpen door Rauenbusch en Bereiter (1991). Voorbeelden van wat studenten zelf zoal weten te melden over hun verschillende toestanden van begrip en onbegrip, zijn te vinden bij Entwistle en Entwistle (1992). In hun onderzoeksverslag wordt melding gemaakt van verschillende gradaties van de 'diepte' van begrip (depth of understanding), maar ook van de 'breedte' van het begrip (breadth of understanding) en van de 'structuur' van het begrip dat door studenten wordt verkregen in de vakinhouden die zij bestuderen voor een tentamen. Uit de bewoordingen die studenten kiezen, wordt duidelijk dat er voor veel van de fijnere gradaties van cognitieve toestanden die studenten (moeten) kunnen onderscheiden, nog geen gangbare woorden en uitdrukkingen bestaan. Het arsenaal aan wat men 'cognitive internal state words' noemt (Hall, Scholnick, & Hughes, 1987), zal moeten worden uitgebreid om trainers de beschikking te geven over termen waarin zij met cursisten kunnen praten over de cognitieve toestanden die zij moeten gaan 'meemaken' en moeten leren onderscheiden. Daarnaast zullen trainers zich, voordat ze cursisten gaan werven, goed moeten bezinnen op de kennis die zij willen overdragen omtrent de peroons-, taak en strategie-kenmerken die kunnen bijdragen tot het bereiken van de diverse toestanden.

De grootte van het effect dat van de cursus verwacht mag worden, hangt af van de ernst van de schade die lerende ondervinden van hun geringe besef van wat er aan 'cognitieve toestanden' te bereiken valt en van de daarmee samenhangende geringe kennis omtrent de factoren die voor het bereiken van de diverse toestanden van belang zijn. Er zijn geen algemene uitspraken over te doen evenmin als over de tijd die voor de training moet worden uitgetrokken. Men hoeft er echter niet bij voorbaat van uit te gaan dat het wel weer zal gaan om enkele weken intensief trainen.

### **Metacognitieve kennis uitbouwen tot meer algemeen toepasbare strategieën**

Lerenden die in bepaalde omstandigheden ervaren dat bepaalde factoren tot voortgang van het cognitieve proces bijdragen, kunnen die kennis tot onderdeel maken van een metacognitieve strategie. Van de verschillende manieren om te bevorderen dat dit gebeurt, bespreken we er slechts één, een schijnbaar eenvoudige die echter veel omvat. De werkwijze houdt in dat men mensen die minder goed zijn in leren en studeren laat kennis maken met de algemene (niet met vakken en domeinspecifieke vaardigheden verweven) opvattingen en strategieën van mensen die er beter in zijn.

Wat zijn algemene 'het leren bevorderende' opvattingen en strategieën waarin mensen die 'goed leren' zich onderscheiden van anderen? Opmerkelijk is de consensus die hierover bestaat in een aantal recente publicaties (Resnick, 1989; Bereiter 1990; Pressley, Borkowski & Schneider 1989, VanLehn, 1991). VanLehn (1991) vat zijn bevindingen als volgt samen: "In short, it may be that good students are good because they assume that they are ignorant, they want to become less ignorant, and they know how to solve the ignorance problem in a variety of instructional settings" (VanLehn, 1991, p.39). De eerste twee punten die VanLehn hier noemt, zijn in het voorgaande al herhaaldelijk ter sprake geweest. Om goed te kunnen leren moeten lerenden onderkennen dat zij nog wat kunnen leren (van anderen die ergens meer van weten of beter in zijn) en dat het de moeite waard is om daar moeite voor te doen (hetzij door zelf dieper in leerstof door te dringen, hetzij door actief deel te nemen aan een training). Het laatste punt is nog niet met zoveel woorden ter sprake gekomen. Als VanLehn spreekt over het oplossen van het 'onwetendheid-probleem', doelt hij op metacognitieve strategieën die (zoals voor zoveel metacognitieve strategieën geldt) zowel een procesbesturende als een procesuitvoerende functie vervullen. VanLehn noemt als voorbeeld van een strategie die geschikt is om bepaalde onwetendheid-problemen op te lossen de 'wetenschappelijke onderzoeksmethode', die bestaat uit het systematisch stellen en toetsen van hypotheses. Goede lerenden gebruiken deze methode bij het bestuderen van complex materiaal frequenter dan minder goede en zij hebben er groot voordeel van. Gebruik van de methode leidt tot planmatig gedrag en bevordert dat de lerende zelf begrijpt wat hij doet en toezicht houdt op de vorderingen. Een ander voorbeeld van een methode die helpt om onwetendheid te ontdekken en te bestrijden, is die van het systematisch stellen en beantwoorden van de vragen 'hoe, wat, wanneer en waarom?' bij het doornemen van een praktische handleiding.



Een training in het opbouwen van 'het leren bevorderende'-opvattingen en strategieën (of kortweg 'leren leren') vergt eigenlijk een combinatie van alle eerder besproken typen trainingen. Lerenden moeten bereid zijn om zich de meer algemene methoden 'ter bestrijding van onwetendheid' eigen te maken. Dit houdt in dat zij, onder deskundige leiding en intensief, moeten oefenen in het hanteren van een strategie waarin metacognitieve procesbegeleiding en cognitieve procesuitvoering verweven zijn. Lerenden moeten daarnaast leren discrimineren tussen de toestanden van onwetendheid waarin de verschillende algemene methoden ter bestrijding ervan bruikbaar zijn. Eigen metacognitieve ervaringen moeten immers de aanleiding vormen tot het inzetten van de strategieën. In de derde plaats moeten lerenden gaan merken dat het dikwijls de moeite loont om eigen onwetendheid op te merken en te bestrijden. VanLehn (1991) noemt het een van de meest opvallende kenmerken van goed lerenden dat zij dikwijls actief proberen wat te leren op momenten waarop zij niet zichtbaar verkeren in een impasse. Blijkbaar gaan zij uit van een interne norm (vgl ook Bereiter, 1990) en nemen zij het initiatief tot leren zonder dat de situatie dit van hen verlangt. Dit derde vereiste is vanuit het gezichtspunt van training misschien wel het moeilijkst grijpbare.

De combinatie van de drie eisen rechtvaardigt de stelling dat 'leren leren' een proces van jaren is waaraan velen een bijdrage moeten leveren. Door het beschrijven van deeltaken wordt de kans dat de taken op produktieve wijze verdeeld zullen worden misschien wat groter.

#### DE BASIS VOOR SUCCES VAN TRAINING IN ZELFREGULATIESTRATEGIEËN

In dit artikel is betoogd dat op diverse manieren kan worden gewerkt aan de ontwikkeling van de metacognitieve kennis en zelfregulatievaardigheden van lerenden. Om duidelijkheid te verschaffen over wat men precies beoogt en om tot een realistische schatting te komen van de te verwachten effecten, is van belang dat expliciet wordt aangegeven wat de relatie is tussen metacognitie en cognitie, zowel bij de start van de training, als na afloop ervan. We lopen de drie gevallen die we hebben onderscheiden nog even na. Is het de bedoeling dat metacognitieve en cognitieve strategieën in onderlinge samenhang worden ontwikkeld? Dan kan men, vooral bij zwakke lerenden, goede resultaten verwachten van de intensieve trainingsvormen, die in het voorgaande beschreven zijn aan de hand van het werk van Brown. Gaat het om inlopen van een achterstand die is ontstaan op metacognitief vlak, terwijl er op het cognitieve vlak geen duidelijke achterstanden zijn? Dan biedt wellicht de weg van de betere bewustwording van eigen cognitieve toestanden goede perspectieven. Is het, tenslotte, de bedoeling om metacognitieve ontwikkelingen te laten vooruitlopen en te laten bijdragen tot toename van nieuwe kennis en cognitieve vaardigheden, dan vraagt dat weer een andere trainingsopzet. Er kan dan bijvoorbeeld gekozen worden voor een training waarin verschillende manieren om 'onwetendheid te bestrijden' en hun toepassingsmogelijkheden worden besproken. Door de verschillende gevallen conceptueel strikt te onderscheiden, kan voorkomen worden dat men, ten onrechte, bij elke afzonderlijke training succes op alle fronten verwacht. Bovendien helpt het bij het afstemmen van de trainings-opzet en -inhoud op verschillende doelgroepen en omstandigheden.

#### LITERATUUR

- Anderson, J.R. (1985). *Cognitive psychology and its implications*. New York: Freeman and Company.
- August, D.L., Flavell, J.H., & Clift, R. (1984). Comparison of comprehension monitoring of skilled and less skilled readers. *Reading Research Quarterly*, Fall, XXII, 39-53.
- Baker, L. & Brown, A.L. (1984). Metacognitive skills in reading, In P.D. Pearson (Ed.). *Handbook of reading research* (pp. 353-395). New York/London: Longman.
- Bereiter, C. (1990). Aspects of an educational learning theory. *Review of Educational Research*, 60 (4), 603-624.
- Biggs, J.B. (1985). The role of metalearning in study processes. *British Journal of Educational Psychology*, 55, 185-212.

- Boekaerts, M. (1991). Subjective competence, appraisals and selfassessment. *Learning and Instruction, Vol 1*, 1-17.
- Brown, A.L., Campione, J.C., & Barclay, C.R. (1979). Training self-checking routines for estimating test readiness: Generalization from list learning to prose recall. *Child Development, 50*, 501-512.
- Campione, J.C., Brown, A.L., & Ferrara, R.A. (1985). Mental retardation and Intelligence. In J.R. Sternberg (Ed.), *Handbook of human intelligence* (pp. 392-493). Cambridge: Cambridge University Press.
- Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.
- Collins, A., Brown, J.S. & Newman, S. (1989). Cognitive apprenticeship. In L.B. Resnick (Ed.), *Knowing, Learning and Instruction; Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Corno, L. (1986). The metacognitive control components of selfregulated learning. *Contemporary Educational Psychology, 11* (4), 333-346.
- Derry, S.J. & Murphy, D.A. (1986). Designing systems that train learning ability: From theory to practice. *Review of Educational Research, 56*, 1, 1-39.
- Dweck, C.S. & Leggett, E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273.
- Elshout, J.J. (1983). Een beginner is meer dan iemand die het nog niet kan. In Drenth e.a. (Eds.), *Psychologie in Nederland* (pp. 177-184). Amsterdam/Lisse: Swets & Zeitlinger.
- Elshout, J.J. (1987). Probleemoplossen als context voor leren probleemoplossen. *Nederlands Tijdschrift voor de Psychologie, 42*, 344-353.
- Elshout, J.J. (1990). Hogere-orde vaardigheden: discussie. In M.J. Ippel & J.J. Elshout (Eds.), *Training van hogere orde denkprocessen* (pp. 149-155). Amsterdam/Lisse: Swets & Zeitlinger.
- Elshout-Mohr, M. & van Daalen-Kapteijns, M.M. (1985). Van volgen van onderwijs naar zelfstandig studeren. In J.G.L.C. Lodewijks & P.R.J. Simons (Eds.), *Zelfstandig leren* (pp. 45-73). Amsterdam/Lisse: Swets & Zeitlinger.
- Elshout-Mohr, M. & van Daalen-Kapteijns, M.M. (1989). Self-regulated learning in the first year of higher education. In P. Span, E. de Corte, & B. van Hout-Wolters (Eds.), *Onderwijsleerprocessen* (pp. 47-56). Amsterdam/Lisse: Swets & Zeitlinger.
- Elshout-Mohr, M., Van Daalen-Kapteijns, M.M. & Stawski, S. (in druk). Expliciete aandacht voor eenvoudige denkvaardigheden binnen het hoger onderwijs. In Hoeks en anderen (in druk) *Studievaardigheden in kwaliteitsperspectief* (12e Landelijke Dag Studievaardigheden). Breda.
- Entwistle, A. & Entwistle, N. (1992). Experiences of understanding in revising for degree examinations. *Learning and Instruction, vol.2*, 1-22.
- Flavell, J.H. (1979). Metacognition and cognitive monitoring, A new area of cognitive-developmental inquiry. *American Psychologist, vol 34, 10*, 906-911.
- Flavell, J.H. & Wellman, H.M. (1977). Metamemory. In R.V. Kail & J.W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3-33). Hillsdale, N.J.: Erlbaum.
- Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. *Review of Educational Research, vol 60, 4*, 517-529.
- Greeno, J.G. (1988). *Mathematical Cognition*. Stanford, CA: Stanford University.
- Gruwel, F.L.J.H. (1992). Functionele geleterdheid: Een interventiestudie. Presentatie op de AIO/OIO dag, Amsterdam, februari 1992.
- Ippel, M.J. & Elshout, J.J. (Eds.), (1990). *Training van hogere orde denkprocessen*. Amsterdam/Lisse: Swets & Zeitlinger.
- Janssen, P.J. & De Neve, H. (1988). *Studeren en doceren aan het hoger onderwijs: vakmanschap als leeropdracht*. Leuven: Acco.
- Marton F., Hounsell, B., & Entwistle, N. (1984). *The experience of learning*. Scottish Academic Press.
- Ng, E. & Bereiter, C. (1992). Three levels of goal-orientation in learning. *The Journal of the Learning Sciences, 1, 3, 4*, 243-273.
- Nicholls, J.G. (1978). The development of the concepts of effort and ability, perception of own attainment and understanding that difficult tasks require more ability. *Child Development, Vol. 49*, 800-814.
- Nolen, S.B. (1988). Reasons for studying: Motivational orientations and study strategies. *Cognition and Instruction, 5*, 269-287.
- Oosterhuis-Geers, J. (1991). Formatieve evaluatie van PROBES. In W. Smit (1991) *Aansluitingsproblemen en Studievaardigheid* (210-231). Amsterdam: Huisdrukkerij VU.
- Palincsar, A.S. & Brown, A.L. (1984). Reciprocal teaching of comprehension fostering and comprehension monitoring activities. *Cognition and Instruction, 1*, 117-175.

- Pintrich, P.R. (1989). The dynamic interplay of student motivation and cognition in the college classroom. In Maehr, M. & Ames, C. *Advances in Motivation and Achievement: Motivation Enhancing Environments*, Vol. 6, 117-160. JAI Press.
- Prawat, R.S. (1989). Promoting access to knowledge, strategy, and disposition in students: a research synthesis. *Review of Educational Research*, 59, 1, 1-41.
- Pressley, M. (1986). The relevance of the good strategy user model to the teaching of mathematics. *Educational Psychologist*, 21 (1 & 2), 139-161.
- Pressley, M., Borkowski, J.G., & Schneider W. (1989). Good information processing: what it is and how education can promote it. *International Journal of Educational Research*, 13, 857-869.
- Rauenbusch, F. & Bereiter, C. (1991). Making reading more difficult: A degraded text microworld for teaching reading comprehension strategies. *Cognition and Instruction*, 8, 2, 181-206.
- Resnick, L.B. (1984). Comprehension and learning: Implications for a cognitive theory of instruction. In H. Mandl, N.L. Stein & T. Trabasso (Eds.). *Learning and comprehension*. Hillsdale/New York: LEA Ass.
- Resnick, L.B. (1989). Introduction. In Resnick, L.B. (Ed.). *Knowing, Learning and Instruction, Essays in honor of Robert Glaser*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Rothkopf, E.Z. (1988). Perspectives on study skills training in a realistic instructional economy. In C.E. Weinstein, E.T. Goetz, & P.A. Alexander (Eds.). *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 275-286). San Diego: Academic.
- Schliemann, A.D. & Acioly, N.M. (1989). Mathematical knowledge developed at work: The contribution of practice versus the contribution of schooling. *Cognition and Instruction*, 6 (3), 185-221.
- Schmeck, R.R., (Ed.). (1988). *Learning Strategies and Learning Styles*. New York: Plenum.
- Slobbe, F.J., Blankenstijn, J. Zeeuwen, O.S.C. & Willemsen, T.J.M. (1990). *Toekomstige ontwikkelingen. LDS'90 Eindhoven*.
- Snow, R.E. (1990). Aptitude, instruction, and individual development. *International Journal of Educational Research*. Vol. 13, 869-883.
- Van Daalen-Kapteijns, M.M. & Elshout-Mohr, M. (1989). Tekstgebonden leercriteria en zelfregulatie. In: P. Span, E. de Corte, & B. van Hout-Wolters (Red.). *Onderwijsleerprocessen. Strategieën voor verwerking van informatie* (pp. 85-93). Amsterdam/Lisse: Swets & Zeitlinger.
- Van Hout-Wolters, B.H.A.M. (1992). Cognitieve Strategieën als onderwijsdoel. Groningen: Wolters-Noordhoff.
- VanLehn, K. (1991). Rule acquisition events in the discovery of problem-solving strategies. *Cognitive Science*, 15, 1-47.
- Van Oostendorp, H. (1988). Regulatieprocessen bij tekstverwerking. Academisch proefschrift, Universiteit van Amsterdam.
- Van Someren, M.W. & Elshout, J.J. (1985). Het effect van zelfreflectie op leren probleemoplossen. In J.G.L.C. Lodewijks & P.R.J. Simons (Eds.), *Zelfstandig leren* (pp. 110-118). Amsterdam/Lisse: Swets & Zeitlinger.
- Vermunt, J.D.H.M. (1989). The interplay between internal and external regulation of learning and the design of process-oriented instruction. Paper presented at the third EARLI-conference, Madrid.
- Vermunt, J.D.H.M. (1991). Leerstrategieën van studenten in een zelfinstructie-leeromgeving. *Pedagogische Studiën*, 68, 7, 315-326.
- Wang, M.C., Haertel, G.D., & Walberg, H.J. (1990). What influences learning? A content analysis of review literature. *The Journal of Educational Research*, 84, 1, 30-43.
- Weinstein, C.E. & Mayer, R.F. (1986). The teaching of learning strategies. In M.C. Wittrock (Ed.). *Handbook of research on teaching (3rd edition)* (pp. 315 -328). New York: MacMillan.
- Zimmerman, B.J. & Pons, M.M. (1986). Development of structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23, 4, 614-628.

Manuscript ontvangen 14-10-1992

Definitieve versie ontvangen 3-12-1992

# Training voor transfer van statistische vaardigheden: toepassing van een vier-componenten instructie-ontwerpmodel

Fred G.W.C. Paas en Jeroen J.G. van Merriënboer  
*Vakgroep Instructietechnologie, Universiteit Twente\**

## ABSTRACT

The effect on transfer of statistical knowledge and skills of three instructional strategies was investigated. The instructional strategies were: (a) a traditional strategy which emphasized the solving of goal-specific open problems, (b) a study strategy which emphasized the study of worked-out problem solutions, and (c) a completion strategy which emphasized the completion of partly worked-out problem solutions. A four-component instructional design model was used for the development of the instructional strategies. The main hypothesis of the research was that the study and completion strategy would lead to more inductive processing and schema acquisition and hence to better performance on a transfer test; it was further predicted that this effect would be larger for "far" transfer problems than for "near" transfer problems. These hypotheses were supported by the results.

## INLEIDING

Dit artikel beschrijft een onderzoek waarin het effect van drie instructiestrategieën op de bereikte mate van transfer van statistische vaardigheden onderzocht wordt. Het onderzoek richt zich op het domein van de descriptieve statistiek en heeft betrekking op het leren toepassen van centrale tendentiematen (gemiddelde, modus en mediaan). De drie onderzochte instructiestrategieën zijn: (a) een traditionele *oplosstrategie*, waarbij het zelfstandig oplossen van doelspecifieke, open problemen de belangrijkste activiteit is, (b) een *bestudeerstrategie*, waarbij het bestuderen van problemen met geheel uitgewerkte oplossingen centraal staat, en (c) een *completeerstrategie*, waarbij het completeren van slechts gedeeltelijk uitgewerkte oplossingen benadrukt wordt.

Statistisch analyseren moet beschouwd worden als een complexe cognitieve vaardigheid omdat er sprake is van een prominente probleemoploscomponent. Statistische problemen zijn vaak slecht gedefiniëerd; het oplossen van deze problemen vereist aanzienlijk méér dan het toepassen van rekenkundige algoritmen. Stel bijvoorbeeld dat een grote hoeveelheid ongeorganiseerde gegevens beschreven moet worden in centrale tendentie- en spreidingsmaten. De probleemoplosser moet dan relevante en irrelevante gegevens van elkaar weten te scheiden; verantwoorde beslissingen nemen omtrent de te gebruiken tendentie- en spreidingsmaten; juiste sekwenties van te hanteren vergelijkingen genereren; de gegevens koppelen aan de juiste constanten en variabelen in deze vergelijkingen, enzovoort. Dit betekent dat iedere probleemsituatie weer nieuwe, onverwachte eisen zal stellen aan de probeemoplosser of de lerende. Het is ook precies om deze reden dat het onderwijzen van complexe cognitieve vaardigheden, zoals statistisch analyseren, impliciet *transfer* van de verworven kennis en vaardigheden veronderstelt. Elke nieuwe probleemsituatie kenmerkt zich immers door het feit dat er onbekende aspecten aan te onderscheiden zijn, die slechts tegemoet getreden kunnen worden door reeds verworven kennis en vaardigheden op een "nieuwe" manier te gebruiken.

\* Fred Paas, Universiteit Twente, Faculteit der Toegepaste Onderwijskunde, Vakgroep Instructietechnologie, Postbus 217, 7500 AE Enschede.



Een algemeen probleem binnen de instructietechnologie is het vrijwel ontbreken van instructie-ontwerpmoedellen die richtlijnen verschaffen voor de vormgeving van training in complexe cognitieve vaardigheden. De weinige beschikbare modellen (zie bijv., Fabiani, Buckley, Gratton, Coles, Donchin & Logie, 1989; Gopher, Weil & Siegel, 1989; Myers & Fisk, 1987; Schneider, 1985) zijn dikwijls niet erg gedetailleerd en houden niet expliciet rekening met het belang van transfer van de verworven vaardigheden naar nieuwe probleemsituaties. In een recent artikel presenteren Van Merriënboer, Jelsma en Paas (in druk; zie ook Jelsma, Van Merriënboer & Bijlstra, 1990) een vier-componenten instructie-ontwerpmoedel (4C-moedel) dat meer gedetailleerde richtlijnen bevat voor de ontwikkeling van training voor complexe cognitieve vaardigheden. Het moedel laat tevens predicties toe met betrekking tot de te verwachten mate van transfer van de onderwezen vaardigheden. Onderzoek naar het 4C-moedel heeft tot op heden plaats gevonden binnen de domeinen "fault management" in de procesindustrie (Jelsma, 1989; Jelsma & Bijlstra, 1988, 1990) en inleidend programmeren (Van Merriënboer, 1990a, 1990b; Van Merriënboer & De Croock, in druk).

In het onderhavige onderzoek naar het onderwijzen van vaardigheden in het statistisch analyseren is het 4C-moedel gebruikt om de drie verschillende instructiestrategieën te ontwikkelen. Alvorens over te gaan tot een bespreking van het uitgevoerde onderzoek, zullen wij een summier overzicht geven van het gebruikte 4C-moedel. Vervolgens wordt een beschrijving gegeven van de drie onderzochte instructiestrategieën, hun theoretische achtergronden en de predicties met betrekking tot de verwachte mate van transfer.

#### Het vier-componenten instructie-ontwerpmoedel

Figuur 1 bevat een weergave van de basiselementen van het 4C-moedel (voor een uitgebreide beschrijving van het moedel, zie Van Merriënboer, Jelsma, & Paas, in druk). Het moedel postuleert dat transfer van verworven vaardigheden vereist dat de training expliciet gericht is op (a) de automatisering van "gesloten" deelvaardigheden (closed skills, Fisk & Gallini, 1989), welke consistent zijn over probleemsituaties, alsmede (b) de verwerving van cognitieve schemata die de basis vormen voor de uitvoering van "open" deelvaardigheden, welke variabel zijn over probleemsituaties. De identificatie van-gesloten en open-deelvaardigheden in een proces van *principled skill decomposition* ligt aan de basis van het trainingsontwerp.

Vervolgens vindt een analyse plaats van elk van de onderscheiden deelvaardigheden. Gesloten vaardigheden (Component 1) worden geanalyseerd in termen van "sterke" methoden; de gespecificeerde procedures of beslisregels zijn in hoge mate taak-specifiek en algoritmisch van aard. De gerelateerde kennisanalyse (Component 3) houdt in dat iedere stap in een gespecificeerde procedure, of iedere propositie in een beslisregel, wordt onderworpen aan de vraag wat de lerende moet *weten* om die stap correct te kunnen zetten of om die regel correct te kunnen toepassen. Open vaardigheden (Component 2) worden geanalyseerd in termen van "zwakke" methoden; de gespecificeerde procedures (bijv. in de vorm van een Systematische Probleemaanpak, Mettes, Pilot & Roossink, 1981) of vuistregels zijn niet (volledig) taak-specifiek en heuristisch van aard. De gerelateerde kennisanalyse (Component 4) houdt in dat wordt beschreven welke cognitieve structuren de kans op succes van de gespecificeerde "zwakke" methoden verhogen.

De ontwerpfase volgt pas nadat een analyse heeft plaats gevonden van alle deelvaardigheden. Het moedel veronderstelt dat voor elk van de vier componenten een verzameling instructietactieken bestaat waaruit een selectie gemaakt moet worden. Elke tactiek beschrijft een instructiemethode alsmede de voorwaarden en de doelen waarbij deze methode gebruikt kan worden. Tactieken voor Component 1 betreffen methoden voor de vormgeving van oefening op dusdanige wijze dat kenniscompilatie (Anderson, 1983, 1987) bevorderd wordt; tactieken uit Component 3 betreffen methoden voor de presentatie van informatie zodat deze voor de lerende optimaal beschikbaar is tijdens oefening. De toepassing van combinaties van geschikte tactieken uit de componenten 1 en 3 stimuleert automatisering van de betrokken deelvaardigheden door middel van *regelautomatisering*. Tactieken uit Component 2 betreffen methoden voor de vormgeving van oefening op dusdanige wijze dat kennisinductie, dat wil zeggen, de (re)organisatie of (her)structurering van kennis in cognitieve schemata, bevorderd wordt; tactieken uit Component



## Training voor transfer van statistische vaardigheden

19

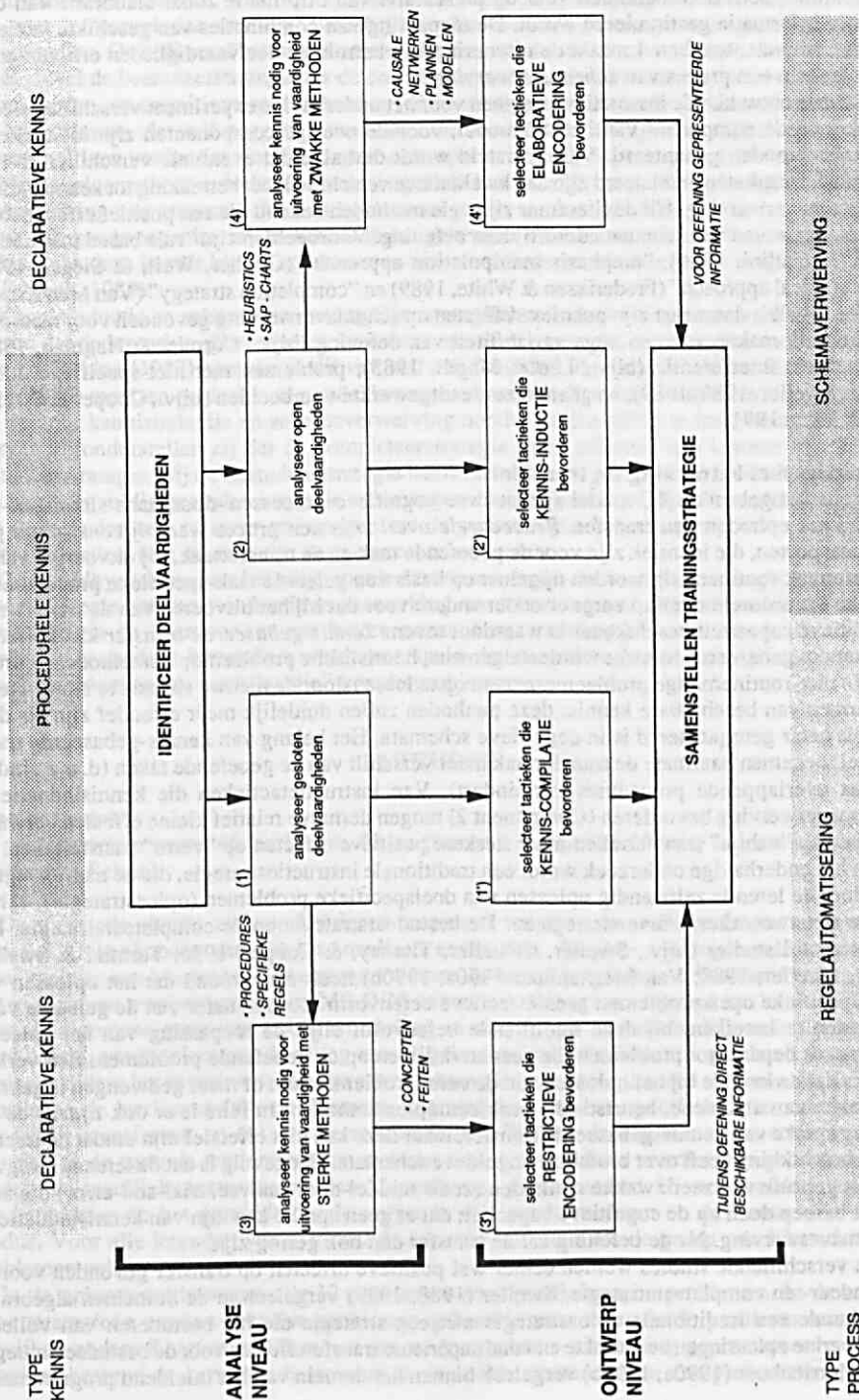


Fig. 1. Het vier-componenten instructie-ontwerpmoedel.

4, tenslotte, betreffen methoden voor de presentatie van informatie zodat elaboratie van deze nieuwe informatie gestimuleerd wordt. De toepassing van combinaties van geschikte tactieken uit de componenten 2 en 4 maakt de uitvoering van betrokken deelvaardigheden efficiënter ten gevolge van een proces van *schemaverwerving*.

De drie ontwikkelde instructiestrategieën voor het onderhavige experiment verschillen slechts op de tweede component van het 4C-model; voor de overige componenten zijn identieke instructiemethoden gehanteerd. Verondersteld wordt derhalve dat eventuele verschillen in leer- en transferresultaten gerelateerd zijn aan kwalitatieve verschillen met betrekking tot kennisinductie en schemaverwerving. Uit de literatuur zijn vele methoden bekend die een positief effect hebben op het verwerven van schemata door/tijdens oefening. Voorbeelden zijn "rule based instruction" (Fisk & Gallini, 1989), "emphasis manipulation approach" (Gopher, Weil, & Siegel, 1989), "hierarchical approach" (Frederiksen & White, 1989) en "completion strategy" (Van Merriënboer, 1990a, 1990b); daarnaast zijn positieve effecten op schemaverwerving gevonden voor methoden die gebruik maken van een hoge variabiliteit van oefening (bijv., Cormier & Hagman, 1987), contextuele interferentie (bijv., Lee & Magill, 1985), problemen met niet-specifieke doelen (bijv., Sweller, 1988, 1989), en geannoteerde uitgewerkte voorbeelden (bijv., Cooper & Sweller, 1987; Paas, 1991).

### Predicties met betrekking tot transfer

Volgens het gebruikte 4C-model spelen twee cognitieve processen—doorgaans simultaan—een rol bij het optreden van transfer. *Procedurele overlap* is een proces waarbij consistente probleemaspecten, die identiek zijn voor de geoefende taak en de transfertaak, bij uitvoering van de transfertaak routinematig worden opgelost op basis van geleerde taak-specifieke procedures of regels. Procedurele overlap zorgt er onder andere voor dat bij het uitvoeren van de transfertaak cognitieve capaciteit beschikbaar is waardoor tevens *kennis-gebaseerde transfer* kan optreden. Bij kennis-gebaseerde transfer worden algemene, heuristische probleemoplossmethoden gebruikt om de niet-routinematige probleemaspecten op te lossen door de nieuwe situatie te interpreteren in termen van beschikbare kennis; deze methoden zullen duidelijk méér effectief zijn als deze kennis beter georganiseerd is in cognitieve schemata. Het belang van kennis-gebaseerde transfer zal toenemen naarmate de transfertaak méér verschilt van de geoefende taken (d.w.z., indien de set overlappende procedures vermindert). Van instructietactieken die kennisinductie en schemaverwerving bevorderen (Component 2) mogen derhalve relatief kleine effecten verwacht worden op "nabije" transfertaken maar sterkere positieve effecten op "verre" transfertaken.

In het onderhavige onderzoek wordt een traditionele instructiestrategie, die de nadruk legt op het door de lerende zelfstandig oplossen van doelspecifieke problemen (oplosstrategie), vergeleken met twee alternatieve strategieën: De bestudeerstrategie en de completeerstrategie. Een groot aantal studies (bijv., Sweller, Chandler, Tierney, & Cooper, 1990; Tarmizi & Sweller, 1988; Sweller, 1988; Van Merriënboer 1990a, 1990b) heeft aangetoond dat het oplossen van doelspecifieke open problemen geen effectieve oefenvorm is om transfer van de geleerde vaardigheden te bereiken; bij deze traditionele oefenvorm blijft de toepassing van het geleerde doorgaans beperkt tot problemen die zeer sterk lijken op de geoefende problemen. Een verklaring is dat de lerende bij het oplossen van de oefenproblemen min of meer gedwongen is gebruik te maken van algemene, heuristische probleemoplossmethoden. In feite is er ook *tijdens* de oefening sprake van kennis-gebaseerde transfer, maar deze kan niet effectief zijn omdat de lerende geen beschikking heeft over bruikbare cognitieve schemata. Het gevolg is dat de lerende terugvalt op het gebruik van zeer zwakke methoden (zoals middel-doel analyse, trial-and-error) die zo'n groot beroep doen op de cognitieve capaciteit dat er geen sprake kan zijn van kennisinductie en schemaverwerving. Na de oefening zal de transfer dan ook gering zijn.

In verschillende studies werden echter wel positieve effecten op transfer gevonden voor de bestudeer- en completeerstrategie. Sweller (1988, 1989) vergeleek in de domeinen algebra en meetkunde een traditionele oplosstrategie met een strategie die het bestuderen van volledig uitgewerkte oplossingen benadrukte en vond superieure transfereffecten voor de bestudeerstrategie. Van Merriënboer (1990a, 1990b) vergeleek binnen het domein van het inleidend programmeren

een traditionele strategie, waarbij lerenden zelfstandig computerprogramma's moesten schrijven, met de "completion strategy", waarbij lerenden onvolledige computerprogramma's moesten aanvullen. Ook hij vond betere transferresultaten voor de completeerstrategie. Een verklaring is dat zowel de bestudeerstrategie als de completeerstrategie de lerenden expliciet de mogelijkheden bieden om cognitieve schemata af te leiden uit de gepresenteerde probleembenaderingen en -oplossingen; dat wil zeggen, kennisinductie bevorderen. De (gedeeltelijke) oplossingen bewerkstelligen verder dat niet alle cognitieve capaciteit wordt opgesoupeerd door zeer zwakke probleemoplosmethoden, zodat de aandacht gericht kan worden op de relevante taakaspecten en de verwerving van schemata die richting kunnen geven aan het oplossen van de problemen. Zo wordt een goede organisatie van kennis in het geheugen bereikt, waardoor tevens de toegankelijkheid van deze kennis bij het oplossen van nieuwe transferproblemen wordt vergroot (Bassok & Holyoak, 1989).

Kennisinductie en schemaverwerving zijn strategisch te beïnvloeden processen, hetgeen wil zeggen dat de lerenden *effort* moeten investeren om deze processen te doen plaats vinden. Salomon en Perkins (1987) spreken in dit verband van "mindful abstraction". Van Merriënboer en Paas (1989) merkten op dat het niet reëel is om te veronderstellen dat een lerende altijd bereid is om de voor kennisinductie en schemaverwerving noodzakelijke *effort* te investeren. Op grond hiervan veronderstellen zij dat de completeerstrategie meer effectief zou kunnen zijn dan de bestudeerstrategie. Bij de bestudeerstrategie heeft men immers geen enkele controle over het al dan niet zorgvuldig bestuderen van de oplossingen door de lerenden; bij de completeerstrategie zijn de lerenden min of meer gedwongen om de gedeeltelijke oplossingen goed te bestuderen omdat zij anders niet in staat zijn deze correct te voltooien.

De hoofdhypothese van het uitgevoerde onderzoek is dat de bestudeerstrategie en de completeerstrategie tot meer inductieve verwerking en schemaverwerving zullen leiden dan de oplosstrategie en daardoor tot een betere prestatie op transferproblemen die na de training worden aangeboden; bovendien wordt voorspeld dat dit effect groter zal zijn voor "verre" transferproblemen dan voor "nabije" transferproblemen.

## METHODE

### Proefpersonen

De proefpersonen waren 46 leerlingen, 45 jongens en 1 meisje (16-18 jaar) van een Middelbaar Technische School.

### Materialen.

De instructie richtte zich op het leren toepassen van drie centrale tendentiematen uit de statistiek (rekenkundig gemiddelde, mediaan en modus). De drie trainingscondities werden ontwikkeld op een Macintosh IIcx computer binnen de auteursomgeving Authorware Professional. Het experiment werd uitgevoerd op MS-DOS computers. Direct voorafgaand aan het experiment werd de algemene gang van zaken mondeling en schriftelijk toegelicht. Bijzondere aandacht werd daarbij geschonken aan de uitleg van de gebruikte mentale inspanningsschaal, waarop de leerlingen tijdens het experiment de geïnvesteerde mentale inspanning dienden te scoren. Vervolgens doorliepen de leerlingen (a) een algemene instructiefase, (b) een specifieke instructiefase, (c) een nabije transfertest en (d) een verre transfertest. Tijdens de algemene instructie werd de theorie van de centrale tendentiematen met een aantal illustratieve voorbeelden gepresenteerd. Tijdens de specifieke instructie werden 12 problemen gepresenteerd. Hiervan hadden de eerste zes betrekking op het gemiddelde, de volgende drie op de mediaan en de laatste drie op de modus. Voor alle lerenden waren zowel de formele instructie als de groep van 12 statistische problemen gelijk.

In de oplosstrategie waren alle 12 problemen doelspecifieke open problemen. In de andere twee strategieën werden dezelfde problemen gebruikt, maar ieder derde probleem was open, terwijl de andere 8 samen met hun - al dan niet complete - uitwerking gepresenteerd werden. Hierdoor bestond de kern van de training in de oplos-, bestudeer- en completeerstrategie res-

pectievelijk uit het oplossen van open problemen, het bestuderen van uitgewerkte probleemoplossingen, en het completeren van gedeeltelijk uitgewerkte probleemoplossingen.

De nabije transfertest alsook de verre transfertest bestond voor alle condities uit dezelfde 12 problemen. A-priori werden 12 problemen gekarakteriseerd als "nabije" transferproblemen en 12 problemen als "verre" transferproblemen. De classificatie, die afhankelijk was van de mate van overeenkomst van de probleempresentatie met de presentatie van de trainingsproblemen, werd uitgevoerd door de onderzoeker in samenwerking met twee statistiek-leraren. Met uitzondering van de gegeven waarden waren de nabije transferproblemen gelijk aan de open trainingsproblemen. Bij de verre transferproblemen was de probleempresentatie verschillend van de trainingsproblemen. De bruikbare data in de nabije transferproblemen werden gepresenteerd op een gestructureerde manier (bijv., een tabel) in een vorm die bekend was van de trainingsproblemen en waarbij bovendien geen twijfel bestond over de toekenning van de data aan bepaalde variabelen. De bruikbare data in de verre transferproblemen dienden door de lerenden zelf uit een voor hen onbekende presentatievorm (bijv., een histogram) geselecteerd te worden en aan de geschikte variabelen te worden toegekend. Figuur 2 toont een voorbeeld van een nabij transferprobleem (A) en een verre transferprobleem (B).

### Onderzoekopzet en procedure

De proefpersonen werden op basis van toeval aan één van de drie condities toegewezen. Binnen een programmagestuurde minimale en maximale tijd konden de leerlingen in hun eigen tempo

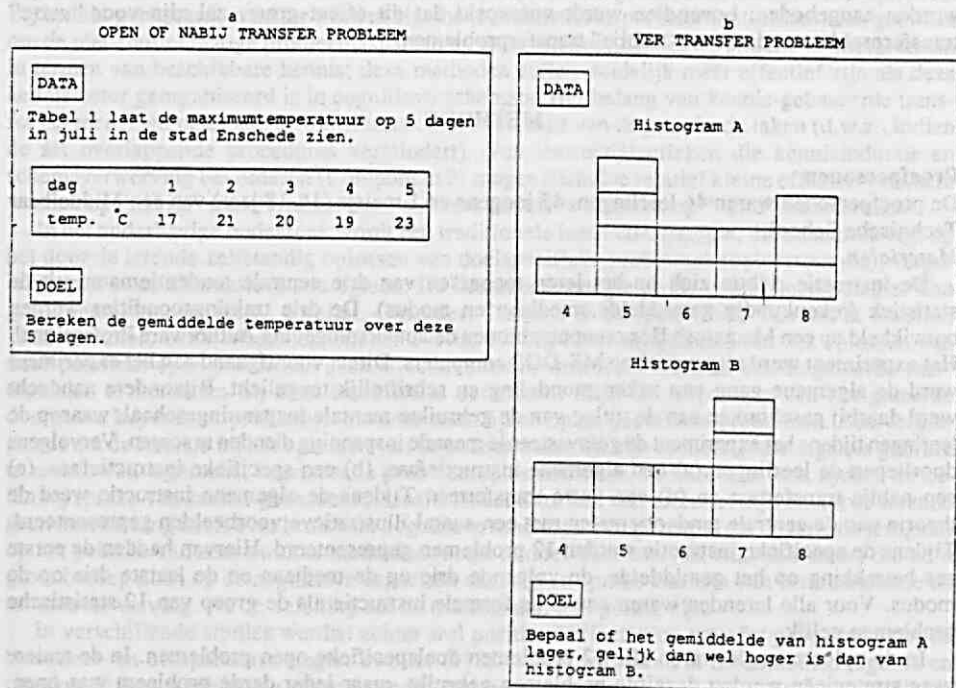


Fig. 2. Voorbeeld van een in dit onderzoek gebruikt (A) nabij transferprobleem en (B) verre transferprobleem.



door de verschillende fasen van het experiment gaan. Wanneer een leerling er na zeven minuten niet in geslaagd was een probleem op te lossen of te completeren, of wanneer drie foute oplossingen gegeneerd waren, werd in de oplos- en completeerstrategie automatisch de correcte oplossing gepresenteerd. Dit betekent dat de lerenden in alle drie de condities altijd geconfronteerd werden met correcte oplossingen voor *alle* problemen. Leerlingen in de bestudeerconditie kregen vijf minuten de tijd om een probleemoplossing te bestuderen. Hen werd duidelijk gemaakt dat de open problemen tijdens de training gelijksoortig waren aan de te bestuderen uitgewerkte problemen. De transfertest diende direct na de specifieke instructiefase gemaakt te worden. Eerst werden de nabije transferproblemen aangeboden en vervolgens de verre transferproblemen. Tijdens de transfertest, waarvoor de leerlingen 40 minuten de tijd kregen, werd géén feedback gegeven. De probleemoplossingen van de leerlingen werden vergeleken met die van de statistiekleraars. De prestatie op de nabije en verre transfertest werd vervolgens door de onderzoeker uit de goed/fout scores samengesteld.

Na ieder probleem in de specifieke instructie en na ieder probleem in de transfertest, werd de leerlingen gevraagd de mentale inspanning die zij investeerden in het oplossen, bestuderen, of completeren van het probleem te scoren op een 9-puntsschaal. De waarden 1 en 9 kwamen respectievelijk overeen met zeer, zeer lichte mentale inspanning en zeer, zeer zware mentale inspanning.

## RESULTATEN

De data werden geanalyseerd met behulp van éénweg Anova's. Vier proefpersonen werden uitgesloten van de analyses omdat hun computers tijdens het experiment niet naar behoren functioneerden. De oplos-, bestudeer- en completeerstrategie bevatten daardoor respectievelijk 13, 14 en 15 proefpersonen. Bij significante  $F$ -waarden werden post-hoc multiple vergelijkingen uitgevoerd volgens de Scheffé methode.

Figuur 3 toont de resultaten waaruit blijkt dat bij de nabije transfertoets ( $F(2,39) = 6.17$ ,  $MS_e = 1.85$ ,  $p < .01$ ) de resultaten van de bestudeerstrategie hoger zijn dan die van de oplos- en completeerstrategie. De resultaten van de oplos- en completeerstrategie verschillen niet van elkaar. Verder blijkt uit Figuur 3 dat de score op de verre transfertoets ( $F(2,39) = 8.32$ ,  $MS_e = 8.87$ ,  $p < .001$ ) hoger was voor de bestudeer- en completeerstrategie dan voor de oplosstrategie. De transferresultaten van de bestudeerstrategie en de completeerstrategie verschillen niet significant van elkaar.

Andere belangrijke resultaten zijn dat de leerlingen in de bestudeerstrategie minder tijd nodig hebben voor de training (specifieke instructie) dan de leerlingen in de oplos- en completeerstrategie ( $F(2,39) = 6.86$ ,  $MS_e = 55.54$ ,  $p < .01$ ), dat de mentale inspanning tijdens de specifieke instructie niet verschilt over condities ( $F(2,39) = 0.74$ ), en dat de mentale inspanning tijdens de verre en nabije transfertest lager is in de bestudeer- en completeerstrategie dan in de oplosstrategie (nabije transfer,  $F(2,39) = 9.14$ ,  $MS_e = 1.01$ ,  $p < .001$ ; verre transfer,  $F(2,36) = 12.13$ ,  $MS_e = 1.09$ ,  $p < .001$ ).

## DISCUSSIE

De resultaten ondersteunen de hypothese dat de bestudeer- en completeerstrategie tot hogere transfer leiden dan de oplosstrategie. De predictie dat het positieve effect van de bestudeer- en completeerstrategie groter is voor verre transferproblemen dan voor nabije transferproblemen wordt eveneens door de resultaten ondersteund. Op de nabije transfertest halen de leerlingen uit de bestudeerstrategie de hoogste score; bovendien wordt dit resultaat bereikt na een kortere (trainings)tijd en met minder mentale inspanning dan bij de oplosstrategie. De score voor de completeerstrategie op de nabije transfertest is weliswaar niet hoger dan de score voor de oplosstrategie, maar deze wordt wel met minder mentale inspanning bereikt. Op de verre



## Statistisch analyseren oplos-/completeer-/bestudeerstrategie

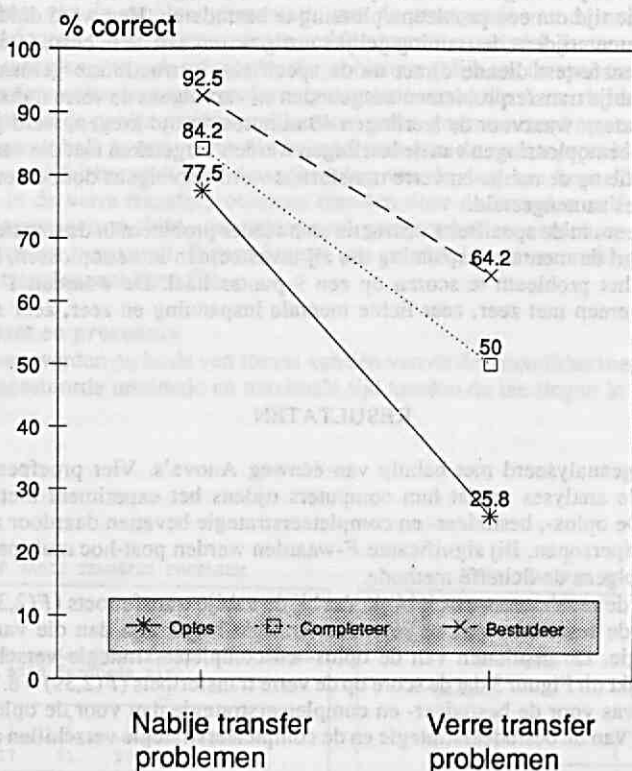


Fig. 3. Percentages correcte antwoorden op de nabije en de verre transfertest, na training van statistische analysevaardigheden met de oplosstrategie, de completeerstrategie en de bestudeerstrategie.

transfertest wordt voor zowel de bestudeer- als de completeerstrategie een hogere score bereikt dan voor de oplosstrategie; bovendien worden ook deze scores bereikt met minder mentale inspanning.

De scores op de transfertests in combinatie met de gerapporteerde mentale inspanningsscores zijn volledig in overeenstemming met de geopperde verklaring dat oefening met deels of geheel uitgewerkte problemen meer kennisinductie en schemaverwerking tot gevolg heeft dan oefening met doelspecifieke open problemen. Ten eerste verklaart de verwerving van schemata het feit dat het positieve effect van de bestudeer- en completeerstrategie groter is voor verre transferproblemen dan voor nabije transferproblemen: Het belang van cognitieve schemata, die een verschillend gebruik van dezelfde kennis mogelijk maken, wordt immers verondersteld in belang toe te nemen naarmate de transferproblemen méér afwijken van de geoefende problemen.

Daarnaast verklaart superieure schemaverwerving bij de bestudeer- en completeerstrategie de lagere mentale inspanning bij het oplossen van de transferproblemen. Doordat bruikbare schemata verworven zijn kan gebruik gemaakt worden van meer efficiënte probleemoplossmethoden die een kleiner beroep doen op de te investeren mentale inspanning. Kortom, de cognitieve structuur die ontstaat ten gevolge van oefening met (gedeeltelijk) uitgewerkte probleemoplossingen vormt een meer efficiënte kennisbasis voor het oplossen van transferproblemen dan de cognitieve structuur die ontstaat ten gevolge van oefening met doelspecifieke open problemen.

Voor de verwachting dat de completeerstrategie betere transferresultaten kan opleveren dan de bestudeerstrategie wordt in dit onderzoek geen steun gevonden. Hierbij werd verondersteld dat lerenden niet altijd geneigd zijn om effort te investeren in het bestuderen van oplossingen; bij de completeerstrategie zou dit gevaar minder groot zijn omdat de lerenden hier min of meer gedwongen worden om de gedeeltelijke oplossingen te bestuderen omdat zij deze anders niet correct kunnen aanvullen. Uit de gerapporteerde mentale inspanningsscores blijkt dat in de drie instructiestrategieën tijdens de oefening eenzelfde mentale inspanning werd gerapporteerd. Mogelijk is dit een gevolg van de korte duur en de experimentele setting van het onderzoek, waardoor alle lerenden optimaal gemotiveerd waren om de opgedragen taken uit te voeren. Dit zou betekenen dat een positief effect van de completeer- boven de bestudeerstrategie pas manifest wordt bij langer durend onderzoek of bij onderzoek in een meer schoolse setting.

Mogelijke alternatieve verklaringen voor de gevonden resultaten zijn de volgende. Een eerste alternatieve verklaring heeft te maken met het genereren van incorrecte oplossingen. In tegenstelling tot leerlingen die problemen moesten oplossen of completeren was het voor leerlingen die oplossingen moesten bestuderen namelijk onmogelijk om incorrecte oplossingen te genereren. T-testen toonden aan dat het gemiddeld aantal incorrect gegenereerde oplossingen in de oplosstrategie ( $M = 0.98$ ) significant hoger was dan in de bestudeerstrategie,  $t(12) = 6.23, p < 0.001$ , en hoger was dan in de completeerstrategie ( $M = 0.07$ ),  $t(14) = 2.57, p < 0.025$ . Ondanks het feit dat de leerlingen in de oplos- en completeerstrategie ook altijd de correcte probleemoplossing konden bestuderen, is het mogelijk dat het genereren van incorrecte oplossingen in de oplosstrategie en in mindere mate in de completeerstrategie een negatief effect heeft op de organisatie van de verworven kennis in cognitieve schemata en daardoor op de bereikte transfer. Aangezien de ontwikkeling en kwaliteit van de schemata niet direct werden bepaald, zou de in dit artikel gehanteerde schema-gebaseerde verklaring in toekomstig onderzoek vergeleken kunnen worden met deze alternatieve schema-gebaseerde verklaring door bijvoorbeeld verbale protocollen tijdens en na de training af te nemen.

Een tweede alternatieve verklaring voor de transferresultaten heeft te maken met de volgorde waarin de lerenden de zogenaamde studie-test cycli doorlopen. "Studie" heeft daarbij betrekking op het bestuderen of completeren van geheel of deels uitgewerkte probleemoplossingen en "test" op het oplossen van problemen. Na het bestuderen of completeren van twee probleemoplossingen moesten de leerlingen in de bestudeer- en completeerconditie óók steeds een open probleem oplossen. Het is mogelijk dat deze directe mogelijkheid om de verworven cognitieve schemata toe te passen tot de superieure transferresultaten heeft geleid. In toekomstig onderzoek zou onderzocht kunnen worden of het gebruik van "zuivere" condities, waarin leerlingen uitsluitend oplossingen bestuderen of uitsluitend deels uitgewerkte oplossingen completeren, vergelijkbare resultaten oplevert.

Tot slot kan opgemerkt worden dat onderzoek in verschillende domeinen inmiddels heeft aangetoond dat de bestudeerstrategie en de completeerstrategie als veelbelovende instructiestrategieën voor het bereiken van transfer beschouwd kunnen worden. In vergelijking met traditionele strategieën leiden zij tot betere transferresultaten met lagere mentale inspanning.

#### NOOT

1. De auteurs danken Henny Kramers-Pals voor haar bijdrage aan dit artikel.

## LITERATUUR

- Anderson, J.R. (1983). *The architecture of cognition*. Cambridge: Harvard University Press.
- Anderson, J.R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review*, 94, 192-210.
- Bassok, M., & Holyoak, K.J. (1989). Interdomain transfer between isomorphic topics on algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 153-166.
- Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, 79, 347-362.
- Cormier, S.M., & Hagman, J.D. (Eds.). (1987). *Transfer of learning: Contemporary research and applications*. San Diego, CA: Academic Press.
- Fabiani, M., Buckley, J., Gratton, G., Coles, M.G.H., & Donchin, E. (1989). The training of complex task performance. *Acta Psychologica*, 71, 259-299.
- Fisk, A.D., & Gallini, J.K. (1989). Training consistent components of tasks: Developing an instructional system based on automatic/controlled processing principles. *Human Factors*, 31, 453-463.
- Frederiksen, J.R., & White, B.Y. (1989). An approach to training based upon principled task composition. *Acta Psychologica*, 71, 89-146.
- Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica*, 71, 147-177.
- Jelsma, O. (1989). *Instructional Control of Transfer*. Enschede, The Netherlands: Bijlstra & Van Merriënboer.
- Jelsma, O., & Bijlstra, J.P. (1988). Training for transfer in learning to detect, diagnose, and compensate system failures. *Proceedings of the Seventh European Annual Conference on Human Decision Making and Manual Control* (pp. 256-262). France, Paris.
- Jelsma, O., & Bijlstra, J.P. (1990). PROCESS: Program for Research on Operator Control in an Experimental Simulated Setting. *IEEE Transactions on Systems, Man, and Cybernetics*, 20, 1221-1228.
- Jelsma, O., van Merriënboer, J.J.G., & Bijlstra, J.P. (1990). The ADAPT design model: Towards instructional control of transfer. *Instructional Science*, 19, 89-120.
- Lee, T.D., & Magill, R.A. (1985). Can forgetting facilitate skill acquisition? In D. Goodman, R.B. Wilberg, & I.M. Franks (Eds.), *Differing perspectives in motor learning, memory and control* (pp. 3-22). Amsterdam: Elsevier Science Publishers.
- Mettes, C.T.W., Pilot, A., & Roosink, H.J. (1981). Linking factual knowledge in solving science problems: A case study in a thermodynamics course. *Instructional Science*, 10, 333-361.
- Myers, G.L., & Fisk, A.D. (1987). Training consistent task components: Application of automatic and controlled processing theory to industrial task training. *Human Factors*, 29, 255-268.
- Paas, F.G.W.C. (1991). *Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach*. IST-MEMO-91-02.
- Salomon, G., & Perkins, D.N. (1987). Transfer of cognitive skills from programming: When and how? *Journal of Educational Computing Research*, 3, 149-169.
- Schneider, W. (1985). Training high-performance skills: Fallacies and guidelines. *Human Factors*, 27, 285-300.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257-285.
- Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*, 4, 457-466.
- Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, 119, 176-192.
- Tarmizi, R.A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, 80, 424-436.
- Van Merriënboer, J.J.G. (1990a). Strategies for programming instruction in high school: Program completion vs. program generation. *Journal of Educational Computing Research*, 6, 265-287.
- Van Merriënboer, J.J.G. (1990b). *Teaching introductory computer programming: A perspective from instructional technology*. Enschede, The Netherlands: Bijlstra & Van Merriënboer.
- Van Merriënboer, J.J.G., & De Croock, M.B.M. (in druk). Strategies for computer-based programming instruction: Program completion vs. program generation. *Journal of Educational Computing Research*.
- Van Merriënboer, J.J.G., Jelsma, O., & Paas, F.G.W.C. (in druk). Training for reflective expertise: A four-component instructional design model for complex cognitive skills. *Educational Technology, Research & Development*.

Van Merriënboer, J.J.G., & Paas, F.G.W.C. (1989). Automation and schema acquisition in learning elementary computer programming: Implications for the design of practice. *Computers in Human Behavior*, 6, 273-289.

Manuscript ontvangen 5-9-1991

Definitieve versie ontvangen 9-12-1991

## ABSTRACT

In this paper, three different approaches to teaching elementary computer programming in a first semester university course are compared by means of the transfer paradigm. It is assumed that teaching directly influences learning success. In order to test this assumption, the transfer paradigm is used as a means to check whether the teaching approaches are transferable. The transfer paradigm is about studying variables in the classroom, teaching the material and checking whether the learning behavior, but not the variables themselves, between teaching and learning, also appears in learning time and post-test results. Finally, in the third section, practice implications are proposed in terms of their possible contribution to teaching effectiveness. These developments involve a shift in the operationalization of teaching rather than powerful effectiveness criteria.

## INTRODUCTION

Since almost a century scientists in the field of education have been searching theoretical frameworks, models, methods and criteria to define teaching effectiveness. However, not universal agreement has been reached. These authors would like to refer to the work of Snyggan, a King of Canaan, who tried to tell a story to a lot of people only to have it fall down again on every side by falling, not even as a whole, period through several phases, dependent on duration of study and educational level (see, for instance, 1989; Coomans, 1991; De Corte & Loyens, 1993; Dossela & Biddle, 1976; Gage, 1988; Gage & Gage, 1999; Lippitt & White, 1959; Salomon, 1981; Stebbins, 1979). Though numerous criteria have been proposed, it is very hard to find the precise-product approach (see, for instance, 1989, p. 1008), since researchers are still in a state of confusion.

In the days of new psychology and personality studies, effectiveness was being re-perceived as a by-product of the teacher's personality. Not teaching behavior, but rather the teacher's personality is the paramount indicator for effectiveness. The research design is a correlational one, which is strong correlation between teachers' characteristics and perceived teaching quality is hypothesized. The outcomes of ability tests and questionnaires are supposed to be a sign of both teachers' and students' opinions on the teacher's quality. Distorted results, however, indicate four variable results; therefore, more results to account for this situation are needed. In fact, 1983: First, the results tell little about the teaching profession since, instead of being pupil-oriented, teacher- or enthusiastic or psychologically stable, open to all and whose interaction with people is predominantly loving, subjective and idiosyncratic, people who are often people, mostly reflect traditional and superficial influences of quality. These being personality traits do not fit in with elementary, discrete teaching behaviors. Each teacher's personality is highly different. Fourth, research reports and meta-analysis comparing groups variable learning and most that greatest characteristics are derived from individual studies.

In this article, more attention is paid to the other side of both design and criteria. A clear example is the need for a diagnostic teacher to provide a solution to the teacher's formal content using the indirect and indirect to the student's response. Generally, articles or letters (leadership style) and research literature studies, which have shown the persistent subject of study (see, for instance, 1989). The operation of this process is not only operationalized in

Tijdschrift voor Onderwijsresearch, 19 (1994), nr. 1, pp. 17-25

## Teaching Effectiveness: An overview of studies

Joost Lowyck

*Centre for Instructional Psychology and Technology, Leuven University*

### ABSTRACT

In this contribution, different approaches to teaching effectiveness studies are described. In a first section, effectiveness interpreted by 'process-product' studies is reviewed. It is assumed that teaching directly influences learning output. Isolated effective teaching behaviors are described as well as models in which effective teaching behaviors are grouped. The second section is about mediating variables. In this conception, learning outcomes are not directly influenced by teaching behaviors, but by variables mediating between teaching and learning, like 'academic learning time' and pupil cognitions. Finally, in the third section, teacher cognitions are reviewed in terms of their possible contribution to teaching effectiveness. These descriptive studies reveal the idiosyncratic nature of teaching rather than powerful effectiveness criteria.

### INTRODUCTION

Since almost a century scholars in the field of education use many and divergent theoretical frameworks, models, methods and criteria to define teaching effectiveness without, however, any univocal or definite answer. Their stubborn endeavours bear a likeness to the Greek myth of Sisyphus, a king of Corinth condemned forever to roll a stone up a hill in Hades only to have it roll down again on nearing the top. Indeed, research on teaching passed through several phases, dependent on dominant ideas about education and human behavior (see Anderson, 1989; Creemers, 1991; De Corte & Lowyck, 1983; Dunkin & Biddle, 1974; Gage, 1978; OECD, 1990; Lowyck & Pieters, 1993; Scheerens, 1991; Shulman, 1986). Though teaching effectiveness research came at its very height within the process-product approach (see Wittrock, 1986, p. 1008), some predecessors are still worthy of mention.

In the days of test psychology and personality studies effective teaching is perceived as a by-product of the teacher's personality. Not teaching behavior in itself but the teacher's personality is the paramount indicator for effectiveness. The research design is a correlational one, since a strong correlation between teachers' characteristics and estimated teaching quality is hypothesized. The outcomes of ability tests and questionnaires are compared with expert's and students' opinions on the teacher's quality. Literature reviews, however, indicate few valuable results; there are some reasons to account for this shortage (see Getzels & Jackson, 1963). First, the results tell little about the 'teaching' profession since qualities like being pupil-centered, friendly, enthusiastic or psychologically stable apply to all jobs where interaction with people is predominant. Second, subjective and idiosyncratic perceptions by school people, mostly reflect traditional and superficial indicators of quality. Third, broad personality traits do not fit in with elementary, discrete teaching behaviors, since the unit of analysis highly differs. Fourth, research reviews and meta-analyses emphasize grouped results leaving out more fine-grained characteristics gathered from individual studies.

In later studies more attention is paid to the refinement of both design and criteria. A clear example is the search for a democratic teachers' style as a reaction to the European fascist scene during the thirties and forties. In this context teachers' democratic attitudes or beliefs ('leadership style') and concomitant democratic classroom behaviors became the paramount object of study (see Lippitt & White, 1943). The criterion of effectiveness is not only operationalized in

\*Address: Vesaliusstraat 2, B-3000, Leuven.



terms of opinions of school people, but in measurements of learning results as well. Here, the criterion-of-effectiveness is no more the mere perception of quality, but the learning output itself. Anderson (1959) analyzed two types of effect in 32 studies on democratic leadership styles: productivity in terms of learning gains and positive classroom climate. For learning outcomes 11 studies report gains with a democratic (integrative) style; 8 studies indicate effects with the authoritarian ('dominative') style and 13 reports are not conclusive at all. Regarding the classroom climate the results are more in favour of the democratic style though the criteria used were very divergent. Following Anderson (1959) the main conclusion is clear: "We were not fortunate enough to find that one style is consistently different from the other; thus we are forced to explore new avenues" (p. 212).

This lack of consistency in research outcomes has been interpreted in different ways (see Anderson, 1959; Dunkin & Biddle, 1974; Flanders, 1965). First, the methodology of many studies was weak or inadequate. Second, the concepts 'authoritarian' and 'democratic' held value judgments which in a subtle way influenced the design and interpretation of the studies. Third, the concepts were differently operationalized due to the lack of a precise description of classroom behaviors which account for one or another style. Fourth, a binary dimension (authoritarian-democratic) seems inadequate for a valid interpretation of classroom complexity, since the classroom is multidimensional.

#### PROCESS-PRODUCT STUDIES: TEACHER'S CLASSROOM BEHAVIORS

The lack of clear research outcomes together with the rise of behaviorism as the dominant psychological paradigm forced researchers towards the study of effective classroom 'behaviors'.

The basic question is which specific teaching behaviors (process variables) are effective so as to bring about learning gains in pupils (product variables). In order to make teaching observable, the complex teaching activity is fragmented into elementary or discrete teaching behaviors in line with the analytical paradigm as advocated by Gage (1963). Within this view, scientific problems can be solved by being analyzed into small problems, whose variables are less complex. So-called 'micro-criteria' are the central issue leaving out the search for over-all effectiveness of teachers in the many, varied facets of their roles.

The research method is a correlational one. A lot of time and energy is invested in category systems for the registration of process variables. Indeed, almost each researcher worked out an idiosyncratic observation system or modified already existing ones, which resulted both in the proliferation of observation systems (see Simon & Boyer, 1970) and in a shortage of conceptually consistent models and frameworks (Lowyck, 1980). Moreover, the product variables are weakly operationalized which brought Heath & Nielson (1974) to qualify the operational definitions of student achievement as 'shallow'.

In their meta-analysis of fifty well selected studies Rosenshine & Furst (1973) report nine 'powerful' process variables, like clarity, variability, enthusiasm, task orientation, criticism, indirectness, student opportunity to learn, structuring comments, and use of different levels of questions and cognitive activities. Notwithstanding possible positive interpretations, the problems with this approach are recognized by many scholars in the field of teaching research. First, the analytical and formal nature of the process variables dismisses two important types of variables, namely subject-matter content and pupil characteristics (Heath & Nielson, 1974). It seems obvious that one single set of teaching behaviors cannot account for learning outcomes in all subject-matter and for pupils from different socio-economic background or from all grades, the research being 'content-proof' and 'pupil-proof'. Second, almost all observational systems register behaviors in a traditional 'classroom teaching' context with emphasis on lecturing and question-answer cycles. Individual work or groupwork are not studied and, hence, the research design is 'method-proof'. Third, teaching is conceived of as a unidirectional activity of teachers

neglecting any influence of the situation on the teacher, as recognized by Doyle's mediating paradigm (Doyle, 1978). Since, this approach is 'situation-proof'. Fourth, because emphasis is exclusively on observable classroom behaviors, no information is provided on cognitive processes of teachers and pupils nor on the teachers' complex information processing during the planning, interactive and post-interactive phase of teaching (Jackson, 1968). Consequently, the approach is 'cognition-proof'.

As a solution to the analytical deficiency, researchers try to 'chunk' isolated teaching behaviors into more encompassing models or patterns, like 'active teaching' (Brophy, 1986) or 'adaptive teaching' (Good, 1979; Wang, 1980). In these models, several single teaching behaviors are grouped and a clear interdependency of discrete behavioral elements within one teaching model is hypothesized. Active teaching, for example, claims that students learn best in classes where they spend most of the time being taught or supervised by their teacher rather than working on their own or not working at all. The teacher presents information, develops concepts through lecture and demonstration, offers feedback, prepares the student for follow-up assignments, and monitors progress. There is a great deal of academic teacher talk rather than procedural or managerial, and much of it involves asking questions and giving feedback rather than extended lecturing (Brophy, 1986). This active teaching model seems very promising, mainly for groups of students with a need for intensive instructional support, like in the first grades of elementary education or in special education settings, for well-structured subject-matter (mathematics, reading and grammar) and in teacher-centered approaches (see Brophy, 1991; Carnine, 1989).

A possible disadvantage of this kind of model-driven teaching research is the fact that effective teaching models are defined by means of research reviews, without strong empirical validation of the composite behaviors, or as Gage (1978) states "From hundreds of correlations, especially the significant ones, the investigators and reviewers then synthesize the style or pattern of teaching that seems to be associated with desirable kinds of pupils achievement and attitudes". Further experimental studies with these composite behaviors indicated that not all effective teachers used the 'desirable' behaviors and that not all skills trained showed stable correlations with learning gains. Consequently, not all composites seemed necessary for effective teaching (Shulman, 1986).

Despite efforts to remediate the deficiencies mentioned above, the direct correlation of teacher behavior and pupil achievement remains problematic. It is hypothesized, then, that the search for direct correlations should at least be complemented by studying the 'intervening' or 'mediating' variables which determine classroom effectiveness (see Doyle, 1986).

#### MEDIATING VARIABLES: LEARNING TIME AND TEACHER/PUPIL COGNITIONS

Since the direct link between instructional input (teacher behavior) and output (learning outcomes) caused problems for defining effectiveness, emphasis is laid on the throughput, the variables mediating between teaching activities and learning outcomes. The learner himself becomes the decisive factor in the definition of effectiveness "What a teacher does at any one moment while working in a circumscribed content area affects a student primarily at only that particular moment and in that particular content area. The link between teacher behavior and student achievement is, therefore, the ongoing student behavior in the classroom learning situation" (Berliner, 1979, p. 123).

Both studies on 'Aptitude-Treatment-Interaction' (Snow, 1986) and the 'Beginning Teacher Evaluation Study' (Fisher, Berliner, et al., 1978) revealed the importance of individual differences and their concomitant cognitive and affective processes for active learning. This especially means that information delivered in classrooms results in differential effects in pupils due to the complexity of their processing characteristics and motivation. This increased attention to individual differences was compatible with the focus on individualization in the sixties and

seventies. The democratization of education with its increasing number of students in secondary and higher education, brought about the need for differentiated learning routes. Moreover, it was expected that individualization could enhance both creativity and competence in 'basic' subject-matter.

In a first approach 'time-on-task' is considered to be the decisive predictive effectiveness variable, in line with Carroll's model (1963) and the mastery-learning strategy (Bloom, 1974). It is hypothesized that powerful indicators for teaching effectiveness are found in observable and measurable actual pupil behaviors. If researchers could find a 'direct' criterion of effectiveness, they could define effectiveness 'on-the-spot' without having to wait for outcomes from achievement tests at the end of the year. In this view, pupil competence is defined as the amount of learning time necessary for task fulfilment. Brophy and Good (1986), in their review of effects of teaching on pupil achievement, report time-on-task as the most general positive finding from research on teaching: "At least two common themes cut across the findings, despite the need for limitations and qualifications. One is that academic learning is influenced by the amount of time that students spend engaged in appropriate learning tasks..." (p. 366). However, the amount of time is not a definite measure for effectiveness, since learning processes activated during the allocated time strongly determine the learning outcomes.

The cognitive-mediational perspective supports conceptions of learning and instruction that are associated with meaningful and 'higher-order' student learning, teacher efficacy, and equity of educational opportunities. Teaching is interpreted by the pupils and their interpretation influences to a high degree the effects of any teaching activity. Here, the most important findings from cognitive research enter the field of teaching effectiveness. The criterion of effectiveness is the cognitive quality of instruction. Anderson (1989) describes the criterion as follows "Learning occurs when learners actively transform incoming information and construct meaning in terms of their prior knowledge. Teachers and others influence learning not through mere transmittal of information but through their influence on students' cognitive processing of that information" (p. 101). In this phase of research, most emphasis is laid on the direct translation of findings from psychological literature, like stimulation of metacognition and self-regulation, schema activation, linking prior knowledge with new information by means of advance organizers, supporting cognitive processes through scaffolding, modelling, cognitive apprenticeship (see Brown, Collins, & Duguid, 1989).

Concerning the social mediation, attention is paid to the analysis of meaning and purposes attributed by the participants to social settings. One example of this approach is the 'hidden curriculum' of schools and classrooms. It is not the classroom context in itself, but the complex attribution of meaning by the actors which is of paramount importance. Anderson (cit. in Shulman, 1986) analyzed ways in which primary-grade children cope with seatwork. "Whereas the process-product researcher would have correlated frequency of seatwork with student achievement, and the ALT scholar would have observed the proportion of seatwork time during which the student was apparently engaged along with the degree of difficulty of the tasks, Anderson proceeds to record student comments to themselves and to classmates during assignments and to interview the students" (p. 16). Her focus is on what students are thinking and feeling as they work on their tasks, and on what these phenomena reveal about the mediation of instruction by the students.

In line with the social mediation paradigm, emphasis is laid on support activities of the peer group. Examples are: reciprocal teaching (Palincsar & Brown, 1984), collaborative knowledge construction in the classroom (Scardamalia & Bereiter, 1992), and cooperative learning (Slavin, 1990). Common effectiveness criteria are explicit goals, individual responsibility, equal opportunities for peers, clear procedure and task relevance. Effectivity studies reveal interesting and lasting outcomes of the original training programs (see Palincsar & Brown, 1984). This effect, however, fades the more the experimental training procedure is implemented under usual classroom conditions (see Palincsar, Brown & Martin, 1987).

Interesting though this line of research may seem, some problematic issues remain as to the definition of effectiveness criteria. In most cases, a direct translation of descriptive findings into

prescriptions for instruction take place without, however, a strong control of the transition steps necessary to validate the effectiveness criteria. As Shulman (1986) remarks "We get too little sense of how variations among tasks, or among forms of teaching, relate to differences in how students mediate those instructional presentations. As with the paradigms discussed earlier, far too little attention is devoted to differences in the content of subject matter being taught. ... We thus learn important new things about teaching from this research, but also forgo parts of the portrayal available from the work in other research programmes" (p. 16).

A positive aspect of this recent line of effectiveness research, however, is the combination of the process-product paradigm with the cognitive and social approach on teaching and learning. It seems a challenging endeavour indeed to search for the most essential effectiveness variables within the complex teaching and learning situation, making use of a parsimonious model.

### TEACHER COGNITIONS

The shift in the psychological scene from behaviorism to cognitivism, along with a refinement of qualitative research methods during the seventies paved the way for studies on teachers' cognitions. Teaching is no more perceived as a quality of a teachers' personality nor as a set of observable behaviors, but as a profession with a call for higher-order cognitive skills in order to cope with the complexity of an educational environment. It is a job resembling the one of physicians, engineers and architects with emphasis on judgment, problem-solving and decision-making. Although this more qualitative line of research started in the sixties (see Smith & Geoffrey, 1969; Jackson, 1968) it was first acknowledged as a research field by the National Institute for Education in its publication: "Teaching as Clinical Information Processing" (1975). In spite of the fact that most studies of teacher's thinking are explorative and descriptive in nature, they nevertheless offer some indications as to effective teaching. At the beginning, most attention is paid to the isolated description of the 'naturalistic' cognitive processes of teachers during the planning, interactive and post-interactive phase of teaching, defined in terms of information processing. The methods used are protocol-analysis, retrospection, stimulated recall, thinking aloud, rep-grid, logbook keeping, etc. Because these methods focus almost exclusively on isolated cognitive processes, no links between cognition and action are studied. Later on, the interaction between the phases becomes object of study, like the interaction between planning and classroom interaction, or the relationship between planning, interaction and post-interactive reflection. One of the most common observations is the psychological way in which teachers function which seems to counter a mere logical teaching model (Lowyck, 1980).

Whereas the description of cognitive processes tends to be content-free, the study of teacher's subjective or implicit theories (teacher beliefs, teacher's practical knowledge) throws light on the very content of their cognitions. It is shown that teachers do not possess a complete nor clear model of teaching, but an eclectic mix of rules of the thumb, generalizations, opinions, values and expectations. An important issue is how the teacher's cognitive understanding of subject-matter content can be elucidated. Shulman (1986) suggests to avoid any unnecessary isolation of types of knowledge and to focus on the intrinsic interaction between subject-matter knowledge, pedagogical knowledge and curricular knowledge.

One of the main problems with this kind of research is the lack of conceptual refinement (see Pajares, 1992) and the closed-circuit characteristics of the research object, namely its reduction to cognitive processes as internal representations without any relationship to teaching behaviors. Moreover, shortage of validity and methodological strength in many idiosyncratic research endeavours caused a lot of free-wheeling descriptions, lacking model building and hypothesis testing. The danger of an over-estimation of idiosyncratic narratives of teachers lays in its disconnection from any search for professional expertise or teaching effectiveness, the latter labeled as a 'positivist assumption' (see Elbaz, 1991, p. 5).

It is clear that insight in the subjective theories of teachers can elucidate their knowledge-base about teaching which allows clarification and classification of subjective teaching models.



However, this line of research is not clear at all about the effectiveness criteria of teaching used. Even if one tries to deduce criteria of effectiveness from expert's cognition and behavior, it is an enormously difficult task due to the idiosyncrasy of the expert's functioning, the vagueness of criteria that should warrant the experienced teacher under study is an expert, the influence of the complex context or task and the difficulty for experts to fully report their cognitive processes or practices.

The integration of teacher cognitions in process-product research was advocated by Berliner (1989), but empirical research in this tradition is scarce. Clark & Peterson (1986), for example, mention only three studies that bear on the relationship between teacher's decision making in classrooms and student achievement. And, the solution proposed by Peterson (1988) to include teacher's cognitional knowledge as a variable in effective teaching did not lead towards solid empirical investigations neither. It seems time, now, to reconsider the paradigms of research on teaching and to "invite scholars in the field to reflect carefully upon the possible effects of centrifugal tendencies in research on teaching so as to look for more integration and cooperation" (Lowyck & Clark, 1989). We do not need a dissension of the research domain on teaching, but as Berliner (1989) contends "There is no doubt that continued descriptive and taxonomic work to better understand the teacher's interactive thoughts and decisions is highly desirable... But I believe that it will be the establishment of relations between interactive thoughts and decision making on the one hand, and criteria of effectiveness on the other, that will lead to the greatest growth of knowledge in this field" (p. 12).

In short, the whole domain of teaching research suffers from a lack of a well functioning apparatus of concepts, models and theories. "The vagueness of the conceptual framework, the babel of unclear, idiosyncratic definitions, the proliferation of models, the isolation of paradigms, and the suggested incompatibility of qualitative and quantitative methodology all hinder the consolidation of thousands of research outcomes. How can we further the understanding of teaching in its complexity, if the concepts of teaching are apparently more complex than the teaching reality itself" (Lowyck, 1990, p. 92-93).

## DISCUSSION

As has been illustrated in the previous sections, research on teaching is strongly influenced by the "Zeitgeist", the state-of-the-art in research, its available methodology and the expectations towards or the frustrations with research outcomes.

First, societal needs, like democratization of citizens ('authoritarian-democratic' paradigm), the competitive edge of society (back-to-the-basics), the adaptation of schools to new target groups (eg. special education), and the fit between school and industry (cognitive flexibility, transfer of knowledge) all are in a subtle way filtered through education policy and plugged into research endeavours. Needless to say that not only social aspects, but also economical-financial factors influence the way education and schooling are perceived.

Second, evolutions within research itself determine to a high degree the focus, expectations and frustrations with regard to paradigms, theories and models. Examples are the successive shift from 'trait psychology' toward behaviorism and cognitivism. These shifts are generated both by disappointing outcomes of the research paradigm at hand and the attractiveness of new avenues.

Third, methodological refinement and availability of new research techniques in different domains, like psychology, sociology or antropology all challenge researchers to find new technical solutions. Examples are ethnography, symbolic interactionism, verbal reports, protocol analysis, rep-grids, as well as multi-level analysis, etc.

Fourth, the conceptualization of teaching dominates the choice of research design. In the case teaching is perceived as the behavioral output of the teacher's personality, emphasis is laid on the study of teachers' characteristics. If teaching is conceptualized as the sum of overt but discrete teaching skills, process-product studies seem appropriate. When teaching is seen as the idiosyncratic model of any individual teacher, teacher cognitions are focused at.



The preceding examples of influences, however, cannot explain all variability in approaches of teaching research. Often more psychological factors determine the research at hand. Examples are the dominant research groups, or the brilliant ideas of individual researchers, like Carroll, Dewey, Gage, Jackson, Vygotsky, who unexpectedly launch influencing research contrary to the dominant line of thought.

Consequently, any review of teaching effectiveness struggles with conceptual, methodological and operational difficulties and affects opinions about the possibility to build synergy in this particular field of research. Some researchers emphasize incompatibility between research paradigms, epistemological backgrounds and suitable methodology. Examples are the opposition of qualitative and quantitative methodology (whatever this means), or the study of teaching as an idiosyncratic attribution of meaning by teachers versus the so-called positivist assumption with a search for more generalizable outcomes. Others claim openness and cooperation between researchers in order to broaden the scope as well as to challenge the integration of originally controversial paradigms. Examples are the recent studies on cognitive apprenticeship, who try to combine a process-product approach with recent insights in cognitive psychology, or the mediating paradigm which inserts new but important variables between teaching and learning outcomes.

This contribution reviewed some lines of thought and empirical work which could contribute to a more synergic approach of the study of teaching effectiveness. It has been revealed that many factors influence both the conceptualization and the design of effectiveness studies. Moreover, the knowledge accumulated through reflection and empirical research during more than half a century, enables at present an incremental view on teaching effectiveness.

## REFERENCES

- Anderson, R.C. (1959). Learning in discussions: a resume of the authoritarian-democratic studies. *Harvard Educational Review*, 29, 20-215.
- Anderson, L.M. (1989). Classroom instruction. In M.C. Reynolds (Ed.). *Knowledge base for the beginning teacher* (pp. 11-115). Oxford/New York: Pergamon.
- Berliner, D.C. (1979). Tempus educare. In P.L. Peterson & H.J. Walberg (Eds). *Research on Teaching*. Berkeley: McCutchan.
- Berliner, D.C. (1989). The place of process-product research in developing the agenda for research on teacher thinking. In J. Lowyck & C.M. Clark (Eds). *Teacher Thinking and Professional Action* (pp. 3-21). Leuven: Leuven University Press.
- Bloom, B.S. (1974). Time and learning. *American Psychologist*, 29, 682-688.
- Brophy, J. (1986). Teacher influences on student achievement. *American Psychologist*, 41, 10, 1069-1077.
- Brophy, J. (1991). Effective schooling and disadvantaged students. In M.S. Knapp & P.M. Shields (Eds). *Better schooling for children of poverty: Alternatives to conventional wisdom* (pp. 211-234). Berkeley: McCutchan.
- Brophy, J. & Good, T.L. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.). *Handbook of Research on Teaching, Third Edition* (pp. 328-375). New York: Macmillan.
- Brown, J.S., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18, 1, 32-42.
- Carnine, D. (1989). Direct instructional model: special education. In T. Husèn & T.N. Postlethwaite (Eds). *International Encyclopedia of Education, Research and Studies. Suppl. Vol. 1* (pp. 251-257). Oxford: Pergamon.
- Carroll, J.B. (1963). A model of school learning. *Teachers College Record*, 64, 724-733.
- Clark, C.M. & Peterson, P.L. (1986). Teachers' thought processes. In M.C. Wittrock (Ed.). *Handbook of Research on Teaching, Third Edition* (pp. 255-296). New York: Macmillan.
- Creemers, B.P.M. (1991). *Effectieve instructie: een empirische bijdrage aan de verbetering van het onderwijs in de klas*. 's Gravenhage: Instituut voor het Onderzoek van het Onderwijs.
- De Corte, E. & Lowyck, J. (1983). Heroriëntatie in het Onderzoek van het Onderwijs. *Tijdschrift voor Onderwijsresearch*, 8, 242-260.
- Doyle, W. (1978). Paradigms for research on teacher effectiveness. In L.S. Shulman (Ed.). *Review of Research on Teaching*. Itasca: Peacock.

- Doyle, W. (1986). Classroom organization and management. In M.C. Wittrock (Ed.). *Handbook of Research on Teaching, Third Edition* (pp. 392-431). New York: Macmillan.
- Dunkin M.J. & Biddle, B.J. (1974). *The Study of Teaching*. New York: Holt, Rinehart & Winston.
- Elbaz, F. (1991). Research on teacher's knowledge: the evolution of a discourse. *Journal of Curriculum Studies*, 23, 1, 1-19.
- Fisher, C.W., Berliner, D.C., et al. (1978). *Teaching and learning in the elementary school: a summary of the Beginning Teacher Evaluation Study*. San Francisco: Far West Laboratory for Educational Research and Development.
- Flanders, N.A. (1965). *Teachers influence, pupil attitudes and achievement*. Washington: U.S. Department of Health, Education and Welfare, Office of Education.
- Flanders, N.A. (1970). *Analyzing teaching behavior*. Reading, Mass: Addison-Wesley.
- Gage, N.L. (1963). Paradigms for research on teaching. In N.L. Gage (Ed.). *Handbook of Research on Teaching* (pp. 94-141). Chicago: Rand McNally.
- Gage, N.L. (1978). *The scientific basis of the art of teaching*. New York: Teachers College Press.
- Getzels, J.W. & Jackson, P.W. (1963). The teacher's personality and characteristics. In N.L. Gage (Ed.). *Handbook of Research on Teaching*. Chicago: Rand McNally.
- Good, T.L. (1979). Teacher effectiveness in elementary school: What we know about it now. *Journal of Teacher Education*, 30, 52-64.
- Heath, R.W. & Nielson, M.A. (1974). The research basis for performance-based teacher education. *Review of Educational Research*, 44, 4, 463-484.
- Jackson, P.W. (1968). *Life in classrooms*. New York: Holt, Rinehart & Winston.
- Lippitt, R. & White, K. (1943). The 'social climate' of children's groups. In R. Barker, J.S. Kounin & H.F. Wright (Eds). *Child behavior and development*. New York: McGraw Hill.
- Lowyck, J. (1980). *A process analysis of teaching (report n° 21): EDRS-ED 190513*. Leuven: K.U. Leuven, Afdeling Didactiek en Psychopedagogiek.
- Lowyck, J. (1990). Teacher thinking studies: bridges between description, prescription and application. In C. Day, M. Pope & P. Denicolo (Eds). *Insights into teachers' thinking and practice* (pp.85-103). London: The Falmer Press.
- Lowyck, J. & Clark, C.M. (1989). Preface. In J. Lowyck & C.M. Clark (Eds). *Teacher thinking and professional action* (pp. I-III). Leuven: Leuven University Press.
- Lowyck, J. & C.M. Clark (Eds). (1989). *Teacher thinking and professional action*. Leuven: Leuven University Press.
- Lowyck, J. & Pieters, J.M. (1993). *The quality of teaching. A research programming study on the quality and the functioning of teachers conducted for the Institute of Educational Research in the Netherlands*. Den Haag: SVO.
- National Institute for Education (1975). *Teaching as clinical information processing*. Washington: U.S. Department of Health, Education and Welfare.
- OECD (1990). *The teacher today. Tasks, conditions, policies*. Paris: Organisation for Economic Co-operation and Development.
- Pajares, M.F. (1992). Teachers' beliefs and educational research: cleaning up a messy construct. *Review of Educational Research*, 62, 307-332.
- Palincsar, A.S. & Brown, A.L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117-175.
- Palincsar, A.S., Brown, A.L. & Martin, S.M. (1987). Peer interaction in reading comprehension instruction. *Educational Psychologist*, 22, 231-253.
- Peterson, P.L. (1988). Teachers' and students' cognitional knowledge for classroom teaching and learning. *Educational Researcher*, 17, 5, 5-14.
- Rosenshine, B. & Furst, N. (1973). The use of direct observation to study teaching. In R.M. Travers (Ed.). *Second Handbook of Research on Teaching* (pp. 122-183). Chicago: Rand McNally.
- Scardamalia, M. & Bereiter, C. (1992). An architecture for collaborative knowledge building. In E. De Corte, M.C. Linn, H. Mandl & L. Verschaffel (Eds). *Computer-based learning environments and problem solving* (NATO/ASI Series F: Computer and System Sciences, Vol. 84) (pp. 41-66). Berlin: Springer.
- Scheerens, J. (1991). Schoolorganisatiekenmerken en condities voor effectieve instructie. In S. Dijkstra, H.P.M. Krammer & J.M. Pieters (Eds). *De onderwijskundig ontwerper* (pp. 19-35). Amsterdam/Lisse: Swets & Zeitlinger.
- Shulman, L.S. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M.C. Wittrock (Ed.). *Handbook of Research on Teaching, Third Edition* (pp. 3-36). New York: Macmillan.

- Simon, A. & Boyer, E.G. (Eds) (1970). *Mirrors for behavior: an anthology of classroom observation instruments*. Philadelphia: Research for Better Schools.
- Slavin, R.E. (1990). *Cooperative learning: Theory, research and practice*. Englewood Cliffs, NJ: Prentice Hall.
- Smith, L.M. & Geoffrey, W. (1968). *The complexities of an urban classroom. An analysis toward a general theory of teaching*. New York: Holt, Rinehart & Winston.
- Snow, R.E. (1986). Individual differences and the design of educational programs. *American Psychologist*, 41, 10, 1029-1039.
- Wang, M.C. (1980). Adaptive instruction: building on diversity. *Theory into Practice*, 19, 122-128.
- Wittrock, M.C. (Ed.) (1986). *Handbook of Research on Teaching, Third Edition*. New York: Macmillan.

Manuscript submitted: October 13, 1993

Accepted for publication: December 22, 1993

## REFERENCES

De opvoeding van kinderen in de kinderopvang wordt in Nederland steeds meer erkend als een belangrijke taak van de overheid. Dit is vooral te zien aan de toename van de kinderopvang plaatsen in de kinderopvanginstellingen. De kinderopvanginstellingen worden steeds meer erkend als een belangrijke taak van de overheid. Dit is vooral te zien aan de toename van de kinderopvang plaatsen in de kinderopvanginstellingen.

In dit artikel wordt vooral aandacht besteed aan de opvoeding van kinderen in de kinderopvanginstellingen. Het wordt onderzocht hoe de opvoeding van kinderen in de kinderopvanginstellingen kan worden verbeterd. Hiervoor wordt gebruik gemaakt van een onderzoek naar de opvoeding van kinderen in de kinderopvanginstellingen.

## THEORETISCHE ACHTERGROND

De opvoeding van kinderen in de kinderopvanginstellingen wordt steeds meer erkend als een belangrijke taak van de overheid. Dit is vooral te zien aan de toename van de kinderopvang plaatsen in de kinderopvanginstellingen.

In dit artikel wordt vooral aandacht besteed aan de opvoeding van kinderen in de kinderopvanginstellingen. Het wordt onderzocht hoe de opvoeding van kinderen in de kinderopvanginstellingen kan worden verbeterd. Hiervoor wordt gebruik gemaakt van een onderzoek naar de opvoeding van kinderen in de kinderopvanginstellingen.

\* Dit onderzoek was onderdeel van het onderzoek naar de opvoeding van kinderen in de kinderopvanginstellingen, uitgevoerd door de Universiteit van Amsterdam.

Tijdschrift voor Onderwijsresearch 13 (1988), nr. 3, pp. 141-150

# Effecten van Klasseconsultatie op Curriculumimplementatie en Leerresultaten

J. Snippe

RION, Instituut voor Onderwijsresearch,  
Rijksuniversiteit Groningen.

## ABSTRACT

In a (quasi-) experimental study\* we examined the effectiveness of facilitator-strategies in improving the implementation of a curriculum as well as the learning results of pupils. Thirty teachers from the fifth grade of the elementary school implemented a new math curriculum. Three groups were formed, two experimental and one control group:

1. *Consultation group*: teachers attended meetings where they received information about the curriculum and discussed problems with fellow teachers. They also received classroom consultation: a classroom consultant observed their lessons and provided them with feedback on their behaviour.
2. *Information group*: teachers attended meetings without classroom consultation.
3. *Control group*: teachers implemented the curriculum without any support.

The results of the experiment show that the consultation group attained the highest implementation as measured by the content covered and the amount of time spent on the curriculum. The learning results of the pupils were also highest in the consultation group. The teachers from the experimental groups exhibited significantly more curriculum related behaviour than did the teachers of the control group. Classroom consultation appears to be an effective strategy to influence curriculum implementation as well as learning outcomes of pupils.

## 1. INLEIDING

De opbrengsten van onderwijsvernieuingsprojecten zijn betrekkelijk teleurstellend te noemen. Er zijn diverse factoren aan te wijzen die hiertoe bijdragen (Van de Grift, 1987). Als een van de belangrijkste oorzaken wordt de gebrekkige implementatie van een onderwijsvernieuwing genoemd. Onder implementatie verstaan we het feitelijk gebruik van een vernieuwing in de praktijk (Fullan en Pomfret, 1977). Voor het welslagen van een onderwijsvernieuingsproject is een goed verloop van het implementatieproces van cruciaal belang.

In dit artikel wordt verslag gedaan van een onderzoek waarin is nagegaan hoe de implementatie door middel van een gerichte begeleiding aan leerkrachten kan worden verbeterd. Hiertoe zijn in een veldexperiment twee begeleidingsinterventies met elkaar vergeleken op hun effectiviteit ten aanzien van de implementatie van een curriculum 'cijferend vermenigvuldigen' en ten aanzien van de leerprestaties.

## 2. THEORETISCHE ACHTERGRONDEN

Uit onderzoek naar implementatie van onderwijsvernieuwingen blijkt dat leerkrachten over het algemeen de vernieuwing niet of slechts gedeeltelijk uitvoeren zoals bedoeld.

Er zijn verschillende factoren aan te wijzen die bijdragen tot het ontstaan van een dergelijk implementatieprobleem (Van de Grift, 1987; Nijhof en Harskamp 1987). In de eerste plaats zijn er factoren in de context waarbinnen een onderwijsvernieuwing wordt geïntroduceerd. Een tweede belangrijke bron voor het ontstaan van implementatieproblemen wordt gevormd door

\* Dit onderzoek was mogelijk dankzij een subsidie van het Instituut voor Onderzoek van het Onderwijs (SVO) te Den Haag, projectnummer 201.4226.

de aard van de vernieuwing. De vaagheid die in vele onderwijsvernieuwingsprogramma's is te bespeuren omtrent de doelstelling en de wijze waarop implementatie dient te geschieden is niet bevorderlijk voor de implementatie (Creemers, 1987).

Als derde factor van belang voor implementatie noemen we de gebruiker van de vernieuwing, de leerkracht. In diverse onderzoeken is gewezen op het feit dat leerkrachten niet of in onvoldoende mate over de kennis en vaardigheden beschikken om een onderwijsvernieuwingsprogramma c.q. methode te gebruiken volgens de bedoelingen van de auteurs (Appelhof, 1979; Nijhof, 1978). Begeleiding van leerkrachten bij de invoering van methoden zou een belangrijke rol kunnen spelen in het voorkomen van implementatieproblemen.

De vierde en laatste belangrijke factor in het implementatieproces is de strategie waarmee een onderwijsvernieuwingsprogramma in de praktijk wordt geïntroduceerd. Hierbij denken we zowel aan innovatiestrategieën die door de diverse begeleidingsinstanties worden gehanteerd als aan concrete begeleidingsinterventies die door schoolleiders en/of schoolbegeleiders worden gebruikt (Van de Grift, 1987).

In het experiment waarvan we hier verslag doen staat de beïnvloeding van het gedrag van leerkrachten centraal als middel om implementatie te bevorderen. Alvorens over te gaan tot de beschrijving van dit experiment, gaan we eerst nader in op de inhoud en vormgeving van een doelmatige curriculumbegeleiding.

### 2.1. Inhoud van curriculumbegeleiding

Met een doelmatige curriculumbegeleiding wordt een zodanige beïnvloeding van het gedrag van de leerkracht beoogd dat deze in staat wordt gesteld tot het implementeren van het curriculum volgens de bedoelingen van de auteurs. Om vast te stellen welke leerkrachtgedragingen essentieel zijn voor de implementatie van het curriculum gaan we uit van de veronderstelling dat curriculuminnovatie uiteindelijk tot doel heeft te komen tot verbetering van de leerprestaties van leerlingen. We gaan daarom na welke componenten van het curriculum bij toepassing in de klas situatie zullen leiden tot een verbetering van de leerprestaties van leerlingen.

Uit onderzoek is inmiddels naar voren gekomen, dat de hoeveelheid leerstof die wordt behandeld, de hoeveelheid tijd die hieraan wordt besteed, alsmede de mate waarin de leerlingen taakgericht bezig zijn tijdens de les, belangrijke factoren zijn voor het verklaren van verschillen in leerprestaties (Veenman, 1986; Harskamp & Suhre, 1986).

We leggen vervolgens de relatie met leerkrachtgedrag door na te gaan welke leerkrachtgedragingen leiden tot een grotere hoeveelheid behandelde leerstof, lestijd en tot een hoge taakgerichtheid van leerlingen. Rosenshine en Stevens (1986) noemen in dit verband het regelmatig controleren van leerlingen en hun leerprestaties, het duidelijk stellen van doelen voor leerlingen, duidelijkheid in het geven van opdrachten en het stellen van vragen aan leerlingen waarop directe feedback mogelijk is. Dergelijke leerkrachtgedragingen zijn gelegen in de voorwaardelijke sfeer en zijn voor de implementatie van elk curriculum noodzakelijk.

Daarnaast zijn er componenten aan te wijzen die specifiek voor het juiste gebruik van één bepaald curriculum zijn. Zij geven bij toepassing in de klas vorm aan het unieke en onderscheidende karakter van dit ene curriculum. In de moderne reken-wiskunde methoden bijvoorbeeld is het voeren van een leergesprek met de leerlingen naar aanleiding van een probleemsituatie een vorm van curriculumspecifiek leerkrachtgedrag dat essentieel is voor een juiste uitvoering van dergelijke methoden. Van deze curriculumspecifieke gedragingen is niet vast komen te staan dat ze bij toepassing in de klas leiden tot betere leerprestaties van leerlingen (zie Harskamp & Suhre, 1986).

De inhoud van een effectieve curriculumbegeleiding zal derhalve gericht moeten zijn op zowel algemeen onderwijskundige als curriculumspecifieke gedragingen.

### 2.2. Vormgeving van curriculumbegeleiding

In de jaren '70 heeft een aantal onderzoeken plaatsgevonden naar in-service training van leerkrachten. Door middel van het oefenen van geïsoleerde onderwijsvaardigheden bracht men gedragsveranderingen bij leerkrachten teweeg. Een voorbeeld hiervan is de minicursus (zie



bijvoorbeeld Veenman en Heeringa, 1977; de Vries, 1980).

Om de transfer van de geoeffende vaardigheden naar de klaspraktijk te bevorderen, wordt de laatste tijd meer nadruk gelegd op een vorm van in-service training waarbij het geven van feedback aan leerkrachten over de toepassing van vaardigheden in de klas centraal staat (Tillema, 1986).

In een analyse van een aantal onderzoeken naar begeleiding van leerkrachten geven Joyce en Showers (1980) aan dat begeleiding van leerkrachten in de eigen classesituatie een onmisbare aanvulling op cursusbijeenkomsten is, indien het doel van de begeleiding het invoeren van een nieuw curriculum betreft. Het gaat dan om een intensievere vorm van begeleiding, klasseconsultatie genaamd, die tot doel heeft om aan de hand van gestructureerd observeren en het geven van feedback over het onderwijsgedrag van de leerkracht, te komen tot een verandering van het lesgeefgedrag in de richting die het curriculum vereist. Daar de leerkracht in de eigen classesituatie wordt geobserveerd, is het mogelijk om gerichte feedback te verschaffen over de mate waarin deze het bedoelde gedrag vertoont en over aanpassingen ten behoeve van de context waarin de leerkracht werkzaam is. Hoewel vooral in Amerika klasseconsultatie als begeleidingsmiddel sterk wordt gepropageerd, ontbreekt de empirische basis omtrent de effectiviteit ervan vrijwel geheel. Showers (1982) doet verslag van een experiment waarin klasseconsultatie als aanvullend begeleidingsmiddel is onderzocht op de training van onderwijsvaardigheden. Hiertoe bezochten twee groepen leerkrachten cursusbijeenkomsten waar men informatie ontving over nieuwe onderwijsvaardigheden. Na afloop ontving de ene groep leerkrachten klasseconsultatie; de andere groep fungeerde als controlegroep. Uit klasseobservaties bleek dat de klasseconsultatiegroep significant meer van de gewenste gedragingen vertoont dan de controlegroep. Ten aanzien van de leerlingprestaties in beide groepen was er slechts sprake van een gering verschil. De onderzoeker concludeert dat klasseconsultatie een noodzakelijke aanvulling op de bijeenkomsten vormt, wil er sprake zijn van overdraagbaarheid van de geleerde vaardigheden naar de classesituatie. Ten aanzien van de opzet van een effectieve curriculumbegeleiding, kunnen we naar aanleiding van het bovenstaande twee begeleidingsvormen onderscheiden:

1. cursusbijeenkomsten met klasseconsultatie en
2. cursusbijeenkomsten zonder klasseconsultatie.

Wij voerden een experiment uit met deze twee begeleidingsvormen als experimentele condities en daarnaast een controlegroep zonder cursusbijeenkomsten.

De onderzoeksvragen voor dit experiment luiden als volgt:

1. Leidt de begeleiding van leerkrachten tot een grotere mate van implementatie van het curriculum dan geen begeleiding?
2. Leidt de begeleiding in de vorm van cursusbijeenkomsten, aangevuld met klasseconsultatie tot een grotere mate van implementatie van het curriculum dan cursusbijeenkomsten zonder klasseconsultatie?
3. Leidt begeleiding van leerkrachten tot hogere leerprestaties bij leerlingen dan geen begeleiding?
4. Leidt begeleiding in de vorm van cursusbijeenkomsten, aangevuld met klasseconsultatie tot hogere leerprestaties bij leerlingen dan cursusbijeenkomsten zonder klasseconsultatie?

### 3. ONDERZOEKSOPZET

Het onderzoek kent een (quasi-) experimenteel design. Leerkrachten van groep 5 van de basisschool voeren een nieuw curriculum 'cijferend vermenigvuldigen' in de klas in. Zij zijn verdeeld over drie begeleidingscondities, twee experimentele condities en een controlegroep. In de eerste experimentele conditie volgen leerkrachten cursusbijeenkomsten die worden aangevuld met klasseconsultatie (consultatiegroep). In de tweede experimentele conditie volgen leerkrachten eveneens cursusbijeenkomsten; er wordt geen klasseconsultatie toegevoegd (cususgroep). De leerkrachten uit de controlegroep ontvangen geen begeleiding bij de invoering van het curriculum. De onafhankelijke variabele wordt gevormd door de begeleidingsconditie en de afhankelijke

variabelen zijn respectievelijk de mate van implementatie van het curriculum en de leerprestaties van leerlingen.

### 3.1. Het curriculum

In het experiment wordt gebruik gemaakt van een curriculum 'cijferend vermenigvuldigen', dat binnen het RION is ontwikkeld. In de 20 lessen die dit curriculum telt, wordt op inzichtelijke wijze het cijferend vermenigvuldigen behandeld. Dit gebeurt volgens het principe van progressieve schematisering: vanuit concreet visueel weergegeven schema's van handeling ontwikkelt zich langzamerhand via het aanbrenge van verkortingen in de schema's het eindalgoritme. De rekenopgaven worden aan de leerlingen aangeboden in verhaalvorm. Tevens wordt de leerlingen geleerd om systematisch gebruik te maken van schattingen als controle op de uitgevoerde rekenbewerking.

Het werken met deze leergang is op een aantal punten verschillend van het werken met een traditionele rekenmethode. In de leergang heeft de instructie een interactief karakter: de leerstof wordt in de vorm van een leergesprek met de leerlingen aan de orde gesteld. In traditionele methoden heeft de leerkracht tijdens de instructie vaak een sterk docerende rol. Na de instructie werken de leerlingen met de leergang aan een aantal oefeningen die zeer gevarieerd zijn samengesteld, in tegenstelling tot traditionele methoden waar de zelfstandige verwerkingsopgaven per les gelijksoortig van aard zijn.

### 3.2. Inhoud van de begeleidingscondities

De leerkrachten uit de beide experimentele condities (de consultatiegroep en de cursusgroep) volgen vijf cursusbijeenkomsten in het tijdvak van drie maanden waarin het curriculum wordt uitgevoerd.

Tijdens deze bijeenkomsten worden de leerkrachten geïnformeerd over de achtergronden en inhoud van het curriculum en krijgen ze de gelegenheid om ervaringen uit te wisselen met collega-leerkrachten en knelpunten te bespreken bij de uitvoering van de leergang. In deze bijeenkomsten worden tevens de lessen van de leergang voorbesproken, waarbij aandacht wordt besteed aan de curriculumspecifieke gedragingen, namelijk het geven van interactieve instructie en het organiseren van zelfstandig werken van de leerlingen aan gevarieerde leertaken en aan de gedragingen in de voorwaardelijke sfeer, namelijk het optimaal besteden van de beschikbare lestijd, duidelijkheid in het geven van opdrachten, controleren van de leerlingen. Er worden voorbeelden gegeven van deze gedragingen die vervolgens worden besproken.

In de consultatiegroep worden de cursusbijeenkomsten aangevuld met een drietal klassebezoeken van de klasseconsultant. Deze observeert een les van de leerkracht aan de hand van een checklist die items bevat met betrekking tot het algemeen onderwijskundig handelen van de leerkracht tijdens de les en met betrekking tot de curriculumspecifieke gedragingen van de leerkracht.

Na afloop van de les vindt er een nabespreking met de leerkracht plaats, aan de hand van de punten uit de checklist. In overleg met de leerkracht worden aanbevelingen gedaan ter verbetering van eventuele knelpunten die zijn signaleerd. Het feedback-gesprek wordt afgesloten met een aantal concrete afspraken over de aanbevelingen die zijn gedaan.

Zowel de cursusbijeenkomsten als de klasseconsultaties zijn verzorgd door één onderzoeksmedewerker, die verbonden is aan het project. De leerkrachten uit de controlegroep hebben – evenals de overige leerkrachten – het curriculum inclusief de handleiding ontvangen. In een begeleidende brief is de technische gang van zaken rond het onderzoek toegelicht. Eens in de drie weken zijn ze telefonisch benaderd door de onderzoekers om de voortgang te bewaken. De controlegroep heeft geen inhoudelijke begeleiding ontvangen, noch was er sprake van systematische overdracht van kennis en vaardigheden.

### 3.3. Samenstelling van de steekproef

De leerkrachten zijn door de onderzoekers telefonisch benaderd met het verzoek om medewerking aan het onderzoek. In verband met het bezoeken van de cursusbijeenkomsten, heeft de werving

per regio plaatsgevonden.

De in totaal 30 leerkrachten zijn allen werkzaam in groep 5 van de basisschool en afkomstig uit het noorden van het land.

De leerkrachten die deelname aan het onderzoek hebben geweigerd, deden dit in de eerste plaats vanwege een gebrek aan tijd en in de tweede plaats vanwege de inhoud van het curriculum. (Men heeft geen interesse in het vak rekenen of men gebruikt al een moderne reken/wiskundemethode.) Het is derhalve niet ondenkbaar, dat er enige vorm van selectie heeft plaatsgevonden bij de samenstelling van de steekproef. Mogelijkerwijs zijn de leerkrachten uit de steekproef iets meer gemotiveerd om hun rekenonderwijs te vernieuwen dan de leerkrachten die hebben geweigerd.

De begeleidingscondities zijn willekeurig aan de regio's toegewezen.

## 4. INSTRUMENTEN

### 4.1 Implementatie-instrumenten

De eerste twee vragen voor dit experiment hebben betrekking op de invloed van (de soort) begeleiding op de mate van implementatie van het curriculum.

Om de afhankelijke variabele, de implementatie, te operationaliseren is in de eerste plaats een korte vragenlijst voor leerkrachten ontwikkeld, waarmee wordt nagegaan of de leergang daadwerkelijk in gebruik is genomen. Er wordt aan de leerkrachten gevraagd hoeveel lessen ze van de leergang hebben behandeld en hoeveel tijd ze aan de leergang hebben besteed. Deze vragenlijst is na afloop van het experiment door alle leerkrachten ingevuld.

In de tweede plaats is een observatiesysteem ontwikkeld om aan de hand hiervan klasse-observaties te verrichten. Het observatiesysteem bestaat uit een checklist met acht items (dichotoom) die zijn ondergebracht in de twee subgroepen. Voor de ene subgroep van curriculumspecifieke gedragingen wordt de somscore van 5 items berekend. De tweede subgroep van de checklist - de algemeen onderwijskundige gedragingen - behoort tot de gedragingen in de voorwaardelijke sfeer; hiertoe wordt de somscore van 3 items berekend. Tevens wordt het controlerend gedrag van de leerkracht in kaart gebracht door per tijdseenheid van 5 minuten te noteren of de leerkracht controlerend gedrag vertoont; per les wordt hier een somscore van berekend. Ook dit controlerend gedrag hoort tot de gedragingen in de voorwaardelijke sfeer.

Van de leerlingen wordt tijdens de les de niet-taakgerichtheid gescoord als indicator voor de actieve leertijd. Dit gebeurt door elke vijfde minuut te noteren hoeveel leerlingen er zichtbaar niet met de leertaak bezig zijn. Per les wordt dan het percentage niet-taakgerichtheid van de hele klas berekend.

De vier observatoren die zijn ingeschakeld voor de dataverzameling, hebben een training ontvangen in het hanteren van het observatiesysteem. (De onderzoeksmedewerker die de cursusbijeenkomsten en de klasseconsultaties verzorgde is niet betrokken geweest bij de observaties.) Teneinde de interbeoordelaarsbetrouwbaarheid te bepalen is per subgroepje van de gedragingen de Cohen's Kappa berekend. De resultaten van de vier observatoren zijn als volgt:

- curriculumspecifieke gedragingen: Kappa = .91
- algemeen onderwijskundige gedragingen: Kappa = .83
- controlerend gedrag: Kappa = .74

Er blijkt sprake van een redelijke mate van overeenstemming tussen de observatoren. Bij alle leerkrachten heeft in de laatste weken van het experiment een observatie plaatsgevonden.

### 4.2. Prestatietoetsen

De laatste twee vragen voor dit experiment hebben betrekking op de invloed van (de soort) begeleiding op de leerprestaties van leerlingen. Voor beantwoording van deze vragen is er bij de leerlingen een aantal prestatietoetsen afgenomen.

Om het beginniveau van de leerlingen te kunnen inschatten, is voorafgaand aan het experiment een toets algemene rekenvaardigheid van het CITO afgenomen (M3B, midden leerjaar 3, multiple choice) en een toets Redactie-opgaven die op het RION is ontwikkeld. Deze toets bevat tien

betrokelijk eenvoudige vermenigvuldigings- en deelsommen (open items). De interne consistentie van deze toets is vastgesteld door de betrouwbaarheidscoëfficiënt alpha te berekenen; deze bedraagt . 71. De toets is redelijk consistent.

Na afloop van het experiment is opnieuw een CITO-toets algemene rekenvaardigheid afgenomen, (E3A, eind leerjaar 3, 40 multiple choice items) en een RION-toets. Deze toets bevat 15 redactie-opgaven (open items) die voor het merendeel betrekking hebben op het cijferend vermenigvuldigen. Bij de helft van de items wordt van de leerlingen gevraagd om een schatting van het antwoord op de opgave te geven. De belangrijkste doelstellingen van het curriculum zijn in deze toets vertegenwoordigd. De interne consistentie van deze toets is hoog ( $\alpha = . 87$ ). De afname van de toetsen heeft zowel bij de voor- als bij de nameting plaatsgevonden in twee gescheiden dagdelen.

## 5. RESULTATEN

De gegevens betreffende de implementatie van het curriculum, zijn geanalyseerd om de vraag naar een eventueel differentieel effect van de begeleidingsconditie op de mate van implementatie van het curriculum te beantwoorden.

In de eerste plaats is gekeken naar de gegevens van de vragenlijst over de uitvoering van het curriculum. In onderstaande tabel worden de antwoorden op de vraag naar het aantal lessen dat de leerkracht van het curriculum heeft behandeld en de hoeveelheid tijd die hieraan is besteed in gemiddelden weergegeven, met tussen haakjes de standaarddeviaties vermeld.

TABEL 1 *Hoeveelheid behandelde leerstof van het curriculum en hoeveelheid tijd, besteed aan het curriculum. Uitgedrukt in gemiddelden en standaarddeviaties.*

|                         | gem. aantal behandelde lessen<br>(maximaal 20) | gem. bestede lestijd (in minuten) |
|-------------------------|--|-----------------------------------|
| consultatiegroep (n=13) | 17.7 (3.3)                                     | 1263.1 (311.4)                    |
| cursusgroep (n=10)      | 12.8 (3.3)                                     | 893.0 (207.0)                     |
| controlegroep (n=7)     | 11.4 (6.9)                                     | 777.9 (386.2)                     |

Er is een statistisch significant verschil tussen de condities met betrekking tot het aantal behandelde lessen ( $F$ -waarde  $(2,27) = 6.2$ ;  $p=.0$ ). Ook het verschil ten aanzien van de hoeveelheid tijd die aan de leergang is besteed, is significant ( $F$ -waarde  $(2,27) = 7.4$ ;  $p=.01$ ). Het blijkt dat de consultatiegroep significant meer lessen van het curriculum heeft behandeld en meer tijd eraan heeft besteed dan de overige condities. De controlegroep scoort het laagst op de behandelde lesstof en de bestede lestijd.

Vervolgens zijn de gegevens geanalyseerd die uit de observatie zijn verkregen. Het gaat hier om

TABEL 2: *Observatiescores per conditie, uitgedrukt in gemiddelden en standaarddeviaties.*

|                         | curriculumspecifieke gedragingen<br>(range 0-5) | algemeen onderw. kundige gedragingen<br>(range 0-3) | controlerend gedrag van leerlingen<br>(range 1-12) | niet-taakgerichtheid van leerlingen<br>(range 1-100%) |
|-------------------------|---|---|--|---|
| consultatiegroep (n=12) | 3.4<br>(0.7)                                    | 1.0<br>(0.9)  | 4.2<br>(3.1)                                       | 17.6<br>(13.3)  |
| cursusgroep (n=9)       | 3.6<br>(1.0)                                    | 1.6<br>(0.9)  | 4.7<br>(2.6)                                       | 17.1<br>(10.1)  |
| controle-groep (n=4)    | 1.8<br>(1.3)                                    | 1.3<br>(1.3)  | 2.3<br>(2.1)                                       | 13.4<br>(8.3)   |



de curriculum specifieke gedragingen, de algemeen onderwijskundige gedragingen en het controlerend gedrag van de leerkracht alsmede de niet-taakgerichtheid van de leerlingen.

In tabel 2 worden de observatiegegevens per conditie voor de onderscheiden categorieën vermeld, uitgedrukt in gemiddelde scores en standaarddeviaties. In totaal zijn er observatiegegevens van 25 leerkrachten beschikbaar.

Er is sprake van een significant verschil tussen de condities op de categorie curriculum specifieke gedragingen. (F-waarde (2,22) = 6.22;  $p = .0$ ). Dit verschil is gelegen tussen de consultatie- en cursusgroep enerzijds en de controlegroep anderzijds. De controlegroep scoort het laagst op de curriculum specifieke gedragscategorie.

Samenvattend kunnen we ten aanzien van de eerste twee onderzoeksvragen concluderen dat begeleiding leidt tot een grotere mate van implementatie in de zin van hoeveelheid behandelde leerstof en lestijd. Er is tevens sprake van een extra effect van klasseconsultatie op de implementatie.

Begeleiding van leerkrachten leidt er eveneens toe dat zij meer curriculum specifieke gedragingen vertonen en derhalve het curriculum meer uitvoeren in overeenstemming met de bedoelingen. Op dit punt is er geen extra effect van klasseconsultatie gebleken. Ten aanzien van de overige leerkrachtgedragingen zijn er geen significante verschillen tussen de begeleidingscondities geconstateerd, noch ten aanzien van de taakgerichtheid van leerlingen.

We gaan vervolgens na wat de invloed is van de begeleidingsconditie op de leerprestaties van leerlingen. Hiertoe beschouwen we eerst de relatie tussen de mate van implementatie en de leerresultaten van leerlingen. De mate van implementatie is in dit experiment geoperationaliseerd als de hoeveelheid behandelde leerstof en de hoeveelheid tijd die aan de leerstof is besteed en de mate waarin de leerkracht curriculum specifieke en algemeen onderwijskundige gedrag vertoont.

Om zicht te krijgen op de samenhang tussen de implementatiefactoren enerzijds en de leerprestaties van de leerlingen anderzijds, zijn correlaties berekend. De resultaten worden in onderstaande tabel weergegeven. Hierbij is voor beide toetsen gecorrigeerd voor de resultaten van de leerlingen op de voormeting van de CITO-toets.

TABEL 3: *Partiële correlaties tussen implementatiefactoren en leerprestaties (gecorrigeerd voor CITO-voortoets)*

|  | natoets CITO         | natoets Redactie-opgaven |
|--|----------------------|--------------------------|
| aantal behandelde lessen uit de leergang | .19<br>( $p=.196$ )  | .65<br>( $p=.000$ )      |
| hoeveelheid tijd besteed aan de leergang | .23<br>( $p=.145$ )  | .55<br>( $p=.003$ )      |
| curriculum specifieke gedragingen        | .03<br>( $p=.447$ )  | -.08<br>( $p=.367$ )     |
| algemeen onderwijskundige gedragingen    | -.14<br>( $p=.267$ ) | .30<br>( $p=.079$ )      |
| controlerend gedrag                      | .48<br>( $p=.010$ )  | .30<br>( $p=.084$ )      |
| niet-taakgerichtheid van leerlingen      | -.33<br>( $p=.064$ ) | -.22<br>( $p=.154$ )     |



Uit de tabel blijkt dat het controlerend gedrag van de leerkracht significant samenhangt met de (gecorrigeerde) CITO-natoetsscores. Met de (gecorrigeerde) scores op de natoets Redactie-opgaven hangen de hoeveelheid behandelde lessen uit de leergang en de hoeveelheid tijd die hieraan is besteed, significant samen.

Alvorens we deze implementatiefactoren in verband brengen met de leerprestaties van de leerlingen, gaan we eerst na of er verschillen zijn tussen de leerlingen op de toetsscores. We geven een overzicht van de gemiddelde klassescores op de voor- en natoetsen van de drie condities. In dit overzicht zijn ter vergelijking tevens de scores opgenomen van de leerlingen uit de landelijke steekproef die door het CITO als normgroep wordt gehanteerd.

TABEL 4: Gemiddelde klassescores op de voor- en natoetsen van de drie condities, met tussen haakjes de standaarddeviaties.

|  | voortoets<br>CITO<br>(40 items) | natoets CITO<br>(40 items) | voortoets<br>Redactie-<br>opgaven<br>(10 items) | natoets<br>Redactie-<br>opgaven<br>(15 items) |
|--|---------------------------------|----------------------------|---|---|
| consultatie<br>groep (n=13)                          | 28.8<br>(4.0)                   | 26.5<br>(4.2)              | 4.6<br>(1.4)                                    | 8.0<br>(2.4)                                  |
| cursusgroep<br>(n=10)                                | 29.1<br>(2.5)                   | 23.4<br>(3.8)              | 4.6<br>(1.5)                                    | 6.4<br>(2.8)                                  |
| controle-<br>groep (n=6)                             | 28.3<br>(3.9)                   | 23.1<br>(4.5)              | 4.1<br>(0.8)                                    | 5.4<br>(3.2)                                  |
| normgroep<br>CITO voor-<br>toets (687<br>leerlingen) | 31.5<br>(7.1)                   |                            |   |   |
| normgroep<br>CITO na-<br>toets (852<br>leerlingen)   |                                 | 24.9<br>(7.8)              |   |   |

Het blijkt dat onze steekproef een lagere startpositie had dan de landelijke steekproef van het CITO; dit verschil is ingehaald op de natoets. De moeilijkheidsgraad van de CITO-natoets ligt hoger dan van de CITO-voortoets.

Er zijn op de voormeting geen verschillen tussen de drie condities, noch op de voortoets Redactie-opgaven ( $F$ -waarde  $(2,26) = .33$ ;  $p = .723$ ), noch op de voortoets CITO ( $F$ -waarde  $(2,26) = .11$ ;  $p = .898$ ). Er is ook geen sprake van een significant interactie-effect tussen begeleidingsconditie en voormeting, noch voor de toets Redactie-opgaven, noch voor de CITO-toets. Vervolgens is getoetst of er verschillen in gemiddelde klassescores op de natoets bestaan die zijn toe te schrijven aan de onafhankelijke variabele 'begeleidingsconditie'. Met behulp van covariantie-analyse zijn eerst de klassescores op de natoets Redactie-opgaven als afhankelijke variabele genomen, waarbij de klassescores op de voortoets CITO als covariaat zijn ingevoerd. De resultaten van deze analyse geven een significant effect van begeleidingsconditie te zien op de natoets Redactie-opgaven ( $F$ -waarde  $(2,22) = 3.47$ ;  $p = .047$ ).

Deze analyse is eveneens uitgevoerd voor de klassescores op de natoets CITO en hier blijkt ook dat er een significant verschil bestaat dat terug te voeren is op de begeleidingsconditie ( $F$ -waarde  $(2,22) = 8.10$ ;  $p = .002$ ).

Samenvattend kunnen we ten aanzien van de derde en vierde onderzoeksvraag concluderen dat begeleiding van leerkrachten in de vorm van cursusbijeenkomsten aangevuld met klasseconsultatie leidt tot hogere leerprestaties van leerlingen, dan begeleiding zonder klasseconsultatie en geen

begeleiding. Het verschil in leerprestaties tussen de cursusgroep en de controlegroep is niet significant.

### 5.1. Verbanden tussen begeleidingsconditie, implementatie en leerresultaten.

Nu gebleken is dat er verschillen bestaan in leerprestaties voor de verschillende condities, is het zinvol om na te gaan via welke implementatiefactoren deze verschillen kunnen worden verklaard. Uit eerdere analyses was gebleken dat de drie condities van elkaar verschillen op twee belangrijke voorspellende variabelen ten aanzien van leerprestaties, namelijk de 'hoeveelheid tijd aan de leergang besteed' en 'het aantal behandelde lessen'. Het is derhalve zinvol om deze variabelen als controle variabelen in een analyse in te voeren.

Er is een covariantie-analyse uitgevoerd met de klasscores op de natoets Redactie-opgaven als afhankelijke variabele en de klasscores op de CITO-voortoets als covariaat, waarbij door middel van step-down toetsing de invloed van de controlevariabelen - de hoeveelheid tijd aan de leergang besteed en het aantal behandelde lessen - op de leerprestaties is uitgeschakeld.

De resultaten van deze analyse geven aan dat er geen significante verschillen meer zijn tussen de condities op de nameting van de RION-toets ( $F$ -waarde  $(2,20) = .12$ ;  $p = .889$ ). Het oorspronkelijk gevonden effect van de begeleidingsconditie op deze curriculumspecifieke natoets kan worden toegeschreven aan de invloed van de variabelen 'hoeveelheid tijd aan de leergang besteed' en 'hoeveelheid behandelde leerstof'.

Bovenstaande covariantie-analyse met step-down toetsing is eveneens uitgevoerd met de klasscores op de natoets CITO als afhankelijke variabele.

De resultaten van deze analyse geven een significant effect van begeleidingsconditie op de natoets CITO te zien. ( $F$ -waarde  $(2,20) = 6.43$ ;  $p = .006$ ). Hieruit blijkt dat na uitpartialisering van de variantie door de ingevoerde controlevariabelen (hoeveelheid tijd en behandelde leerstof) er nog een significant deel van de variantie wordt verklaard door de factor begeleidingsconditie. Dit effect kan niet worden verklaard met de variabelen die in het onderzoek zijn opgenomen. Daar de leerkrachten uit de verschillende condities niet verschillen op die gedragingen die voorspellende waarde bezitten ten aanzien van de leerprestaties van leerlingen, is het effect van conditie van daaruit niet te verklaren.

Het is echter mogelijk dat de leerkrachten (in de tijd waarin het experiment is uitgevoerd) meer leerstof uit de eigen methode hebben behandeld of op intensievere wijze met de eigen methode werkten. Dit is een factor die in het onderzoek niet is onderzocht. We hebben hier kennelijk te maken met een soort transfer-effect van het werken met het curriculum 'cijferend vermenigvuldigen' op de algemene rekenvaardigheid van de leerlingen. Dit transfer-effect is mogelijkterwijs tot stand gekomen via leerkrachtgedragingen die we niet onderzochten. Het is eveneens mogelijk dat leerkrachten tijdens de klasseconsultaties zijn geweest op een aantal knelpunten in het lesgeefgedrag, welke bij opheffing tot verbetering van de resultaten van de leerlingen leiden.

## 6. CONCLUSIE EN AANBEVELINGEN

De resultaten van het experiment geven aan dat de leerkrachten die begeleiding hebben ontvangen in de vorm van cursusbijeenkomsten, aangevuld met klasseconsultatie, tot een hogere mate van implementatie van het curriculum komen in de zin van de hoeveelheid behandelde leerstof en de hoeveelheid bestede lestijd. In die zin draagt deze begeleidingsinterventie bij tot een hogere mate van implementatie en tevens tot hogere leerprestaties van de leerlingen.

Voorts blijkt dat leerkrachten die begeleiding hebben ontvangen tot een hogere mate van implementatie komen in die zin dat zij meer curriculumspecifieke gedragingen vertonen dan de controlegroep. Dit wijst erop dat door het gezamenlijk voorbereiden en bespreken van lessen bij de leerkrachten een duidelijker beeld is ontstaan van het gewenste gedrag volgens het curriculum, dan in de controlegroep. Daar er geen verschil in leerkrachtgedragingen bestaat tussen de consultatie- en de cursusgroep, kunnen we vooralsnog de stelling, dat klasseconsultatie via gerichte feed-back in de klas tot een grotere gedragsverandering bij leerkrachten leidt, niet ondersteunen.

Het effect van klasseconsultatie is, dat de leerkrachten meer lessen uitvoerden en meer tijd aan

aan de leergang besteedden. Het is mogelijk, dat de leerkrachten door de klasseconsultaties extra gemotiveerd raakten tot het uitvoeren van de lessen. Een andere verklaring is dat de leerkrachten de klasseconsultaties als vorm van controle op de uitvoering hebben beschouwd.

In een volgend experiment zou hierover meer informatie kunnen worden verschaft door twee condities op te nemen waarin klasseconsultatie wordt toegepast.

#### REFERENTIES

- Appelhof, P. (1979). *Begeleide onderwijsvernieuwing*. Tilburg: Zwijzen.
- Creemers, B.P.M. (1987). Formatieve curriculumevaluatie, een heroverweging van de CURVO-strategie. *Pedagogische Studiën*, 64, p. 59-66.
- Fullan, M. en A. Pomfret (1977). Research on curriculum and instruction implementation. *Review of Educational Research*, Vol. 47, no. 1, p. 335-397.
- Grift, W. Van de (1987). *Implementatie van vernieuwingen: de rol van de schoolleider*. Dissertatie. Leiden: Rijksuniversiteit.
- Harskamp, E. en C. Suhre (1986). *Vergelijking van rekenmethoden in het basisonderwijs*. Eindrapport, SVO-project 9.4035. Groningen: RION.
- Joyce, B. en B. Showers (1980). Improving in-service training: the messages of research. *Educational Leadership*, februari, p. 379-385.
- Nijhof, W.J. (1978). *Interne differentiatie als een innovatie*. SVO-reeks, no. 9.
- Nijhof, W.J. en E. Harskamp (1987). Van curriculumontwerp naar curriculumeffect. In: Harskamp, E. en W.J. Nijhof (red.) *Curriculum: ontwerp, implementatie en evaluatie*. Bijdragen aan de Onderwijsresearch, 13. Lisse: Swets en Zeitlinger.
- Rosenshine, B. en R. Stevens (1986). Teaching functions. In: Wittrock, M.C. (ed) *Handbook of research on teaching*. Third edition. New York.
- Showers, B. (1982). *Transfer of training: the contribution of coaching*. Beverly: Oregon University, Eugene. Center for Educational Policy and Management, December.
- Snippe, J. (1986). *Effecten van klasseconsultatie op curriculum-implementatie en leerresultaten*. Subsidie-aanvraag OTG-curriculum. Groningen: RION.
- Snippe, J. en E. Harskamp (1987). *Effecten van in-service training op curriculum-implementatie en leerresultaten*. Eindrapport SVO-project 201.4226. Groningen: RION.
- Tillema, H.H. (1986). *Effectiviteit van nascholingstrainingen voor leerkrachten*. Onderzoeks-aanvraag. Groningen: RION.
- Veenman, S. en J. Heeringa (1977). Ontwikkeling en evaluatie van de minikursus: Het organiseren van zelfstandig leren. *Pedagogische Studiën*, 54, p. 176-189.
- Veenman, S. en P. Lem (1986). Groepen en differentiëren in combinatieklassen. In: A. Reints en P. Span (red.) *Differentiatie in het onderwijs*. Bijdragen aan de Onderwijsresearch, no. 11. Lisse: Swets & Zeitlinger.
- Vries, A.K. de (1980). Inhouden en achtergronden van de in-service trainingscursussen in het GEON-project. *Pedagogische Studiën*, 59, p. 449-459

*Manuscript ontvangen 29-4-1987*

*Definitieve versie ontvangen 16-3-1988*

Tijdschrift voor Onderwijsresearch, 18 (1993), nr. 5, pp. 295-305

## Aanpak van kwalificatieproblemen in organisaties: een onderzoek naar strategieën van opleidingskundigen

F.J. van der Krogt, A.A. Vermulst en I.E.C. ter Woerds  
*Vakgroep Algemene Pedagogiek, Katholieke Universiteit Nijmegen\**

### ABSTRACT

Training managers are involved with qualification problems of individuals in organizations increasingly. Qualifications problems arise when discrepancies exist between individuals' qualifications and the required qualifications for tasks and functions.

Which strategies are used by training managers to solve these problems? On theoretical grounds, three strategies are distinguished. With help of empirical data obtained from fifty-nine training managers the theoretical strategies are further explored.

### INLEIDING

Met de toename van bedrijfsopleidingen zijn steeds meer opleidingskundigen werkzaam binnen arbeidsorganisaties en bij, meer of minder commerciële, opleidingsbureau's. Een belangrijk onderdeel van hun takenpakket is het (doen) ontwikkelen van cursussen, trainingen en andere leersituaties. De aanleidingen daarvoor liggen in problemen bij het uitvoeren van het werk en in veranderingen van arbeid en functies in organisaties. Dit ontwikkelingswerk kan op vele manieren gebeuren. Vaak worden behoeften- of taakanalyses uitgevoerd die de basis vormen voor het opstellen van een opleidingsprofiel waarna het opleidingsprogramma ontwikkeld wordt. Deze taakgerichte ontwikkelingsstrategie is veelvuldig toegepast (Brandsma, Nijhof & Kamphorst, 1990; Mulder ea, 1988). Maar de verscheidenheid aan problemen en arbeid is groot en we vermoeden dat deze gevolgen heeft voor de ontwikkelingsstrategie. Het ontwikkelen van leersituaties voor werk dat bestaat uit kleine taakonderdelen, taakarbeid zoals bij administratief werk of productiewerk, zal anders verlopen dan voor professionele arbeid zoals die van artsen of adviseurs. In dit artikel stellen we de vraag aan de orde welke ontwikkelingsstrategie opleidingskundigen hanteren en of ze deze afstemmen op het soort arbeid.

Eerst maken we plausibel dat er verschillende leerontwikkelingsstrategieën te onderscheiden zijn en dat deze wellicht verband houden met het soort arbeid. Daarna wordt het onderzoek besproken waarin leerontwikkelingsstrategieën van opleiders zijn geoperationaliseerd. Er is gezocht naar de dimensies in de strategieën en er is nagegaan of opleiders in verschillende situaties andere strategieën hanteren.

### THEORETISCHE ACHTERGROND

#### **Kwalificatiestrategieën**

Eén van de voornaamste redenen om in arbeidsorganisaties leersituaties te ontwikkelen is gelegen in aansluitingsproblemen tussen de beschikbare kwalificaties van de medewerkers en de vereiste kwalificaties die nodig zijn om bepaalde taken uit te voeren. Als we kwalificaties opvatten als het geheel van kennis, vaardigheden en houdingen dat iemand in staat stelt ar-

\* Adres: Vakgroep Algemene Pedagogiek, Postbus 9103, 6500 HD Nijmegen.



heidsprestaties te verrichten (Van Hoof & Dronkers, 1980; Brandsma, Nijhof & Kamphorst, 1990), dan zijn beschikbare kwalificaties die kwalificaties die bij functievervullers voorhanden zijn, en vereiste kwalificaties die kwalificaties die nodig zijn om de functie te vervullen. De discrepanties tussen beschikbare en vereiste kwalificaties kunnen onder andere ontstaan doordat de veranderingen in het productieproces of in de organisatiestructuur doorwerken in de taken en functies, de vereiste kwalificaties veranderen. Aan de andere kant kan de situatie op de arbeidsmarkt leiden tot het moeten inzetten van onvoldoende gekwalificeerd personeel.

Om de discrepanties te reduceren kan men maatregelen nemen die betrekking hebben op zowel de beschikbare als op de vereiste kwalificaties. In principe zijn er drie kwalificatiestrategieën mogelijk om die aansluitingsproblemen op te lossen (Van der Krogt, 1990, 1991; Warmerdam & Van den Berg, 1992).

Men kan de beschikbare kwalificaties van de medewerkers aanpassen aan de vereiste kwalificaties. Dit gebeurt bijvoorbeeld door personeel met de vereiste kwalificaties te werven of intern over te plaatsen, maar ook door cursussen en trainingen voor het zittend personeel te verzorgen. Dit noemen we de strategie van de aanpassende kwalificering, de kwalificaties van de medewerkers worden aangepast aan de taken die ze moeten vervullen. Bij de tweede kwalificatiestrategie, die van de aanpassende werkorganisatie, wordt het werk zo veranderd dat de vereiste kwalificaties passen bij de beschikbare kwalificaties van de medewerkers. Het werk wordt aangepast aan de kwaliteiten van de zittende medewerkers. De strategie van de wederzijdse aanpassing tenslotte werkt via veranderingen zowel aan de kant van de vereiste als aan de kant van de beschikbare kwalificaties.

Deze drie kwalificatiestrategieën kunnen het best toegelicht worden aan de hand van de strategieën zoals die in verschillende soorten organisaties voorkomen (Mintzberg, 1989; Van der Krogt, 1990, 1991).

De strategie van de aanpassende kwalificering komen we tegen binnen de machinebureaucratie. Dit is een organisatie, zoals verzekeringsmaatschappijen of massaproductiebedrijven, met zwaar bemande stafafdelingen die het werk ontwerpen, plannen en voorbereiden. Zij ontwerpen o.a. de taken en functies van het uitvoerend personeel. Veel van het uitvoerende werk bestaat uit enkele kleine deeltaken. De wijze waarop de taken dienen te worden uitgevoerd, is van te voren door de voorbereidende staf bepaald. Discrepanties tussen de beschikbare en vereiste kwalificaties worden aangepakt met de strategie van de aanpassende kwalificering, dus werving en/of scholing van personeel. Opleiden bestaat vooral uit instrueren en trainen, tijdens cursussen of in de vorm van trainingen op de werkplek. Aanleiding tot het ontwikkelen van dergelijke leersituaties binnen de machinebureaucratie is vooral wanneer blijkt dat het werk niet goed wordt uitgevoerd of wanneer de functies veranderen. Bij de ontwikkeling van de leersituaties spelen staffunctionarissen, opleidingskundigen en managers in de organisatie een belangrijke rol. De uitvoerende medewerkers, voor wie de leeractiviteiten bedoeld zijn, zijn overwegend passief.

De strategie van de aanpassende werkorganisatie kan het best geïllustreerd worden aan de hand van de professionele bureaucratie. In zo'n type organisatie doen professionals het uitvoerende werk (het primaire proces); denk bijvoorbeeld aan algemene ziekenhuizen en andere instellingen voor gezondheidszorg, grote advocatenkantoren, accountantskantoren. Dit werk is wel complex maar is anderzijds toch ook gestandaardiseerd. De kwalificaties hebben de professionals middels (vaak) jarenlange scholing verworven. Leersituaties voor de professionals worden door instanties in het professionele veld (ontwikkel- en opleidingsinstituten gelieerd aan de professie) ontwikkeld en aangeboden. De wetenschappelijke ontwikkelingen in het vakgebied - de kennis en kunde - worden tijdens cursussen, trainingen, conferenties, symposia, workshops gepresenteerd. De professionals nemen aan deze cursussen en trainingen deel om de ontwikkelingen in hun vak bij te houden. Daarnaast vinden er ook leeractiviteiten op de job plaats; professionals oefenen en evalueren, individueel of in mono-disciplinair verband, hun (standaard)aanpak en toetsen nieuw verworven kennis op bruikbaarheid in de eigen werksituatie. De professionals komen dus als het ware met hun nieuw verworven kennis en vaardigheden de organisatie binnen. Om deze te kunnen toepassen wordt er, door de professie en de professio-



nals, op het management van de organisaties druk uitgeoefend hiervoor de voorwaarden te scheppen, en de taken en functies zo te veranderen dat deze passen bij de kwalificaties die de professionals verworven hebben. Het werk wordt dus aangepast aan de kwalificaties, vandaar de term aanpassende werkorganisatie.

De strategie van de wederzijdse aanpassing kunnen we goed illustreren aan het werk in de adhocatische organisatie. In een adhocratie, zoals in ontwikkelingslaboratoria, in onderzoeksinstituten, in adviesbureau's, wordt veel in projectgroepen gewerkt die, naar gelang de opdracht, ad hoc worden samengesteld. Werken in dergelijke projectgroepen is bijna voortdurend leren. Men wordt voor problemen gesteld die nieuwe uitdagingen bevatten, de standaardkennis en kunde van de professionals is niet direct toepasbaar. Het kan zelfs zo zijn dat het probleem zo ingewikkeld is dat het gheherdefinieerd moet worden, omdat het met de beschikbare kennis en kunde niet op te lossen is. Tijdens het werken aan het probleem wordt duidelijk hoe er verder gegaan kan worden, de taken worden al doende bepaald en verdeeld. Men kan zich het werken voorstellen als een voortdurend wederzijds aanpassen van probleem en kennis/kunde van de projectgroep. Het werk zelf is in de adhocratie dus een belangrijke leersituatie. Daarnaast zullen de (professionele) medewerkers de ontwikkelingen in hun eigen vakdiscipline bijhouden en formele en informele contacten met collega's onderhouden. Het ontwikkelen van netwerken is heel belangrijk voor deze professionals.

In Figuur 1 zijn de verschillen tussen de drie organisaties en de daar (vermoedelijk) gehanteerde strategieën samengevat.

#### Leerontwikkelingsstrategieën

Een centraal onderdeel van een kwalificatiestrategie is het verschaffen van leermogelijkheden aan de medewerkers. Voor de medewerkers moeten mogelijkheden bestaan om de benodigde kwalificaties te verwerven. Organisaties bieden uiteenlopende leersituaties aan, die op verschillende manieren tot stand komen. Dit noemen we leerontwikkelingsstrategieën dat wil zeggen: maatregelen en activiteiten in verband met het ontwikkelen en vormgeven van leersituaties die gericht zijn op het reduceren van discrepanties tussen beschikbare en vereiste kwalificaties.

|                            | MACHINE<br>BUREAUCRATIE  | PROFESSIONELE<br>BUREAUCRATIE   | ADHOCRATIE   |
|----------------------------|--|---|--|
| SOORTEN ARBEID             | <ul style="list-style-type: none"> <li>• taakarbeid</li> <li>• kleine deeltaken</li> <li>• werken volgens voorschriften.</li> </ul>  | <ul style="list-style-type: none"> <li>• prof. bureaucr. arbeid</li> <li>• ingewikkeld werk</li> <li>• individueel uitvoeren volgens (standaard) methodieken geleerd tijdens opleiding.</li> </ul>  | <ul style="list-style-type: none"> <li>• adhocratische arbeid</li> <li>• groepen medewerkers werken aan complexe problemen.</li> <li>• tijdens werken worden taken verhelderd en verdeeld tussen medewerkers.</li> </ul>   |
| KWALIFICATIE-<br>STRATEGIE | <ul style="list-style-type: none"> <li>• aanpassende kwalificering</li> <li>• beschikbare kwalificaties aanpassen aan eisen van werk.</li> <li>• verwerven kwalificaties nav. eisen die taken stellen.</li> <li>• Medewerkers volgen cursussen.</li> </ul> | <ul style="list-style-type: none"> <li>• aanpassende werkorganisatie</li> <li>• werk aanpassen aan beschikbare kwalificaties.</li> <li>• verwerven kwalificaties nav. ontwikkelingen die in beroep plaatsvinden. Professionals volgen cursussen over nieuwe methodieken.</li> </ul> | <ul style="list-style-type: none"> <li>• wederzijdse aanpassing</li> <li>• werk en beschikbare kwalificaties voortdurend aan elkaar aanpassen.</li> <li>• verwerven kwalificaties nav. problemen die zich in werk voordoen. Leren is werk, werk is leren.</li> </ul> |

Fig. 1. Kwalificatiestrategieën voor arbeid in drie typen organisaties.

We spreken over 'leer'ontwikkeling – en niet over opleidingsontwikkeling – omdat het een verscheidenheid aan leersituaties kan betreffen en niet alleen opleidingen, cursussen of trainingen. We denken ook aan leersituaties als intervisie, werkoverleg, training op de werkplek, functioneringsgesprekken. Het gaat om het ontwikkelen van leersituaties en leeractiviteiten die medewerkers kunnen gebruiken om zich te kwalificeren, zowel op de werkplek als daarbuiten, speciaal ontwikkeld voor dit probleem of een standaardcursus enzovoorts.

Leersituaties kunnen voor vele soorten functies en taken ontwikkeld worden. Hier worden deze beperkt tot leerontwikkelingsstrategieën voor uitvoerende functies in verschillende soorten organisaties.

In leerontwikkelingsstrategieën is een viertal aspecten te onderscheiden. Deze zijn (1) de aanleiding voor het ontwikkelen van leersituaties, (2) de werkwijze bij het ontwikkelen van leersituaties, (3) de aard van leersituaties en (4) de organisatie van het proces om het leeraanbod te ontwikkelen. Elk van deze aspecten kan op verschillende wijze ingevuld worden.

Bij de aanleiding speelt datgene waardoor de discrepantie tussen vereiste en beschikbare kwalificaties is ontstaan, een belangrijke rol; het betreft vooral de analyse van de situatie en de achtergronden van het kwalificatievraagstuk. De werkwijze omschrijven we als die activiteiten die leiden tot leeractiviteiten waarmee de geconstateerde aansluitingsproblemen opgeheven kunnen worden. Het gaat hierbij vooral om voorbereidende activiteiten die nodig zijn om te inventariseren en te analyseren waar de aansluitingsproblemen zitten en wat deze precies inhouden. Met de aard van leeractiviteiten worden de kenmerken bedoeld waaraan de leeractiviteiten moeten voldoen om de geconstateerde aansluitingsproblemen te reduceren. Bij dit aspect komen zaken aan de orde als het doel van leeractiviteiten, de leervorm en de leersituatie. De organisatie

#### LEERONTWIKKELINGSSTRATEGIEËN

| ASPECTEN           | TAAKGERICHT  | PROFESSIEGERICHT  | PROBLEEMGERICHT   |
|--------------------|--|---|---|
| <b>AANLEIDING</b>  | Huidige taken niet goed uitgevoerd. Medewerkers beschikken over onvoldoende of onjuiste kwalificaties  | Bijhouden van nieuwe ontwikkelingen binnen het vakgebied.   | Complex en veranderlijk werk vereist voortdurend nieuwe kwalificaties.  |
| <b>WERKWIJZE</b>   | Taak- en functieanalyses bepalen de inhoud van de leeractiviteiten.  | Opleidingsinstututen organiseren in opdracht van beroepsverenigingen leeractiviteiten.  | Medewerkers analyseren zelf problemen en bijbehorende kwalificaties. Dit is tevens een leeractiviteit.  |
| <b>AARD</b>        | Aanleren van taakspecifieke kwalificaties. Doel is taken adequaat uit te kunnen oefenen. Vooral cursussen off the job. Grote groepen lerenden. | Aanleren van breed toepasbare kwalificaties. Zowel cursussen off the job (in grote groepen) als leeractiviteiten on the job (werkbesprekingen, evalueren van handelen). | Doel is vergroten van probleemoplossingsstrategieën. Zowel taakspecifieke als taakgenerieke kwalificaties. Vooral leren op de werkplek, in teams. |
| <b>ORGANISATIE</b> | Stafffunctionarissen en management spelen de hoofdrol. Medewerkers spelen passieve rol.  | De beroepsgroep van de professionals speelt hoofdrol. Binnen organisatie heeft management een bescheiden rol.   | Medewerkers zijn belangrijkste actoren bij ontwikkelen van leeractiviteiten. Zij worden begeleid door coördinator.                                |

Fig. 2. Beschrijving van drie leerontwikkelingsstrategieën naar vier aspecten.

gaat in op de verdeling van taken en verantwoordelijkheden van de actoren (opleidingskundigen, managers, lerenden) bij het ontwikkelen van leeractiviteiten.

We onderscheiden drie leerontwikkelingsstrategieën respectievelijk de taakgerichte, de professioniegerichte en de probleemgerichte leerontwikkelingsstrategie. In Figuur 2 is een korte beschrijving van deze strategieën gegeven.

Welke strategieën zijn nu het meest geschikt? In de contingentietheorie wordt dit afhankelijk gesteld van de situatie (Mintzberg, 1983, 1989; Scheerens, 1989). Vermoedelijk speelt het soort arbeid een belangrijke rol als het leerontwikkelingsstrategieën betreft.

Het ontwikkelen van leeractiviteiten voor adhocistische arbeid zal anders verlopen dan voor taakarbeid in een machine bureaucratische organisatie. Taakarbeid leent zich voor het opstellen van een gedetailleerde taakanalyse. Deze kan als basis dienen voor het in kaart brengen van de te verwerven kwalificaties. Voor een uitvoerende functie in een adhocratie (bijvoorbeeld organisatie-advieswerk) zal deze werkwijze minder geschikt zijn aangezien deze functie niet uit eenduidig te definiëren en af te bakenen taken bestaat. In zo'n geval zal een andere werkwijze, die beter aansluit bij de aard en de kenmerken van de functie, wenselijk zijn.

We veronderstellen dan dat bij taakarbeid de taakgerichte leerontwikkelingsstrategie zal voorkomen, bij professioneel bureaucratische arbeid een professioniegerichte leerontwikkelingsstrategie en dat bij adhocistische arbeid een probleemgerichte leerstrategie aangetroffen wordt.

### ONDERZOEKSVRAGEN

In het algemeen willen we weten welke dimensies er in de leerontwikkelingsstrategieën voorkomen en of de toepassing van de strategie varieert naar gelang het soort arbeid.

Alvorens dit te onderzoeken zijn de volgende beperkingen aangebracht.

Bij het ontwikkelen van leeractiviteiten spelen diverse actoren in en rondom organisaties een rol: managers, uitvoerende medewerkers, personeelsfunctionarissen, opleidingskundigen en andere staffunctionarissen. Opleidingskundigen zijn het meest betrokken bij het ontwikkelen van leersituaties. Om deze reden is gekozen voor de opleidingskundige als centrale figuur in het onderzoek: welke aanpak hanteren opleidingskundigen bij kwalificatieproblemen en in het bijzonder welke dimensies zijn er in de maatregelen van opleidingskundigen te onderscheiden. We noemen dat de leerontwikkelingsdimensies.

Naast de beperking tot de strategieën van opleidingskundigen, is ook gekozen voor een vergelijking tussen twee (in plaats van drie) typen arbeid die naar ons idee ver uiteenlopen: taakarbeid in de bureaucratie en adhocistische arbeid in de adhocratie. We willen nagaan in hoeverre opleidingskundigen kwalificatieproblemen bij verschillende soorten arbeid ook anders aanpakken.

Samenvattend kunnen we de volgende onderzoeksvragen formuleren: zijn er dimensies in de aanpak van kwalificatieproblemen door opleidingskundigen te onderkennen? In aansluiting hierop gaan we na of die aanpak anders is bij de twee typen arbeid.

### ONDERZOEKSOPZET

Het onderzoek is uitgevoerd onder 59 opleidingskundigen. De ene helft kreeg een kwalificatieprobleem bij taakarbeid voorgelegd, de andere helft een soortgelijk probleem bij adhocistische arbeid. Hen is door middel van een schriftelijke vragenlijst gevraagd hoe zij dit probleem zouden aanpakken. In de vragen zijn de drie strategieën en vier aspecten verwerkt, zie ook Figuur 2.

### Subjecten

De opleidingskundigen zijn geworven via een adressenlijst van een beroepsvereniging voor afgestudeerde onderwijskundigen en adressen van opleidingsbureaus. Verder is gebruik ge-

maakt van de sneeuwbalmethode: opleidingskundigen benaderen en hen vragen naar adressen van collega's. Uiteindelijk is een bestand ontstaan van 80 opleidingskundigen aan wie de door ons gemaakte vragenlijst werd toegestuurd. Er werden 60 vragenlijsten teruggestuurd waarvan er 59 bruikbaar bleken. De hierna gepresenteerde onderzoeksresultaten zijn daarmee gebaseerd op een response van 74%.

Van deze 59 opleidingskundigen bleken 35 personen binnen de eigen arbeidsorganisatie te werken en 24 op een zelfstandig opleidingsbureau. Dit laatste betekent dat de opleidingskundige bezig is met opleiden en leren voor medewerkers uit andere organisaties. Het grootste deel (38) heeft een wetenschappelijke opleiding genoten. Daarnaast hebben 19 opleidingskundigen een HBO-opleiding en 2 opleidingskundigen een MBO-opleiding gevolgd.

### Procedure en onderzoeksinstrumenten

Er zijn twee verschillende casebeschrijvingen gemaakt waarin een kwalificatieprobleem werd beschreven.

In de ene case-beschrijving staat een verzekeringsmaatschappij centraal. Het primaire proces in deze maatschappij kenmerkt zich door een organisatie van de arbeid die vergelijkbaar is met die in een machinebureaucratie: het werk van de verzekeringsagenten bestaat uit kleine deeltaken die op een gestandaardiseerde manier moeten worden uitgevoerd, ze hebben weinig autonomie en hun functie is als weinig complex te omschrijven. Het management signaleert dat de verzekeringsagenten bij moeilijke vragen (die afwijken van het standaardpakket) van cliënten steeds vaker een beroep doen op stafmedewerkers. Dit heeft langere wachttijden tot gevolg en leidt dus tot een vermindering van de kwaliteit van de dienstverlening.

De andere case-beschrijving gaat over een reclame-adviesbureau. De organisatie van de arbeid in dit bureau heeft overeenkomsten met de organisatie van de arbeid in een adhocistische organisatie: de reclameontwerpers werken in teamverband aan complexe opdrachten, tijdens het werken aan deze opdrachten wordt in onderling overleg bepaald wie welke taken uitvoert en in welke mate deze taken verdeeld worden. In het reclamebureau wordt gesignaleerd dat reclames en advertenties vaker een voor de klant specifieke vorm moeten krijgen en dat de opdrachten steeds diverser van aard worden. Het wordt steeds moeilijker om vragen van cliënten om te zetten in ontwerpen. Door de vragen uit de schriftelijke vragenlijst te beantwoorden, kunnen de respondenten aangeven hoe ze de genoemde kwalificatieproblemen aan gaan pakken.

Zowel de vragenlijst als de case-beschrijvingen zijn aan een kritisch vooronderzoek onderworpen alvorens ze aan de gehele steekproef werden voorgelegd. Eerst zijn proefinterviews gehouden met vier opleidingskundigen. Als resultaat daarvan zijn de case-beschrijvingen en een aantal items uit de vragenlijst bijgesteld. In de daarop volgende pilot-studie is het gehele materiaal voorgelegd aan drie opleidingskundigen met het verzoek het geheel kritisch te evalueren en waar nodig wijzigingen voor te stellen. Dit resulteerde in een definitieve vragenlijst met bijbehorende casebeschrijvingen. De vragenlijst bestaat uit 107 items met 7-punts Likert antwoord schalen variërend van 1 (=klopt helemaal niet) tot en met 7 (=klopt helemaal wel). De taakgerichte leerontwikkelingsstrategie bestaat uit 46 items (2 voor aanleiding, 21 voor werkwijze, 16 voor aard en 7 voor organisatie), de professiegerichte uit 24 (respectievelijk 1, 9, 12 en 2) en de probleemgerichte uit 37 items (respectievelijk 2, 13, 19 en 3).

De 80 opleidingskundigen zijn aselekt verdeeld in twee groepen. De ene groep kreeg het kwalificatieprobleem met taakarbeid voorgelegd, de andere groep het probleem met de adhocistische arbeid. Beide groepen moesten dezelfde vragenlijst invullen. Over de case-beschrijving van de verzekeringsmaatschappij zijn van 31 opleidingskundigen de antwoorden verwerkt, over de case-beschrijving van het reclame-adviesbureau 28. Van alle opleidingskundigen werden ook enige achtergrondgegevens verzameld (genoten opleidingen en dergelijke). Geen significante verschillen werden gevonden tussen beide groepen op diverse achtergrondkenmerken.

Voor een meer uitvoerige beschrijving van de case-beschrijvingen, de vragenlijst en de vergelijking op achtergrondkenmerken, zie Keurentjes & Ter Woerds (1992).



### Analyse opzet

De antwoorden van alle respondenten op de 107 items van de vragenlijst vormen de input voor de analyses. Met behulp van iteratieve principale factor analyse wordt geprobeerd factoren of dimensies te realiseren die een weerspiegeling zijn van leerontwikkelingsdimensies. Er wordt niet bij voorbaat aangenomen dat de dimensies onafhankelijke concepten zijn. Daarom zal gebruik worden gemaakt van oblique rotaties (OBLIMIN) voor interpretatie van de onderliggende factorstructuur. Items die niet in de factorstructuur thuis horen zullen worden verwijderd. In eerste instantie worden de minst passende items verwijderd maar als eis stellen we dat een item in de uiteindelijke patroonmatrix minstens een gewicht moet hebben van .40 op de factor waarvoor het item geformuleerd is en een gewicht lager dan .20 op de overige factoren. Items die verwijderd zijn zullen in de loop van het analyseproces weer in de analyse worden opgenomen om te kijken of ze nu wel in de factorstructuur passen. Een dergelijke werkwijze leidt tot duidelijk van elkaar onderscheidbare factoren en daarmee tot leerontwikkelingsdimensies die ten opzichte van elkaar afgebakend zijn. Dit komt de validiteit van het meetinstrument ten goede. Nadat de factorstructuur is vastgesteld worden betrouwbaarheden vastgesteld in termen van Cronbach's interne consistentie maat alpha. Items die de betrouwbaarheid in behoorlijke mate drukken zullen zonnodig verwijderd worden.

## RESULTATEN

### Vier leerontwikkelingsdimensies

Van de 107 geformuleerde items zijn tenslotte 31 items na factor analyses overgebleven.

De overgebleven items zijn vooral afkomstig uit de aspecten werkwijze en aard van de strategie: 27 van de 31. Eén item blijkt nog het aspect aanleiding te vertegenwoordigen, terwijl 3 items betrekking hebben op de organisatie. Dat van de aspecten aanleiding en organisatie zo weinig items zijn overgebleven heeft waarschijnlijk te maken met het gering aantal items dat voor deze aspecten al vooraf was geformuleerd. Er zijn vier factoren geconstrueerd waarvan het resulterend factorpatroon vermeld staat in Tabel 1.

De eerste kolom in Tabel 1 geeft de originele nummering weer van de items. Kolom 2 geeft aan van welk aspect de items afkomstig zijn: al=aanleiding, w=werkwijze, ar=aard en or=organisatie). Kolom 3 geeft een verkorte omschrijving van de inhoud van de items, terwijl de kolommen 4, 5, 6 en 7 de gewichten weergeven van de items op elk van de vier factoren. In de laatste regel van de tabel zijn de betrouwbaarheden vermeld.

De vier factoren verklaren in totaal 42.5% van de variantie in de variabelen, wat voor analyses op itemniveau een hoog percentage genoemd kan worden. De factoren blijken onderlinge correlaties te laten zien van -.07 tot .12 wat betekent dat ze als onafhankelijke factoren kunnen worden beschouwd. Een orthogonaal uitgevoerde rotatie (VARIMAX) laat dan ook qua factorstructuur een identiek resultaat zien.

De betrouwbaarheden zijn voor drie van de vier factoren goed te noemen. De eerste drie factoren hebben respectievelijk alpha's van .81, .87 en .81, terwijl de laatste factor een alpha heeft van .69. Er zijn geen items gevonden die een negatieve invloed hebben op de betrouwbaarheid vandaar dat op basis van de betrouwbaarheidsanalyses geen items verwijderd werden.

De factoren verwijzen ons inziens naar vier leerontwikkelingsdimensies. We duiden deze aan met:

- 1 taakanalytische oriëntatie. (F1).
- 2 professionele leeraanbod oriëntatie. (F2).
- 3 leer-werk oriëntatie. (F3).
- 4 werkaanpassing oriëntatie. (F4).

De taakanalytische oriëntatie geeft aan in welke mate een opleidingskundige een analyse maakt van de taken die moeten worden uitgevoerd. Opleiders met een hoge score op deze dimensie leggen de nadruk op het voortraject, de analyses die nodig zijn alvorens leersituaties te ontwikkelen.



Tabel 1. Factorpatroon van de vier-factoroplossing.

| nr                           | as | verkorte item-omschrijving  | F1   | F2   | F3   | F4   |
|------------------------------|----|---|------|------|------|------|
| 8                            | w  | nieuwe taken moeten in kaart worden gebracht                            | .75  | .02  | -.03 | .06  |
| 7                            | w  | huidige taken moeten geanalyseerd worden                                | .75  | .13  | -.12 | -.02 |
| 9                            | w  | discrepancie tussen huidige en nieuwe taken moet bepaald worden         | .74  | -.11 | .07  | .02  |
| 27                           | w  | taken van werknemers duidelijk beschrijven                              | .73  | -.01 | .19  | .04  |
| 14                           | w  | taakanalyse is grondslag voor het ontwerpen van de leeractiviteit       | .70  | .05  | -.02 | -.07 |
| 38                           | w  | bestaande taakomschrijvingen worden geraadpleegd                        | .63  | .12  | -.03 | -.14 |
| 64                           | ar | eerst kwalificaties verwerven voordat nieuwe taken uitgevoerd worden    | .50  | -.05 | -.05 | -.02 |
| 98                           | or | leeractiviteit in overleg met het management grondig plannen            | .41  | -.02 | -.04 | .04  |
| 81                           | a  | kwalificaties verwerven via cursussen bij eigen beroepsvereniging       | -.11 | .86  | -.10 | .00  |
| 47                           | w  | een beroepsvereniging consulteren                                       | .08  | .84  | -.01 | .14  |
| 36                           | w  | samenwerken met een beroepsvereniging                                   | .03  | .73  | .09  | -.01 |
| 43                           | w  | overzicht van leeractiviteiten van de beroepsvereniging raadplegen      | .12  | .73  | .01  | .20  |
| 65                           | ar | beginnen met het volgen van een cursus bij de eigen beroepsvereniging   | .04  | .68  | -.11 | .02  |
| 22                           | w  | de deskundigheid van de beroepsvereniging is belangrijk                 | .06  | .61  | -.02 | .08  |
| 20                           | w  | het bestaan van een beroepsvereniging speelt een belangrijke rol        | -.13 | .61  | -.12 | .09  |
| 95                           | ar | samen met collega's uit dezelfde branche kwalificaties verwerven        | .03  | .44  | .02  | .06  |
| 2                            | al | ontwikkelingen binnen eigen vakgebied is aanleiding                     | .09  | .42  | .09  | -.07 |
| 57                           | ar | professionele kennis is belangrijker dan specifieke kwalificaties       | -.11 | .40  | .03  | -.14 |
| 50                           | ar | leeractiviteit moet samenhangen met complexe vragen van het werk        | .03  | -.07 | .75  | .03  |
| 48                           | w  | het complexe werk van werknemers vormt uitgangspunt                     | -.19 | -.09 | .73  | -.03 |
| 101                          | or | werknemers spelen een actieve rol bij ontwikkeling van leeractiviteiten | -.09 | -.12 | .62  | .15  |
| 19                           | w  | de mate van complexiteit van het werk is van belang                     | .00  | .06  | .58  | -.08 |
| 89                           | ar | werknemers moeten in kleine groepjes met elkaar reflecteren             | -.12 | -.11 | .54  | -.03 |
| 84                           | ar | de feitelijke werksituatie is belangrijk voor kwalificatieverwerving    | .04  | .17  | .49  | .02  |
| 21                           | w  | het probleemoplossend vermogen moet worden vergroot                     | -.01 | .20  | .44  | -.20 |
| 100                          | or | leeractiviteit is meer procesgeleiden                                   | .17  | -.07 | .44  | .09  |
| 85                           | ar | kwalificaties verwerven via aanpassing van werkopdrachten               | -.05 | -.02 | .12  | .77  |
| 25                           | w  | de mate waarin werknemers opdrachten aanpassen aan hun kwalificaties    | -.02 | .14  | -.09 | .62  |
| 54                           | ar | leeractiviteit betekent aanpassen van werkopdrachten aan kwalificaties  | .02  | .00  | -.11 | .58  |
| 11                           | w  | kwaliteiten analyseren en opdrachten daarop afstemmen                   | -.07 | .03  | -.01 | .57  |
| 90                           | ar | kwalificaties verwerven door aanpassing van het werk aan problemen      | .04  | .04  | .10  | .49  |
| Cronbach's alpha coëfficiënt |    |   | .81  | .87  | .81  | .69  |

De professionele leeraanbod oriëntatie benadrukt de betekenis van de beroepsvereniging en het leeraanbod dat vanuit die kringen wordt aangeboden. De inhoud van de na- en bijscholing is vooral een aangelegenheid van de professie, zo vinden opleidingskundigen met een hoge score op deze dimensie.

De leer-werk oriëntatie brengt de verbondenheid van leren en werken tot uitdrukking. Het kwalificatievraagstuk zou vooral dicht bij de werkplek in samenwerking met de werkenden zelf geanalyseerd en aangepakt moeten worden. Deze aanpak wordt mede ingegeven door de complexiteit van het werk.

De vierde oriëntatie verwijst naar de mate waarin opleidingskundigen mogelijkheden zien om door het werk aan te passen de kwalificatieproblematiek op te lossen. Er zijn diverse elementen van de wederzijdse aanpassing aanwezig.

Tabel 2. Resultaten van de verschiltoetsen.

| leerontwikkelings-<br>dimensie         |              | taak-<br>arbeid | adhocratische<br>arbeid | df | t     | p    |
|--|--------------|-----------------|-------------------------|----|-------|------|
| taakanalytische<br>oriëntatie          | gem.<br>s.d. | 5.40<br>.91     | 4.62<br>1.34            | 47 | 2.58  | .013 |
| professionele<br>leeraanbod oriëntatie | gem.<br>s.d. | 3.34<br>1.09    | 3.98<br>.82             | 56 | -2.54 | .013 |
| leer-werk<br>oriëntatie                | gem.<br>s.d. | 5.54<br>.78     | 5.51<br>.77             | 56 | .16   | n.s  |
| werkaanpassing<br>oriëntatie           | gem.<br>s.d. | 3.17<br>1.12    | 3.41<br>1.15            | 56 | -.78  | n.s  |

n.s = niet significant

### Verschillen in strategieën bij de twee soorten arbeid

Met behulp van t-toetsen voor onafhankelijke steekproeven is nagegaan of op de vier leerontwikkelingsdimensies opleidingskundigen bij taakarbeid anders te werk gaan dan bij adhoc-ratische arbeid. De resultaten staan in Tabel 2.

Bij taakarbeid hebben opleidingskundigen significant een andere taakanalytische oriëntatie dan bij adhoc-ratische arbeid ( $t=2.58$ ,  $p=.013$ ). De cijfers wijzen erop dat de taakanalytische oriëntatie hoger scoort bij taakarbeid dan bij adhoc-ratische arbeid. Voor de professionele leeraanbod oriëntatie vinden we eveneens een significant verschil ( $t=-2.54$ ,  $p=.013$ ) maar nu in omgekeerde richting: bij taakarbeid scoort men gemiddeld lager dan bij adhoc-ratische arbeid. De leer-werk en de werkaanpassingsoriëntatie laten geen significante verschillen zien tussen beide typen van arbeid ( $t=.16$  respectievelijk  $-.78$ , beide dus niet significant).

### Conclusies en perspectieven

De aanpak van kwalificatieproblemen door opleidingskundigen kan langs vier leerontwikkelingsdimensies beschreven worden. Bovendien blijken opleidingskundigen op uiteenlopende manieren met die kwalificatieproblemen om te gaan en daarbij enigermate rekening te houden met de situatie waarin zich het probleem voordoet. De aanpak van kwalificeringsproblemen bij taakarbeid verloopt anders dan bij adhoc-ratische arbeid.

De vier dimensies in de aanpak van kwalificatieproblemen, die het resultaat zijn van de analyses, zijn als volgt te omschrijven:

- de taakanalytische oriëntatie, dat wil zeggen de nadruk die opleidingskundigen leggen op de (taak)analyses in het voortraject. Deze oriëntatie, waarin de uit te voeren taken een belangrijk referentiepunt voor de aanpak van het probleem vormt, is bij de onderzochte opleidingskundigen in sterke mate aanwezig. In de opleidingskundige literatuur en technolo-gie is deze opvatting ook het verst ontwikkeld.
- de professionele leeraanbod oriëntatie, dat wil zeggen de betekenis die opleidingskundigen toekennen aan de professie voor de na- en bijscholing. Het kwalificatieprobleem wordt gerelateerd aan de professie van de uitvoerenden. Daar komen de inhoudelijke impulsen voor de scholing vandaan. Deze oriëntatie is duidelijk minder aanwezig bij de opleidingskundigen.
- de leer-werk oriëntatie, dat wil zeggen de betekenis van de integratie van werken en leren. Ook hier speelt het werk een belangrijke rol maar op een geheel andere manier dan bij de taakanalytische oriëntatie. Het werk is zo complex dat het minder gemakkelijk is te analyseren en er geen standaard leersituaties voor te maken zijn. Het leren en werken wordt meer

met elkaar verbonden. Hierin is het leren op de werkplek onder andere door zelf de problemen te analyseren, herkenbaar. Dit impliceert ook een veel actievere rol van de werknemers dan bij de taakanalytische oriëntatie.

- d. de werkaanpassing oriëntatie, dat wil zeggen de mogelijkheden die opleidingskundigen zien om door middel van het aanpassen van het werk de kwalificatieproblemen aan te pakken en leeractiviteiten te ondernemen. In de taakanalytische en de leer-werk oriëntatie wordt het werk als een gegeven beschouwd. Bij de werkaanpassing wordt het kwalificatieprobleem aan twee kanten aangepakt, zowel de beschikbare als de vereiste kwalificaties staan ter discussie. Door het werk te veranderen wordt er ook geleerd.

In de eerste dimensie is de meest gangbare leerontwikkelingsstrategie, de taak- of functiegerichte, duidelijk herkenbaar. De drie andere oriëntaties laten ook de nieuwe elementen zien namelijk dat veranderingen in het werk mogelijkheden bieden voor de aanpak van kwalificatieproblemen en tevens dat de eigen inbreng van de (lerende) werknemers bij de aanpak van dergelijke problemen van belang is. De verbinding van leer- en ontwikkelingsactiviteiten met de analyse en verandering van de werksituatie blijkt niet alleen in theorie maar ook empirisch herkenbaar in de aanpak van deze opleidingskundigen.

De tweede onderzoeksvraag heeft betrekking op het hanteren van verschillende leerontwikkelingsstrategieën door opleidingskundigen bij verschillende soorten arbeid. Ons vermoeden was dat er weinig of geen verschillen zouden bestaan en dat de taakgerichte leerontwikkelingsstrategie het sterkst vertegenwoordigd zou zijn. Deze verwachting was gebaseerd op de constatering dat in de onderwijskundige literatuur de taakgerichte leerontwikkelingsstrategie sterk vertegenwoordigd is. In deze literatuur, die ook veel in opleidingen voor opleidingskundigen wordt gebruikt, zijn de beide andere strategieën nog niet sterk vertegenwoordigd (zie bijvoorbeeld Robinson & Robinson, 1989; Kessels & Smit, 1989; Thijssen, 1988; Mulder e.a., 1988, Romiszowski, 1989). Uit Tabel 2 blijkt echter dat opleidingskundigen bij beide soorten arbeid verschillend scoren op de taakanalytische oriëntatie en op de professionele leeraanbod oriëntatie. Dit kan het volgende betekenen: opleidingskundigen laten zich bij het ontwikkelen en vormgeven van een leeraanbod niet alleen leiden door de dominante taakgerichte leerontwikkelingsstrategie. Blijkbaar bestaan er verschillende opvattingen over hoe leeractiviteiten ontwikkeld moeten worden; in die opvattingen speelt de (contingente) gedachte dat de situatie mede de aanpak bepaalt een belangrijker rol dan we eerst aannamen.

In dit verband is het goed om ons af te vragen wat we hier gemeten hebben: de oriëntaties en strategieën zoals opleidingskundigen die in hun hoofd hebben of hoe zij in de praktijk te werk gaan? Dit is een vraag naar de "espoused theory", de wenselijke manier van werken en de "theory-in-use" ofwel de feitelijke manier van werken (Argyris & Schön, 1978). Mensen zijn zich niet altijd bewust van hun "theory-in-use" en de "espoused theory". Tijdens proefinterviews bleek dat opleidingskundigen moeilijk het waarom van hun handelen konden beargmenteren. Ze hadden moeite hun praktijktheorie, het geheel aan veronderstellingen dat aan het (keuze-)gedrag van functionarissen ten grondslag ligt (Van Gageldonk & Meyer, 1988, p. 9) te expliciteren. Verder onderzoek hiernaar kan de professionalisering van opleidingskundigen ten goede komen. Door inzicht in de eigen praktijktheorieën kan men de veronderstellingen die aan het handelen ten grondslag liggen ter discussie stellen. Deze vorm van leren is ons inziens van belang omdat opleidingskundigen hiermee inzicht krijgen in hun eigen leerontwikkelingsstrategie maar ook in andere. Het bewust variëren van de strategie naar gelang de situatie, arbeid en organisatie, wordt daarmee bevorderd.

Voorgaande overwegingen en kanttekeningen geven al aan dat dit onderzoek bedoeld is als aanzet voor verdere, empirisch onderbouwde, studies en theorievorming over de samenhang tussen kwalificering van medewerkers en arbeid in verschillende soorten organisaties. Onderzoek en theorievorming op dit terrein staat nog in de kinderschoenen en onze resultaten beschouwen we, door de beperkte omvang van het onderzoek en de gebruikte meetinstrumenten, als empirisch gefundeerde werkhypothesen voor verdere studies.

## LITERATUUR

- Argyris, C., & Schön, D.A. (1978). *Organizational learning: A theory of action perspective*. Massachusetts, California: Addison-Wesley Publishing.
- Brandsma, T.F., Nijhof, W.J. & Kamphorst, J.C. (1990). *Kwalificatie en curriculum*. Amsterdam/Lisse: Swets & Zeitlinger.
- Gageldonk, A. van, & Meyer, R.F. (1988). *Praktijktheorieën van het management in het middelbaar beroepsonderwijs: Reconstructie en analyse van hoofdlijnen*. Leiden: Leids Instituut voor Sociaal Beleidsonderzoek (LISBON).
- Hoof, J.J. van & Dronkers, J. (1980). *Onderwijs en arbeidsmarkt: een verkenning van relatie tussen onderwijs, arbeidsmarkt en arbeidssysteem*. Deventer: Van Loghum Slaterus.
- Kessels, J.W.M., & Smit, C.A. (1989). *Opleidingskunde: Een bedrijfsgerichte benadering van leerprocessen*. Deventer: Kluwer Bedrijfswetenschappen.
- Keurentjes, J.G.M.M., & Woerds, I.E.C. ter. (1992). *Opleidingskundigen en de ontwikkeling van leeractiviteiten. Een onderzoek naar de leerstrategieën die opleidingskundigen toepassen bij een verschillende organisatie van de arbeid*. Nijmegen: Katholieke Universiteit Nijmegen, doctoraalscriptie Vakgroep Algemene Pedagogiek.
- Kraayvanger, G., & Onna B. van. (1985). *Arbeid en leren: Bijdragen tot de volwasseneneducatie*. Baarn: Nelissen.
- Krogt, F.J. van der. (1990). De leerfunctie in verschillende organisaties. *Opleiding & Ontwikkeling*, 11, 17-24.
- Krogt, F.J. van der (1991a). Aansluitingsstrategieën en opleidingsactiviteiten in verschillende organisatie-typen. *Tijdschrift voor Arbeidsvraagstukken*, 4, 3-15.
- Krogt, F.J. van der. (1991b). Autonomie en leren van professionals. *Opleiding & Ontwikkeling*, 9, 15-18.
- Mintzberg, H. (1983). *Structure in Fives: Designing effective organizations*. New Jersey: Prentice-Hall International Editions.
- Mintzberg, H. (1989). *Mintzberg on management*. New York: Free Press.
- Mulder, M., Akkerman, J.S., & Bentvelsen, N. (1988). *Bedrijfsopleidingen in Nederland*. Den Haag: Instituut voor Onderzoek van het Onderwijs.
- Robinson, D.G. & Robinson, J.C. (1989). *Training for impact*. San Francisco: Jossey Bass.
- Romiszowsky, A.J. (1989). Ontwikkelen van opleidingen. Instructie: De integratieve benadering. *Handboek Opleiders in Organisaties*. Deventer: Kluwer Bedrijfswetenschappen.
- Scheerens, J. (1989). Onderzoek naar schoolorganisatorische effectiviteit en de contingentiebenadering. In: Scheerens, J. & Verhoeven, J.C. (red.) *Schoolorganisatie, beleid en onderwijskwaliteit*. Lisse: Swets & Zeitlinger.
- Thijssen, J.G.L. (1988). *Bedrijfsopleidingen als werkterrein*. 's-Gravenhage: VUGA uitgeverij.
- Warmerdam, J., & Berg, J. van den. (1986). *Opleidingsactiviteiten in arbeidsorganisaties*. Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen.
- Warmerdam, J. & Berg, J. van den (1992). *Scholing van werknemers in veranderende organisaties*. Den Haag/Nijmegen: Ministerie van Sociale Zaken/ITS.

Manuscript ontvangen 14-5-1993

Definitieve versie ontvangen 8-9-1993

# Relevante Variabelen bij het Doorverwijzen na de Lagere School; een Structureel Model

H. Blok

*Research Instituut voor de Toegepaste Psychologie<sup>1</sup>*

W.E. Saris

*Vrije Universiteit, Amsterdam<sup>2</sup>*

## *The assignment of elementary school pupils to secondary school types; a linear structural model*

Aim of the study was to determine the effect of the sex of the pupil, the SES of the parents and the scores on a well-known Dutch school achievement test (Cito-schooltoets) on the assignment of elementary school pupils to secondary school types. To that end the relations were explored between sex of the pupil, SES of the parents, teacher predictions of the test score, the obtained test scores, first teacher advice as to the most appropriate school type (given before the test), second teacher advice (given after the test), application of the parents and the final decisions of the assignment committees in a sample of 514 sixth grade pupils. After introduction of a non-measured variable, school achievement (throughout the year), our linear structural model fits in with the data and is easily interpretable. Our conclusions are that sex and SES have influence upon school achievement only, with advantage for girls and for pupils with higher SES. Neither the teachers, nor the parents nor the assignment committees seem to take into account sex or SES of the pupils. As to the well-known school achievement test, the obtained test score has a slight influence upon the assignment proces via the second teacher advice. Somewhat surprisingly the predicted test score seemed to have a greater influence than the obtained test score. After all, the factor which most influenced the decisions of the various persons involved (teachers, parents, assignment committees) is the unmeasured variable school achievement throughout the year.

## 1. Inleiding

De LO-VO overgang vormt al vele jaren een dankbaar terrein van onderzoek. Eén van de grondleggers van de Nederlandse onderwijsresearch, Révész, is de eerste geweest die het probleem van de aansluiting serieus aanvatte. Révész constateerde in 1926 dat de schoolprestaties van H.B.S.-leerlingen bedroevend slecht waren en zocht de oorzaak hiervan in de selectie, die naar zijn oordeel niet streng genoeg was. Ook Kohnstamm heeft zich, intensiever, beziggehouden met het aansluitingsprobleem. Samen met Van Veen bracht hij in 1928 aan de wethouder voor het onderwijs te Amsterdam een rapport uit over 'De aansluiting tusschen Lager en Middelbaar (Gymnasiaal) Onderwijs'. Zoals Deen (1969) beschrijft, heeft dit thema een belangrijke rol gespeeld in het onderzoek van het Nutsseminarium voor Pedagogiek aan de Universiteit van Amsterdam. De ontwikkeling van de

<sup>1</sup> R.I.T.P., Prinsengracht 303, Amsterdam

<sup>2</sup> Afd. Methoden en Technieken, Vrije Universiteit, De Boelelaan 1115, Amsterdam



'Amsterdamse Schooltoetsen' samen met het RITP in 1966 vormde een waardige afsluiting van het werk dat het Nutsseminarium op dit terrein verricht heeft. Het aansluitingsprobleem werd door de schooltoets evenwel niet uit de wereld geholpen en het onderzoek werd voortgezet. Van Heek (1968), Boon van Ostade (1972), Oosterbaan (1973), Peschar (1975), Blok (1978), Dronkers (1978) hielden zich bezig met de invloed van sociaal milieu op de doorstroming en, wijder soms, de schoolloopbaan. Bos & Warries (1971) verrichtten een evaluatie van de Amsterdamse Schooltoetsen. Sandbergen et al. (1972), Lutje Spelberg & Rotteveel (1978) schonken aandacht aan de voorspellende waarde van intelligentie- en schoolvorderingentoetsen, die gebruikt worden bij de LO-VO overgang. Onlangs werden aan deze onderzoeken nog enkele toegevoegd. Het CITO rapporteert over een onderzoek naar het gebruik van en de waardering voor eindprocedures en -methoden bij de LO-VO overgang (CITO, 1978). Hoogstraten & Mellenbergh (1978) en Groeneboom et al. (1978) trachten vast te stellen welke variabelen invloed hebben op en van belang zijn voor het onderwijzersadvies. Blok en Wesdorp (1979) inventariseren welke, oorspronkelijk aan de schooltoets toegedachte functies momenteel vervuld worden door de CITO-toets. Het gaat ernaar uitzien dat deze stroom van onderzoeken nog doorgaat tot het moment dat lager en voortgezet onderwijs zo hecht aaneen gesmeed zijn, dat selectie en determinatie niet meer nodig zijn.

Er zijn drie vraagstellingen met het oog waarop wij ons model ontwikkelden. De eerste twee vragen zijn die naar invloeden van sexe en sociaal milieu op de doorstroming. De derde vraag betreft de invloed van een schooltoets op beslissingen aangaande de schoolkeuze. Op het belang van deze vragen willen wij kort ingaan.

Dat meisjes in het onderwijs een achterstand oplopen is bekend en al vaak aangetoond. Zij sluiten hun onderwijsloopbaan gemiddeld eerder en op een lager niveau af dan jongens (Meijer, 1977). Een antwoord op de vraag naar de oorzaken van deze achterstand kan nog niet gegeven worden, want empirisch onderzoek op dit terrein komt nu pas van de grond. Mogelijke oorzaken van deze achterstand kunnen echter wel aangegeven worden. Meisjes blijken over het algemeen gemiddeld iets lagere scores dan jongens te leveren op schooltoetsen (Kohnstamm, 1973). Het is mogelijk dat meisjes als gevolg daarvan bij de overgang van het ene schooltype naar het volgende minder mogelijkheden hebben. Het kan ook zijn dat de achterstand van meisjes niet zozeer een gevolg is van hun net iets lagere prestaties, maar dat ouders of leerkrachten een succesvolle schoolloopbaan voor meisjes minder belangrijk achten. Wij zullen nagaan of meisjes tijdens de LO-VO overgang achtergesteld worden, en zo ja, welke rol hierbij niet alleen schoolprestaties spelen, maar ook de leerkrachten en de ouders.

Niet alleen van meisjes, ook van kinderen uit lagere sociale milieus is de onderwijsachterstand vaak vastgesteld. Hoewel op dit terrein al meer onderzoek verricht is, zijn er ook hier verschillende gezichtspunten over de bron van deze achterstand. Een strijdvraag is bijvoorbeeld in hoeverre het lager onderwijs tijdens de aansluiting bijdraagt aan de ongelijke verdeling van onderwijs. De belangrijkste conclusies uit empirisch onderzoek (Van Heek, 1968; Van Kemenade en Kropman, 1972; Kropman en Collaris, 1974; Dronkers, 1978) hebben betrekking op de generatie die in 1965 het lager onderwijs verliet. Voor deze generatie werd geconstateerd dat onderwijzers in hun adviezen bijdragen aan de maatschappelijke ongelijkheid. Blok's conclusie (Blok, 1978) dat leerkrachten in hun adviezen leerlingen uit lagere milieus niet benadelen, is gebaseerd op gegevens uit het Amsterdamse onderwijs in 1977. Recente gegevens met een landelijk en representatief karakter ontbreken. Wij zullen nagaan of leerlingen uit de lagere milieus bij de overgang

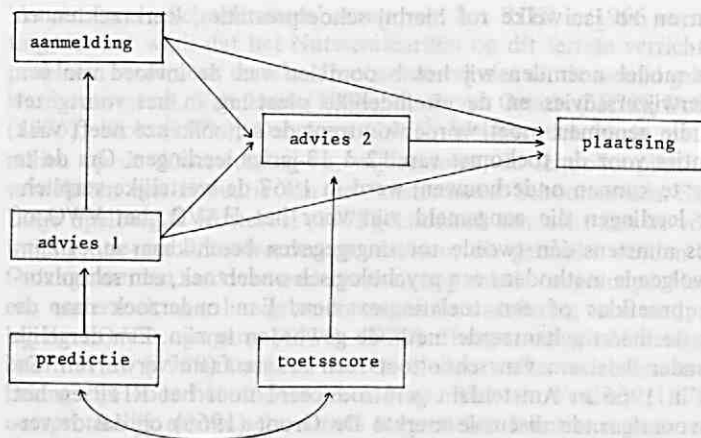
een achterstand oplopen en zo ja, welke rol hierbij schoolprestaties, leerkrachten en ouders spelen.

Als derde doel van ons model noemden wij het beoordelen van de invloed van een schooltoets op het onderwijzersadvies en de uiteindelijke plaatsing in het voortgezet onderwijs. De beslissing die genomen moet worden omtrent de schoolkeuze heeft vaak verstrekkende consequenties voor de toekomst van 12 à 13-jarige leerlingen. Om de te nemen beslissingen beter te kunnen onderbouwen, werd in 1967 de wettelijke verplichting ingevoerd dat over leerlingen die aangemeld zijn voor het HAVO, het VWO of HAVO-VWO combinaties minstens één tweede toetsingsgegeven beschikbaar moet zijn. De keus is hierbij uit de volgende methoden: een psychologisch onderzoek, een schoolvorderingenonderzoek, een proefklas of een toelatingsexamen. Een onderzoek naar de schoolvorderingen blijkt de meest gehanteerde methode geworden te zijn. Een dergelijk onderzoek heeft zich onder de naam van schooltoets een zekere faam verworven. De eerste schooltoets werd in 1966 in Amsterdam geïntroduceerd door het RITP en het Nutsseminarium. In de voorafgaande discussie merkte De Groot (1966) op dat de verwachte verbetering door de schooltoets van de selectie voor het VO niet groot is. Al vormde de selectieve functie de aanleiding tot de introductie van de schooltoets, toch hechtten de oorspronkelijke ontwerpers veel meer belang aan de evaluatieve functies. Het blijkt nu dat schoolhoofden en leerkrachten desondanks de schooltoets veel vaker een selectieve functie toekennen dan een evaluatieve (CITO, 1978). Hoe groot de selectieve functie is, is evenwel onbekend.

## 2. Modelspecificatie

De keuze welke schooltoets in het onderzoek opgenomen zou worden, was niet moeilijk. Er zijn in ons land slechts enkele omvangrijke eindtoetsen voor het lager onderwijs in gebruik. Wij kozen voor de grootste, de CITO-toets. Deze toets wordt door circa 45% van alle doorstromende leerlingen gemaakt. Het CITO doet enkele aanbevelingen aangaande de procedures, die gevolgd worden bij de LO-VO overgang. Wanneer deze aanbevelingen worden opgevolgd, leidt dit tot de volgende gang van zaken. In januari brengen de schoolhoofden voor iedere leerling een eerste advies uit. Dit advies, waarin het schoolhoofd aangeeft welk type voortgezet onderwijs het meest geschikt is voor de leerling, wordt aan het CITO opgegeven. Tevens wordt dit eerste advies ingebracht in het overleg met de ouders, dat in de eerste maanden van het jaar gevoerd wordt. De ouders schrijven hun kind in op de school van hun keuze. Bij de inschrijving dienen ze, naast enkele personalia, ook het onderwijzersadvies op te geven. In de maanden februari of maart wordt meestal de toets afgenomen. Wanneer de toetsscores een maand later aan de school worden meegedeeld, is de reguliere inschrijvingstermijn voor het voortgezet onderwijs al gesloten. Via de school wordt de toetsuitslag aan de ouders gerapporteerd. Vervolgens stellen de leerkrachten, soms nog na overleg met de ouders, voor iedere leerling het tweede en definitieve advies vast. Het is de bedoeling dat ook dit advies en de toetsscore ter beschikking komen van de toelatingscommissies. Deze nemen de toelatingsbeslissingen.

Hoewel geenszins vaststaat dat de door het CITO aanbevolen procedure overal strikt gevolgd wordt, kunnen wij bij gebrek aan nadere gegevens vooralsnog niet anders doen dan ons baseren op de beschreven procedure. Uit de beschrijving van deze procedure blijkt dat er drie centrale partijen te onderscheiden zijn, die ieder hun beslissingen nemen. Leerkrachten dienen een eerste en een tweede advies op te stellen, ouders dienen te



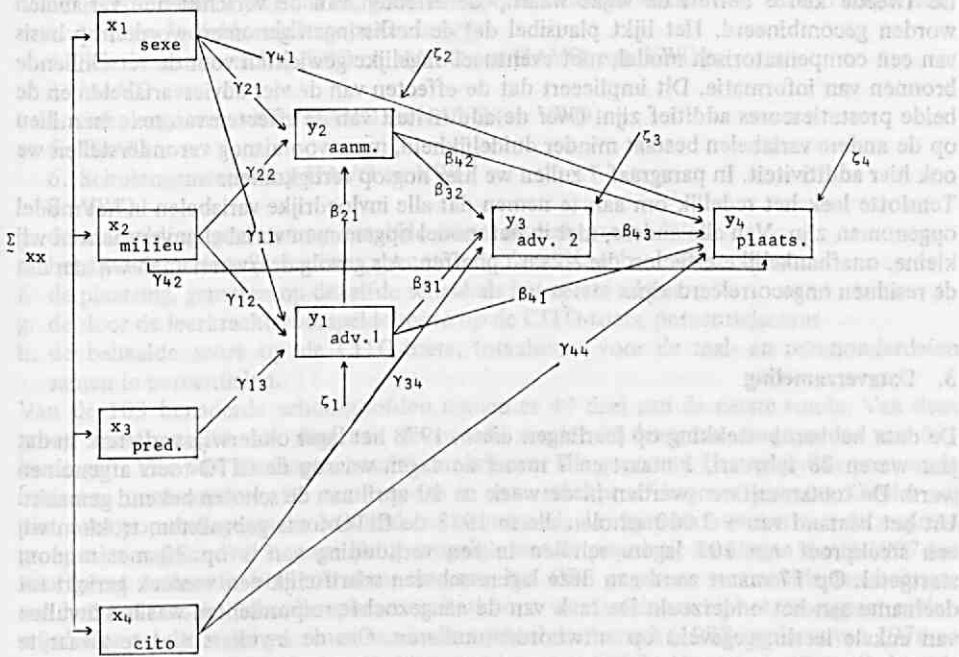
*Figuur 1. Structureel model waarin de beslissingen van leerkrachten, ouders en toelatingscommissies opgenomen zijn benevens de informatie die hen ter beschikking staat*

besluiten op welke school zij hun kind gaan aanmelden, toelatingscommissies beslissen over de toelating. Wanneer de leerkracht zijn eerste advies opstelt, beschikt hij over een in de klas verkregen indruk aangaande de capaciteiten van de leerling. De ouders zijn, wanneer zij hun kind aanmelden, op de hoogte van het eerste advies. Als de leerkracht zijn tweede advies geeft, heeft hij de beschikking over het eerste advies, de aanmelding of schoolkeuze van de ouders en de score op de schooltoets. De toelatingscommissie is in het bezit van de beide onderwijzersadviezen en kent de schoolkeuze van de ouders en de schooltoetsscore van de leerling. Omdat we weten over welke informatie de verschillende partijen beschikken, ligt het op te stellen structurele model voor de hand (zie figuur 1). De variabele 'predictie' in deze figuur staat voor de voorspelde toetsscore. Wij veronderstellen dat de leerkracht in zijn eerste advies rekening houdt met de verwachte of voorspelde score op de schooltoets.

Het is mogelijk dat de beslissingen tijdens de overgangsprocedure niet alleen gebaseerd worden op de informatie waarover de drie partijen beschikken, maar ook afhankelijk zijn van de sexe en het sociale milieu van de leerling. Daarom worden sexe en milieu nog als oorzakelijke variabelen aan het model toegevoegd met een directe invloed op de beslissingen van de drie partijen. In het volledige model, model A (zie figuur 2), worden dus de volgende relaties opgenomen:

- het eerste onderwijzersadvies wordt direct bepaald door de sexe van de leerling, het sociale milieu en de verwachte prestatie op de toets;
- de aanmelding door de ouders wordt direct bepaald door het geslacht van hun kind, het sociale milieu en het eerste advies van de onderwijzer;
- het tweede onderwijzersadvies wordt direct bepaald door het eerste advies (waarin sexe en milieu-invloeden al verdisconteerd zijn evenals de verwachte toetsscore), de aanmelding en de behaalde toetsscore;
- de plaatsing wordt direct bepaald door de sexe en het sociale milieu van de leerling, de beide onderwijzersadviezen, de aanmelding en de behaalde toetsscore.

Om het model volledig te specificeren, dienen we nog een drietal keuzes te maken. De



$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -\beta_{21} & 1 & 0 & 0 \\ -\beta_{31} & -\beta_{32} & 1 & 0 \\ -\beta_{41} & -\beta_{42} & -\beta_{43} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & 0 \\ \gamma_{21} & \gamma_{22} & 0 & 0 \\ 0 & 0 & 0 & \gamma_{34} \\ \gamma_{41} & \gamma_{42} & 0 & \gamma_{44} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{pmatrix}$$

waarbij  $E(y)=0$ ,  $E(x)=0$ ,  $E(\zeta)=0$ ,  $E(x\zeta')=0$ ,  $E(\zeta\zeta')=$ diagonaal.

Figuur 2. Pad diagram voor model A met bijbehorende modelvergelijkingen

eerste betreft de vorm van de relaties tussen de variabelen. Wij achten het plausibel dat de relaties tussen de vier adviesvariabelen, d.w.z. advies 1, aanmelding, advies 2 en plaatsing, lineair zijn: een toename bijvoorbeeld in advies 1 geeft waarschijnlijk aanleiding tot een evenredige toename in aanmelding, in advies 2 en in plaatsing. Hetzelfde geldt voor de relatie tussen de voorspelde en behaalde toetscore. Ook de relaties tussen de vier adviesvariabelen en de twee prestatiescores (predictie, toetscore) zullen waarschijnlijk lineair zijn: de adviezen die leerlingen ontvangen zijn waarschijnlijk evenredig met hun prestaties. Over de lineariteit van de relaties tussen milieu en de andere variabelen kunnen we minder stellig zijn. Maar zolang er voor andersoortige relaties geen aanwijzingen zijn, menen wij er goed aan te doen eenvoudigweg lineaire relaties te veronderstellen. De lineariteit van de relaties tussen sexe en de andere variabelen vormt geen probleem omdat sexe een dichotome variabele is.



De tweede keuze betreft de wijze waarop de effecten van de verschillende variabelen worden gecombineerd. Het lijkt plausibel dat de beslissingen genomen worden op basis van een compensatorisch model, met eventueel ongelijke gewichten voor de verschillende bronnen van informatie. Dit impliceert dat de effecten van de vier adviesvariabelen en de beide prestatiescores additief zijn. Over de additiviteit van de effecten van sexe en milieu op de andere variabelen bestaat minder duidelijkheid, maar vooralsnog veronderstellen we ook hier additiviteit. In paragraaf 7 zullen we hier nog op terugkomen.

Tenslotte leek het redelijk om aan te nemen dat alle invloedrijke variabelen in het model opgenomen zijn. Van alle andere, niet in het model opgenomen variabelen verwachten wij kleine, onafhankelijke effecten die elkaar opheffen. Als gevolg daarvan nemen wij aan dat de residuen ongecorrleerd zijn.

### 3. Dataverzameling

De data hebben betrekking op leerlingen die in 1978 het lager onderwijs verlieten. In dat jaar waren 28 februari, 1 maart en 7 maart de dagen waarop de CITO-toets afgenomen werd. De toetsresultaten werden in de week na 10 april aan de scholen bekend gemaakt. Uit het bestand van  $\pm 3.060$  scholen die in 1978 de CITO-toets gebruikten, trokken wij een steekproef van 103 lagere scholen in een verhouding van 1 op 30 met random startgetal. Op 17 maart werd aan deze lagere scholen schriftelijk een verzoek gericht tot deelname aan het onderzoek. De taak van de aangezochte respondenten was het invullen van enkele leerlinggegevens op antwoordformulieren. Om de inyullers niet te zwaar te belasten kon één groep volstaan met gegevens van leerlingen wier achternaam begon met een van de letters A t/m K, de andere groep diende gegevens te verschaffen over de L t/m Z leerlingengroep. De betrokken gegevens waren het geslacht van de leerling, beroep van de vader, het eerste onderwijzersadvies en de voorspelde score op de toets. Wat dit laatste betreft luidde de instructie als volgt: 'Het bijzondere van dit onderzoek is dat wij u vragen voor deze leerlingen te voorspellen welke score zij gehaald hebben op de toets. De vraag is minder moeilijk dan hij lijkt. Precies voorspellen is onmogelijk. Dat vragen wij u dan ook niet. Waar het ons om te doen is, is dat u een globale voorspelling geeft. U kunt hierbij iedere score tussen 1 en 100 gebruiken, maar misschien verdient het aanbeveling om alleen de vijftallen te gebruiken. Het voorspellen wordt hierdoor gemakkelijker.' In een tweede ronde werd aan de scholen die voor 10 april het eerste antwoordformulier gere-  
tourneerd hadden, een tweede formulier gezonden waarop ingevuld diende te worden de behaalde score, het tweede advies, het schooltype van aanmelding en van plaatsing. Alle voor het onderzoek benodigde gegevens werden verstrekt door schoolhoofden of zesde klasleerkrachten. Het gaat om de volgende variabelen:

- a. de sexe van de leerling
- b. het beroepsniveau van de vader, ingedeeld in zes beroepsniveaus volgens de indeling van Westerlaak e.a. (1975) als indicatie voor het sociaal milieu. De schaal kent de volgende categorieën:
  1. ongeschoolde arbeiders
  2. geschoolde arbeiders
  3. lagere employees
  4. kleine zelfstandigen
  5. hogere employees
  6. academische beroepen



- c. het eerste onderwijzersadvies, gemeten op een ordinale schaal met zeven categorieën:
  1. LBO en Individueel Beroeps Onderwijs
  2. Scholengemeenschap LBO en MAVO (en HAVO, evt. VWO)
  3. MAVO
  4. Scholengemeenschap MAVO en HAVO (en VWO)
  5. HAVO
  6. Scholengemeenschap HAVO en VWO
  7. VWO
- d. de aanmelding, gemeten op dezelfde schaal als het eerste advies
- e. het tweede onderwijzersadvies, gemeten op dezelfde schaal als het eerste
- f. de plaatsing, gemeten op dezelfde schaal als het eerste advies
- g. de door de leerkracht voorspelde score op de CITO-toets, percentielscores
- h. de behaalde score op de CITO-toets, totaalscore voor de taal- en rekenonderdelen samen in percentielen.

Van de 103 benaderde schoolhoofden namen er 49 deel aan de eerste ronde. Van deze groep namen er 45 ook deel aan de tweede ronde. Zij leverden de gegevens van 591 leerlingen. Van 514 leerlingen beschikten wij over alle gegevens. Hun gegevens vormen de basis van onze analyse. De non-respons in onze steekproef is groot geweest. Omdat wij niet over populatiegegevens beschikken, kunnen wij de representativiteit van de responsgroep niet nagaan. Wel is vergelijking mogelijk met de populatie leerlingen die in 1977 het voortgezet onderwijs binnentraden, waarvan het CBS gegevens verzamelde ten behoeve van het onderzoek 'Schoolloopbaan en herkomst van leerlingen bij het voortgezet onderwijs.' Hoewel dit strikt genomen een andere populatie betreft (1977 in plaats van 1978 en een landelijke groep in plaats van de groep door het CITO getoetste leerlingen), kan de vergelijking wel laten zien in hoeverre onze responsgroep grote afwijkingen vertoont van de landelijke situatie. Wij konden een vergelijking maken wat betreft sexe, tweede advies en plaatsing. Bij alle drie de vergelijkingen constateerden wij statistisch significante afwijkingen. Onze responsgroep bevat iets meer jongens ( $\chi^2 = 5.63$ ,  $df = 1$ ,  $p < .05$ ); wat betreft het tweede advies is in onze groep het LBO oververtegenwoordigd ten koste van het MAVO ( $\chi^2 = 11.36$ ,  $df = 2$ ,  $p < .01$ ); wat betreft de plaatsing is in onze groep niet alleen het LBO, maar ook het VWO en HAVO enigszins oververtegenwoordigd ten koste van het MAVO ( $\chi^2 = 22.21$ ,  $df = 2$ ,  $p < .001$ ). Omdat de verschillen niet erg groot zijn, hebben wij de indruk dat onze responsgroep geen ernstige afwijkingen vertoont van de landelijke leerlingengroep.

#### 4. Methode

Gezien de probleemstelling van dit onderzoek, namelijk het bepalen van de grootte van de effecten van de verschillende variabelen op elkaar en de specificatie van het model in lineaire structurele vergelijkingen lijkt de keuze van de LISREL-procedure voor de hand te liggen (Jöreskog & Sörbom, 1978). Een probleem is echter dat de variabelen hoogstens op ordinaal niveau gemeten zijn. Men zou daarom kunnen overwegen om gebruik te maken van programma's voor het analyseren van nominale of ordinale data (Goodman, 1972; Bishop e.a., 1975; Nelder & Wedderburn, 1972). Dit is echter niet mogelijk zonder een groot verlies aan informatie, omdat het aantal cellen in de tabel te groot zou worden om analyseerbaar te blijven. Als gevolg daarvan zou men categorieën samen moeten

voegen. De kans is groot dat men categorieën samenvoegt die niet samengevoegd mogen worden. Het is de vraag of deze fout opweegt tegen de fout die we maken wanneer we van interval-statistiek gebruik maken, ook al zijn de variabelen in zeven ordinaal (Blalock, 1964; Reynolds, 1977).

Een andere mogelijkheid die men zou kunnen overwegen, is om deze data te analyseren met behulp van discriminant analyse zoals gedaan is door Groeneboom e.a. (1978). Hoewel deze keuze op basis van het datamateriaal terecht zou zijn, valt echter toch in verband met de interpretatie-mogelijkheden het gebruik van LISREL te overwegen. We moeten dan echter enig inzicht hebben in de fouten die we daardoor kunnen maken.

Voordat we hierop ingaan, moeten wij allereerst vermelden dat in de twee-dimensionale analyse bleek dat de relaties tussen alle variabelen monotoon stijgend zijn overeenkomstig onze verwachting, behalve bij de variabele milieu waarvoor de categorie 4 (kleine zelfstandigen) gemiddeld lager scoorde op de andere variabelen dan categorie 3 (lagere employees). Aangezien daarmee de reden vervalt om categorie 4 hoger te achten dan categorie 3, hebben we in alle volgende analyses deze twee ranggetallen omgewisseld. Als gevolg daarvan heeft onze variabele milieu misschien meer de betekenis van opleidingsniveau van de ouders dan de oorspronkelijke schaal.

Nu we mogen aannemen dat we voor alle variabelen tenminste de juiste ranggetallen hebben blijft nog het probleem dat de afstanden tussen de verschillende categorieën niet noodzakelijk gelijk hoeven te zijn, zoals het geval is wanneer we de ranggetallen als waarden op een intervalschaal opvatten. Gegeven het ordinale niveau van de variabelen is elke monotone transformatie mogelijk zonder dat dit tot verandering in de ordening van de klassen leidt. Wat echter wel verandert onder dergelijke transformaties is de relatie tussen de variabelen. Omdat we echter veronderstellen dat alle relaties lineair zijn, zou het feit dat foute getallen zijn toegekend aan de verschillende categorieën kunnen betekenen dat de relaties dan niet meer lineair zijn. Als dus de eta-coëfficiënt voor de verschillende paren van variabelen veel hoger is dan de correlatiecoëfficiënt, dan zou dit kunnen bete-

Tabel 1

De p.m.-correlatiecoëfficiënten (boven) en de eta-coëfficiënten (onder) voor 514 leerlingen

|            | sexe          | milieu       | predictie    | Cito         | advies 1     | advies 2     | aanmelding   |
|------------|---------------|--------------|--------------|--------------|--------------|--------------|--------------|
| milieu     | -.052<br>.052 |              |              |              |              |              |              |
| predictie  | -.104<br>.104 | .265<br>.268 |              |              |              |              |              |
| Cito-toets | -.073<br>.073 | .291<br>.292 | .818<br>.820 |              |              |              |              |
| advies 1   | -.155<br>.155 | .332<br>.339 | .817<br>.842 | .779<br>.798 |              |              |              |
| advies 2   | -.150<br>.150 | .326<br>.332 | .817<br>.842 | .818<br>.838 | .914<br>.916 |              |              |
| aanmelding | -.152<br>.152 | .340<br>.350 | .787<br>.801 | .785<br>.801 | .886<br>.892 | .944<br>.945 |              |
| plaatsing  | -.148<br>.148 | .338<br>.348 | .791<br>.806 | .796<br>.811 | .890<br>.895 | .951<br>.952 | .992<br>.992 |

kenen dat foute getallen zijn toegekend aan de verschillende categorieën. Is dit niet het geval dan zouden deze getallen bij benadering juist kunnen zijn. In tabel 1 hebben we voor alle paren van variabelen de p.m.-correlatie-coëfficiënt vermeld en de eta-coëfficiënt. Uit deze tabel blijkt dat de eta-coëfficiënten slechts minimaal groter zijn dan de correlatiecoëfficiënten. Voor de onderlinge relaties tussen de variabelen advies 1, advies 2, aanmelding en plaatsing is dit niet verwonderlijk, omdat ze zijn gemeten met dezelfde schaal. Elke verandering in de ene schaal zal dus ook de andere schalen veranderen en de relaties zullen nauwelijks veranderen, temeer daar de correlaties erg hoog zijn. Voor deze variabelen geeft de tabel ons alleen een indicatie dat de veronderstelling van lineariteit voor deze variabelen niet onjuist was. Hetzelfde geldt voor de onderlinge relatie tussen de variabelen predictie en Cito-toets. Ook deze relatie blijkt lineair te zijn. De relaties tussen de vier adviesvariabelen en de twee prestatie-variabelen vertonen nauwelijks grotere afwijkingen dan de relaties binnen deze twee groepen. Hetzelfde geldt voor de relaties tussen al deze variabelen en de variabele beroep. Voor de variabele sexe mogen we geen verschillen verwachten aangezien deze variabele dichotoom is. Deze resultaten suggereren dat de getallen, die aan de verschillende klassen zijn toegekend, niet dermate fout zijn dat ze de relaties tussen de variabelen sterk vertekend hebben, aannemend dat onze veronderstelling van lineariteit juist is.

Om na te gaan hoe robuust de correlatiecoëfficiënt is bij plausibele niet-lineaire, monotone transformaties hebben we de relatie tussen de variabelen advies 1 en predictie wat nader bestudeerd. Omdat de afstanden tussen de categorieën van de adviesschaal niet gelijk hoeven te zijn aan beide uiteinden van de schaal hebben we deze schaal getransformeerd via een machtsfunctie  $y = (\text{advies } 1)^b$  waarbij aan  $b$  verschillende waarden zijn gegeven. Voor de variabele predictie lijkt het meer plausibel dat de schaal te zeer is uitgerekt in het midden, aannemende dat schoolprestaties een normaal verdeelde variabele is. Daarom hebben we voor elke percentielscore de bijbehorende z-score ingevuld. In tabel 2 is te zien hoe de correlatiecoëfficiënt verandert door deze transformaties. Het is duidelijk dat de effecten van de transformaties slechts zeer gering zijn. De meest extreme verschillen bedragen minder dan 5% ten opzichte van de correlatiecoëfficiënt. Toch zijn de veranderingen in de schaalwaarden voor de klassen aanzienlijk. Aangezien het niet plausibel leek dat de afwijkingen van de schaalwaarden groter zouden zijn dan de afwij-

Tabel 2

De correlatiecoëfficiënt voor de relatie tussen de variabelen advies 1 en predictie onder verschillende transformaties

| transformaties volgens<br>(advies 1) <sup>b</sup> met: | correlatie coëfficiënten van advies 1 en predictie |                       |
|--|--|-----------------------|
|  | ongetransformeerd                                  | predictie<br>z-scores |
| b = .25  | .810   | .795                  |
| .50  | .821   | .811                  |
| .75  | .820   | .817                  |
| 1.00   | .817   | .803                  |
| 1.25   | .794   | .807                  |
| 1.50   | .801   | .811                  |
| 1.75   | .780   | .798                  |
| 2.00   | .776   | .795                  |

kingen van de ranggetallen die we hier hebben gepostuleerd, hebben we geconcludeerd dat de fout bij het analyseren van deze data uitgaande van de hierboven vermelde correlatiematrix slechts tot zeer geringe fouten aanleiding zal geven. Wij zijn van mening dat het gebruik van statistische technieken gebaseerd op interval meetniveau in dit geval zonder veel problemen kan geschieden.

Gegeven dit uitgangspunt hebben we de LISREL procedure verkozen boven schatting van de effecten m.b.v. de kleinste kwadraten methode en wel om de volgende redenen:

1. De resultaten zullen voor ons model nauwelijks verschillen voor de twee methoden;
2. De LISREL-procedure levert automatisch inzicht in de mate waarin het model bij de data past, terwijl men deze informatie zelf moet uitwerken in het andere geval;
3. De LISREL-procedure geeft ook inzicht in de totaal effecten van de verschillende variabelen. Ook deze effecten zouden apart berekend moeten worden in het andere geval;
4. De LISREL-benadering maakt het mogelijk om latente variabelen te introduceren in het model, hetgeen niet mogelijk is in het andere geval.

In de LISREL-procedure wordt de waargenomen correlatiematrix gebruikt als data-basis en aangeduid met  $\underline{S}$ . De correlatiematrix van de populatie wordt aangeduid met  $\underline{\Sigma}$ . Aangegeven kan worden dat deze matrix een functie is van de parameters van het model. We zullen deze parameters aanduiden met de vector  $\underline{\pi}$ . Deze vector kan vrije parameters bevatten, parameters die a priori een vaste waarde hebben gekregen en parameters die gelijk moeten zijn aan andere parameters. Om een schatting te krijgen van de parameters wordt functie  $f(\underline{\pi})$  geminimaliseerd (zie Saris, 1978). Indien de variabelen multivariaat normaal verdeeld zijn levert de schattingsprocedure de meest aannemelijke schatters op van de parameters in  $\underline{\pi}$ . Deze aanname kan hier echter niet worden gemaakt. Dat wil echter nog niet zeggen dat deze procedure daarom niet geschikt zou zijn. Het is namelijk bewijsbaar dat deze procedure ook bij afwijkingen van de multivariate normaalverdeling consistente schatters oplevert (Dijkhuizen, 1978). Echter, vanwege de afwijkende verdeling moet wel de extra informatie die het programma LISREL levert zoals standaardfouten voor de schatters en de 'goodness of fit' test met voorzichtigheid worden gehanteerd. We zullen daarom bij de bestudering van de passing van het model voornamelijk afgaan op de residuen die overblijven wanneer de geobserveerde correlaties worden afgetrokken van de gereproduceerde correlaties die zijn berekend op basis van de geschatte waarden van de parameters. De grootte van deze residuen is een indicatie voor het feit of het model bij de data past.

## 5. Toets en verbetering van het model

Wij eindigden paragraaf 2 met de formulering van model A. In figuur 2 werd een diagram van model A opgenomen met de gespecificeerde vergelijkingen. Op basis van de data in tabel 1 en de procedure, die hiervoor beschreven is, zijn de parameters geschat. Vervolgens zijn ook de residuen berekend om na te gaan of het model past bij de data. Uit de residu-matrix blijkt het model niet volledig te voldoen gezien de grote residuen voor de relaties van de Cito-score met andere variabelen (nl.  $-.102$ ,  $-.183$ ,  $-.135$  en  $-.174$  met resp. advies 1, aanmelding, advies 2 en plaatsing). De overige residuen zijn wel aanvaardbaar en bedragen absoluut gezien maximaal  $.064$ . Ook de 'goodness of fit' test berekend door het programma laat zien dat dit model niet goed past bij de data ( $\chi^2 = 122.4521$ , df

= 7). Dit kan er op wijzen dat de relaties tussen de Cito-score en de endogene variabelen niet correct gespecificeerd zijn. Aangezien echter vastligt dat de toetsscore pas na het eerste advies en de aanmelding maar voor het tweede advies en de plaatsing een rol kan spelen, kunnen er geen andere relaties tussen deze variabelen zijn dan in model A aangegeven is. De slechte verklaring van de relaties tussen de Cito-score en de adviesvariabelen zou echter ook kunnen optreden als gevolg van een variabele, die niet in het model opgenomen is, maar wel invloed uitoefent op de Cito-score en op de adviesvariabelen. De variabele predictie kan deze rol niet vervullen, omdat deze variabele een kenmerk van de leerkracht is en onbekend bij de andere beslissers. De meest voor de hand liggende variabele die invloed kan uitoefenen op zowel de adviesvariabelen als op de toetsscore wordt gevormd door de schoolprestaties van de leerling. De verandering die we daarom in model A zouden kunnen aanbrengen is het opnemen van een ongemeten oorzakelijke variabele, die wij voorlopig Schoolprestaties (SP) noemen. Van het nieuwe model, model B, wordt de modelspecificatie hieronder weergegeven:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\beta_{21} & 1 & 0 & 0 & 0 & 0 & 0 \\ -\beta_{31} & -\beta_{32} & 1 & 0 & 0 & 0 & 0 \\ -\beta_{41} & 0 & -\beta_{43} & 1 & 0 & 0 & 0 \\ -\beta_{51} & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\beta_{63} & -\beta_{64} & -\beta_{65} & 1 & 0 \\ 0 & 0 & -\beta_{73} & -\beta_{74} & -\beta_{75} & -\beta_{76} & 1 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ 0 & 0 \\ \gamma_{31} & \gamma_{32} \\ \gamma_{41} & \gamma_{42} \\ 0 & 0 \\ 0 & 0 \\ \gamma_{71} & \gamma_{72} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \\ \zeta_6 \\ \zeta_7 \end{pmatrix}$$

waarbij  $E(\eta) = 0$ ,  $E(\xi) = 0$ ,  $E(\zeta) = 0$ ,  $E(\xi\xi') = 0$ ,  $E(\zeta\zeta') = 0$ ,  $E(\xi\xi') = \text{diagonaal}$

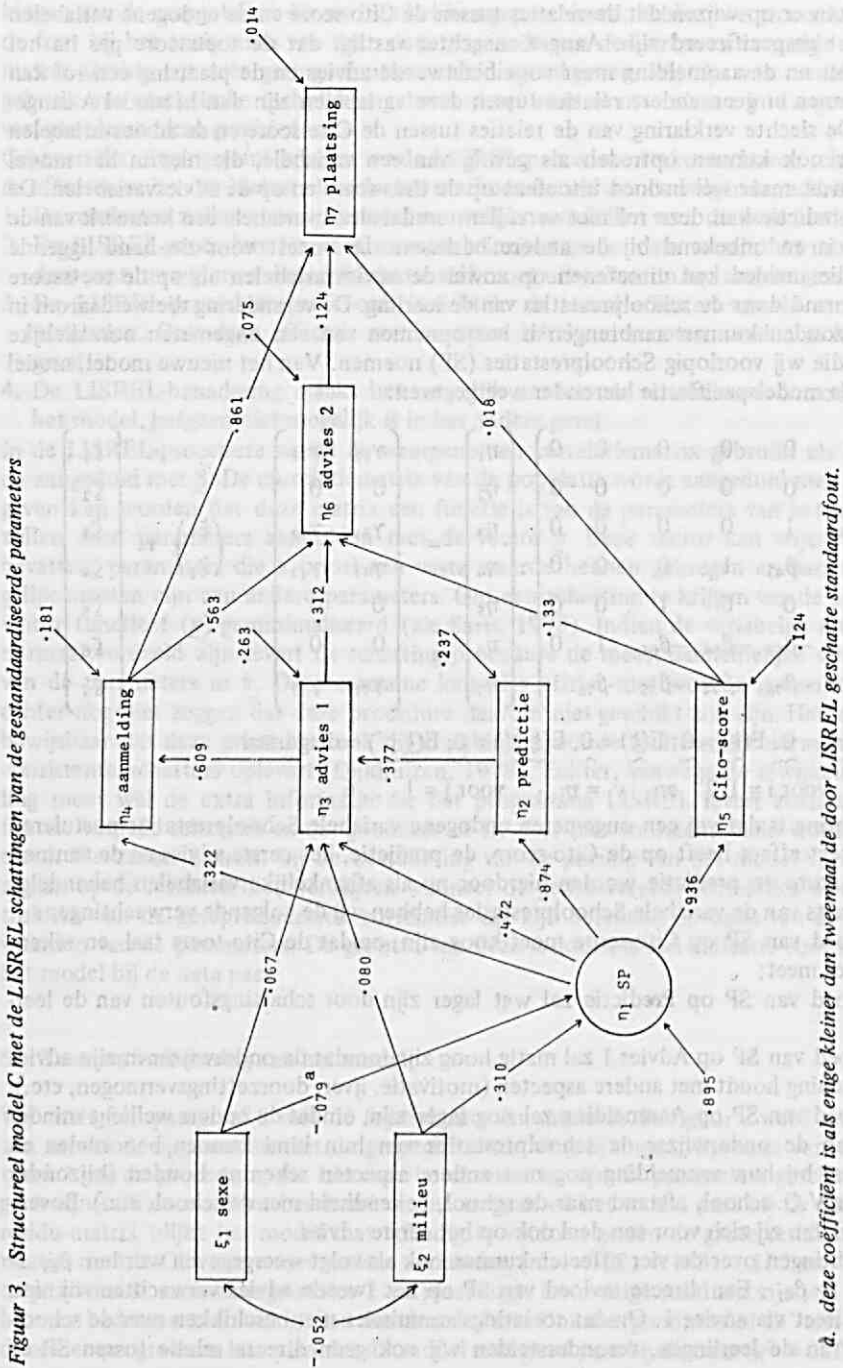
en  $x_i = \xi_i$  voor  $i = 1, 2$  en  $y_j = \eta_{j+1}$  voor  $j = 1 \dots 6$

De verandering is dat wij een ongemeten endogene variabele Schoolprestaties postuleren die een direct effect heeft op de Cito-score, de predictie, het eerste advies en de aanmelding. Cito-score en predictie worden hierdoor nu als afhankelijke variabelen behandeld. Over de plaats van de variabele Schoolprestaties hebben wij de volgende verwachtingen:

- de invloed van SP op Cito-score moet hoog zijn, omdat de Cito-toets taal- en rekenprestaties meet;
- de invloed van SP op Predictie zal wat lager zijn door schattingsfouten van de leerkracht;
- de invloed van SP op Advies 1 zal matig hoog zijn, omdat de onderwijzer in zijn advies ook rekening houdt met andere aspecten (motivatie, ijver, doorzettingsvermogen, etc.);
- de invloed van SP op Aanmelding zal nog lager zijn, omdat de ouders wellicht minder goed dan de onderwijzer de schoolprestaties van hun kind kunnen beoordelen en daarnaast bij hun aanmelding nog met andere aspecten rekening houden (bijzonder karakter V.O.-school, afstand naar de school, bekendheid met de school, etc.). Bovendien verlaten zij zich voor een deel ook op het eerste advies.

De verwachtingen over de vier effecten kunnen ook als volgt weergegeven worden:  $\beta_{51} > \beta_{21} > \beta_{31} > \beta_{41}$ . Een directe invloed van SP op het tweede advies verwachten wij niet, slechts indirect via advies 1. Omdat toelatingscommissies niet beschikken over de schoolprestaties van de leerlingen, veronderstelden wij ook geen directe relatie tussen SP en





plaatsing. Het effect van sexe op SP moet ongeveer nul zijn, het effect van milieu op SP moet laag tot matig zijn en positief.

Om identificatie van het model mogelijk te maken, werd  $\beta_{51}$  op 1 gefixeerd. Daardoor zijn alle effecten in model B weer uniek te bepalen met de LISREL-procedure. Alle elementen van de residu-matrix van model B blijken nu tot bijna nul gereduceerd te zijn. De grootste absolute afwijking is 0.024, een acceptabel verschil. De gemiddelde absolute afwijking is .003. De verandering in het model blijkt de grote residuen voor de Cito-score tot kleine afwijkingen te hebben gereduceerd. Dit resultaat impliceert dat wij er in zijn geslaagd een model op te stellen dat de relaties tussen de variabelen bevredigend verklaart. Dat model B een verbetering is in vergelijking met model A blijkt ook uit de test van de 'goodness of fit' ( $\chi^2 = 6.1570$ ,  $df = 6$ ). In model B kunnen we nog wel enige vereenvoudigingen aanbrengen door de relatief kleine effecten weg te laten. In model C zijn daarom alle effecten die kleiner zijn in absolute waarde dan tweemaal de door LISREL geschatte standaardfout op nul gesteld. Dit betreft het effect van sexe en milieu op aanmelding en op plaatsing en het effect van advies 1 op plaatsing. Deze veranderingen hebben geen nadelige gevolgen voor de passing van het model. Het absoluut grootste element van de residu-matrix is .031 en de gemiddelde absolute afwijking bedraagt .006. Ook de 'goodness of fit' test van model C is acceptabel ( $\chi^2 = 12.2448$ ,  $df = 11$ ). In figuur 3 worden de schattingen voor de gestandaardiseerde parameters gegeven. Aangaande de effecten van de latente variabele, die wij voorlopig Schoolprestaties noemden, valt het volgende te constateren. SP heeft een groot direct effect op de Cito-score en verklaart 88% van de variantie daarvan. Omdat de Cito-score een maat voor reken- en taalprestaties is, vermoeden wij dat de latente variabele veel overeenkomst vertoont met schoolprestaties op de kernvlakken rekenen en taal, al achten wij het mogelijk dat ook schoolprestaties op de andere vakken onze latente variabele mede bepalen.

De door de onderwijzer voorspelde toetscore blijkt ook grotendeels te berusten op Schoolprestaties. De directe invloed van Schoolprestaties op het eerste advies en op de aanmelding is minder groot maar nog in ruime mate aanwezig.

De variabele Schoolprestaties ondervindt een zeer zwak effect van sexe. Het effect van milieu op Schoolprestaties is zwak en positief. Uit dit overzicht blijkt de variabele Schoolprestaties precies die plaats in het model in te nemen die we verwachtten.

Gezien dit resultaat en de kleine residuen en het feit dat ook andere parameterwaarden zinvol blijken te zijn, zoals we hierna zullen laten zien, concluderen we dat het niet nodig is te trachten het model verder te verbeteren. We beschouwen daarom model C als het uiteindelijke model en zullen in de volgende paragraaf het gevonden resultaat proberen te interpreteren.

## 6. Interpretatie

Circa 99% van de variantie van de plaatsing wordt verklaard door de variabelen die in ons model opgenomen zijn. Directe beïnvloeding vindt voor het grootste deel plaats door de aanmelding. Inspectie van de betreffende kruistabel leert ons dat slechts 2% van de 514 leerlingen op een andere school geplaatst zijn dan waar zij werden aangemeld. De aanmelding zou dus zo goed als doorslaggevend voor de plaatsing beschouwd kunnen worden. Dit resultaat mag echter niet al te veel gewicht gegeven worden want ook voor de variabele advies 2 blijkt dat het percentage leerlingen dat anders geplaatst werd dan geadviseerd



geslacht, milieu en predictie, die via de variabele advies 1 van invloed zijn. Hun effecten zijn relatief echter veel kleiner, al valt het indirecte effect van predictie op het tweede advies (.247) en op aanmelding (.230) niet te verwaarlozen. Eigenlijk is het verbazingwekkend dat het eerste advies niet alleen gebaseerd wordt op de schoolprestaties, maar ook op de voorspelde toetsscore. Dit betekent dat de variabele predictie waarschijnlijk iets anders indiceert dan alleen maar schoolprestaties. Bovendien blijkt de invloed van de voorspelde toetsscore op het eerste advies zelfs groter te zijn dan de invloed van de behaalde score op het tweede advies. De redenen hiervoor kunnen velerlei zijn, maar een mogelijke interpretatie luidt dat de status van de Cito-toets zo groot is, dat er enige anticiperende invloed van uitgaat, waardoor correctie in het tweede advies minder nodig is.

Verder kan nog worden opgemerkt dat in ons model niet alle relevante variabelen opgenomen zijn. Het eerste advies en in mindere mate de aanmelding blijken niet volledig verklaard te worden. Advies 1 heeft een onverklaarde variantie van .263, aanmelding van .181. Er is dus nog enige ruimte over voor andere verklarende variabelen zoals bijvoorbeeld motivatie, ijver en doorzettingsvermogen, waar de leerkracht wellicht rekening mee houdt. Via het eerste advies en de aanmelding kunnen dergelijke variabelen indirect nog enige rol spelen bij de uiteindelijke plaatsing.

## 7. Conclusies

In de conclusies knopen wij allereerst aan bij de in de inleiding gestelde vragen. Omdat het effect van sexe op Schoolprestaties uiterst laag is, concluderen wij dat jongens en meisjes niet verschillen in schoolprestaties. Onafhankelijk van schoolprestaties oefent sexe een kleine directe invloed uit op het eerste onderwijzersadvies (zie tabel 3). Dit effect is zodanig dat meisjes iets vaker hogere adviezen krijgen dan jongens. Hoewel een dergelijk direct effect van sexe niet aanwezig is op de aanmelding of de plaatsing, behouden meisjes hun voordeel wel tot en met de plaatsing vanwege indirecte effecten via advies 1. Er is dus geen sprake van dat meisjes bij de LO-VO overgang een achterstand oplopen. Hun schoolprestaties verschillen niet met die van jongens en via het eerste advies van de onderwijzer valt de uiteindelijke plaatsing voor hen iets kansrijker uit.

Methodische overwegingen dwongen ons de schaal van de variabele milieu aan te passen. Als gevolg daarvan hebben wij het milieu van lagere employees een hogere waardering toegekend dan het milieu van kleine zelfstandigen. Het na deze aanpassing geschatte effect van .310 van milieu op Schoolprestaties laat zien dat leerlingen uit lagere milieus lagere Schoolprestaties geleverd hebben. Onafhankelijk van Schoolprestaties blijkt er een klein effect van milieu op advies 1 te bestaan. Een direct effect van milieu op de aanmelding of de plaatsing hebben wij niet gevonden. Maar via het eerste advies blijft er wel tot en met de plaatsing een indirect effect van milieu bestaan (zie tabel 3). Leerlingen uit lagere milieus hebben dus op 12 à 13-jarige leeftijd een achterstand in schoolprestaties. Onafhankelijk van deze achterstand lopen deze leerlingen via een direct effect van milieu op advies 1 een kleine extra achterstand op tijdens de LO-VO overgang.

In ons compensatiemodel blijken de variabelen geslacht en milieu slechts geringe effecten te hebben. Het zou echter ook kunnen zijn dat deze beide variabelen een interactie-effect hebben met de andere variabelen. Teneinde dit na te gaan verdeelden we onze responsgroep in vier subgroepen: meisjes hoog milieu, meisjes laag milieu, jongens hoog milieu en jongens laag milieu. Uit een simultane analyse bleek voor deze vier groepen hetzelfde



model te kunnen gelden als voor de totale groep. Wel bleken sommige padcoëfficiënten niet in alle groepen precies gelijk te kunnen zijn. Omdat de gevonden verschillen echter klein zijn, was er voor ons geen aanleiding de conclusies te differentiëren voor de vier subgroepen.

Uit het feit dat de score op de Cito-toets een uiterst klein direct effect heeft op de plaatsing, menen wij te kunnen afleiden dat plaatsingscommissies geen beslissingen baseren op de toetsscore alleen. Via het tweede advies oefent de Cito-score nog enige indirecte invloed uit op de plaatsing, maar ook dit effect (.016) is zeer gering. Blijkbaar voegt de toetsscore nauwelijks nog iets toe aan de informatie over schoolprestaties en voorspelde toetsscore die reeds in de adviezen en de aanmelding verwerkt zijn. Als gevolg daarvan concluderen wij dat de invloed van de score op de toets tijdens de LO-VO overgang verwaarloosbaar klein is. Dit betekent niet dat de toets zelf geen invloed heeft op de LO-VO overgang. Want van de door de leerkracht verwachte score op de toets blijkt een grotere invloed uit te gaan. Via dit direct effect van predictie op advies 1 heeft de Cito-toets een invloed op de uiteindelijke plaatsing die niet verwaarloosd kan worden (.228). Het merkwaardige feit doet zich dus voor dat de toets wel enige invloed heeft op de overgang, niet via de behaalde maar via de verwachte score. Wanneer men dit geen wenselijke situatie zou vinden, rijst de vraag hoe daarin verandering is aan te brengen. Overwogen zou kunnen worden de toets eerder af te nemen, zodat de toetsscore ter beschikking komt voordat het eerste advies gegeven wordt. Voordeel hiervan is in ieder geval dat de ouders bij de aanmelding beschikken over de toetsscore. Zij zijn dan niet alleen meer afhankelijk van het onderwijzersadvies. De effecten van een vervroegde toetsafname op het onderwijzersadvies zijn niet helemaal zeker. Grote veranderingen zijn echter niet te verwachten, omdat de toets de plaats van de variabele predictie zal gaan innemen in het model. Als gevolg daarvan zal het effect van de voorspelde score op advies 1 waarschijnlijk plaats maken voor het effect van de behaalde score.

Tot slot merken wij op dat de invloed van schoolprestaties gedurende de gehele overgang erg groot is. Zowel de beslissingen van onderwijzers als van ouders als van toelatingscommissies blijken voor het grootste deel direct dan wel indirect gebaseerd te zijn op de schoolprestaties van leerlingen. Alleen in het eerste advies en de aanmelding is nog ruimte aanwezig voor directe effecten van andere variabelen. Over de identiteit van deze variabelen levert ons onderzoek geen aanwijzing op. Het lijkt echter niet onwaarschijnlijk dat het gaat om kwaliteiten als ijver, motivatie, doorzettingsvermogen en de eigen wens van de leerling.

## LITERATUUR

- Bishop, Y.M.M., S.E. Fienberg en P.W. Holland. *Discrete multivariate analysis: Theory and Practice*. Cambridge, Mass.; The MIT press, 1975.
- Blalock, H.M. *Causal inferences in non experimental research*. Chapel Hill: University of North Carolina press, 1964.
- Blok, H. *De Schooltoets en gelijke kansen*. Psychologisch Laboratorium aan de Universiteit van Amsterdam, 1978 (doctoraal werkstuk).
- Blok, H. en H. Westorp. De schooltoets na twaalf jaar. Deel I. Een evaluatie van de op de leerling gerichte functies: selectie voor het v.o., advisering aan de ouders en opsporing van reserve aan intellect. Deel II: Een evaluatie van de op het onderwijs gerichte functies: analyse van de stand van het onderwijs, concretisering van doelstellingen en beïnvloeding van het lager onderwijs. *Pedagogische Studiën*. 1979, 56, 335-347, 388-396.



- Boon van Ostade, A.H.. Het milieu en de overgang van het basis- naar het voortgezet onderwijs in Nederland. *Sociologische Gids*, 1972, 19, 211-219.
- Bos, J. en E. Warries. *De functie van een toetsprogramma: de Amsterdamse Schooltoetsen in 1969 en 1970*. Amsterdam; RITP, 1971.
- C.I.T.O. *Meningen over de eindtoets en plannen voor de toekomst*. Arnhem; CITO, 1978.
- Deen, N. *Een halve eeuw onderwijsresearch in Nederland. Het Nuttseminarium voor Pedagogiek aan de Universiteit van Amsterdam 1919-1969*. Groningen; Wolters, 1969 (proefschrift).
- Dijkhuizen, A.A. *LISREL-procedure: Mathematisch-statistische analyse van lineair structurele vergelijkingmodellen met latente variabelen*. Uitgave van het project onderwijs en sociaal milieu, Rotterdam, 1978.
- Dronkers, J. Manipuleerbare variabelen in de schoolloopbaan. Boeknummer van *Mens en Maatschappij* (1978) over stratificatie.
- Goodman, L.A. A general model for the analysis of surveys. *American Journal of Sociology*, 1972, 77, 1035-1086.
- Gordon, R.A. Issues in multiple regression. *American Journal of Sociology*, 1968, 73, 592-616.
- Groeneboom, P., J. Hoogstraten, G.J. Mellenbergh en J.P.H. van Santen. Relevante variabelen bij het doorverwijzen na de lagere school; een correlatieve analyse. *Tijdschrift voor Onderwijsresearch*, 1978, 3, 262-280.
- Groot, A.D. de. Het nut van een schooltoets in de zesde klas L.O. *Het Schoolblad*, 1966, 1, 288-292.
- Heek, F. van, e.a. *Het verborgen talent*. Meppel: Boom, 1968.
- Hoogstraten, J. en G.J. Mellenbergh. Relevante variabelen bij het doorverwijzen na de lagere school: een experiment. *Tijdschrift voor Onderwijsresearch*, 1978, 3, 161-172.
- Jöreskog, K.G. en D. Sörbom. *LISREL IV: analysis of linear structural relationships by the method of maximum likelihood*. Chicago; International Educational Services, 1978.
- Kemenade, J.A. van en J.A. Kropman. Verborgenen talenten? Kritische kanttekeningen bij een onjuiste interpretatie. *De Sociologische Gids*, 1972, 19, 219-229.
- Kohnstamm, G.A., Geslachtsverschillen in prestaties op schoolvorderingentoetsen en enkele tests aan het einde van de basisschool. *Nederlands Tijdschrift voor de Psychologie*, 1973, 28, 351-367.
- Kropman, J.A. en J.W.M. Collaris. *Van jaar tot jaar*. Nijmegen; ITS, 1974.
- Lutje Spelberg, H.C. en H.J. Rotteveel. De voorspellende waarde van de Groninger Schoolvorderingentoets. *Tijdschrift voor Onderwijsresearch*, 1978, 3, 3-9.
- Meijer, J.L. *Sociale atlas van de vrouw*. Sociaal en Cultureel Planbureau, Den Haag: Staatsuitgeverij, 1977.
- Nelder, J.A. en R.W.M. Wedderburn. Generalized linear models. *Journal of the Royal Statistical Society, Ser. A*, 1972, 135, 370-384.
- Oosterbaan, J.W. De Amsterdamse schooltoets in 1969 in relatie tot enige conclusies van 'Het verborgen talent' *Sociologische Gids*, 1973, 20, 88-97.
- Peschar, J.L. *Milieu school beroep; een achteraf-experiment over de periode 1958-1973 naar de invloed van het sociaal milieu op school- en beroepsloopbaan*. Groningen; Tjeenk Willink, 1975.
- Reynolds, H.T. *The analysis of cross-classifications*. London: Free Press, 1977.
- Sandbergen, S., J.J. Elshout, T. Akkerman en A. van Peet. Enkele relaties tussen een intelligentietest en een studietoets. *Nederlands Tijdschrift voor de Psychologie*, 1972, 27, 509-529.
- Saris, W.E. *Introduction to the use of linear structural equation models in non-experimental research*. Vrije Universiteit, Amsterdam, 1978.
- Schmidt, P. en E.N. Muller. The problem of multicollinearity in a multistage causal alienation model: a comparison of ordinary least squares, maximum likelihood and ridge estimators. *Quality and Quantity*, 1978, 12, 267-297.
- Westerlaak, J.M. van, J.A. Kropman en J.W.M. Collaris. *Beroepenklapper*. Nijmegen; ITS, 1975.

Manuscript ontvangen 3-8-'79

Definitieve versie ontvangen 27-11-'79

Tijdschrift voor Onderwijsresearch, 20 (1995), nr. 2, pp. 115-132

## Deregulering en de marktpositie van scholen

H.W.C.H. van Amelsvoort en B. Witziers

*Universiteit Twente, Enschede*

### ABSTRACT

This article presents the results of two research projects that were conducted within the framework of the present governmental policy with regard to deregulation and decentralization of decision making power. The first research project dealt with the changes schools go through as a result of this policy and the opinions held by school members about aspects like the legitimacy of the governmental policy and the perceived consequences. The results show that most schools react favorably towards this policy in as much that they increasingly fulfill the conditions necessarily to adjust to the new situation. Considering the opinions held by school members the conclusion is that the governmental policy is favored mostly by members of junior vocational schools and members of school boards and the least by members of primary schools and parents. The second research project investigated the opinions held by principals of secondary schools with regard to the planned introduction of a system of output financing. These results show that most principals oppose the introduction of such a system. However, there are no indications that schools will severely obstruct the implementation of output financing.

### INLEIDING

Het onderwijs in Nederland is altijd in sterke mate gebonden geweest aan regelgeving van de kant van de overheid, die dan ook altijd een belangrijke invloed heeft gehad op de vormgeving van het Nederlandse onderwijs. Van Wieringen (1987) en Smets (1985) spreken in dit verband van een relatie tussen overheid en lokale schoolorganisatie die gekenmerkt wordt door een vanaf de onderwijspacificatie toenemende regulering en centralisering van het onderwijsbeleid. Als oorzaken voor deze sterke wet- en regelgeving worden vaak redenen van kwaliteitsbewaking en financieel beheer genoemd. Daarnaast speelt ook het gelijkheidsprincipe en sterke rol; een sterke wet- en regelgeving betekent dat scholen in hoge mate onder dezelfde condities opereren, waardoor de verschillen tussen scholen gering zullen zijn.

Zoals bekend treedt er pas de laatste jaren in deze situatie een verandering op. Kernwoord hier is deregulering die in beheersmatig en inhoudelijk opzicht tot een van de belangrijkste doelstellingen van het Ministerie van Onderwijs en Wetenschappen in de negentiger jaren behoort (Giesbers, 1990). Onder deregulering wordt in het algemeen het streven verstaan naar een beperkter, globaler en helderder regelgeving (Giesbers, 1987).

Voor het primair en voortgezet onderwijs komt het streven naar deregulering tot uiting in de ministeriële nota "De school op weg naar 2000". Uitgangspunt van deze nota is het toekomstig beeld van de autonome school. Naast het beperken van de overheidsuitgaven heeft deregulering namelijk tot doel de autonomie van scholen te vergroten. Daarmee zou met name de doelmatigheid van scholen vergroot kunnen worden, evenals de flexibiliteit waarmee ze op veranderingen in onderwijsbehoeften kunnen inspelen. Met andere woorden, door meer zelfstandigheid zullen scholen beter toegerust zijn om aansluiting te vinden bij een maatschappelijke omgeving die gekenmerkt wordt door snelle veranderingen en een toenemende diversificatie op lokaal en regionaal niveau (Toenders, 1990). Dat juist op het punt van aansluiting tussen school en omgeving het onderwijs tekort schiet, blijkt uit de kritiek hierover. Zo wordt onder andere gewezen op de gebrekkige aansluiting tussen onderwijs en arbeidsmarkt (Van Hoof, 1986; Geurts, 1989) en het verwaarlozen van de belangen van onderwijsconsumenten (Versloot, 1990;

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

Wansink, 1992). Deze kritiek is overigens een internationaal verschijnsel. Zo heeft in de Verenigde Staten deze kritiek geleid tot onderwijshervormingen die voor een belangrijk gedeelte zijn ontleend aan de "public-choice"-theorie. Deze theorie stelt deregulering en het vergroten van de keuzevrijheid van onderwijsconsumenten voorop met als doel de concurrentie tussen onderwijsinstellingen te stimuleren. De verwachting daarbij is een grotere kwaliteit in de dienstverlening en een meer op de wensen van de onderwijsconsumenten toegesneden aanbod (cf. Boyd & Crowson, 1983; Chubb & Moe, 1990; Lieberman, 1989).

Globaal zijn er binnen een onderwijsinstelling vier gebieden te onderscheiden waar regelgeving vanuit de overheid van toepassing is (Giesbers, 1987; Jacobs, 1987; Marx, 1987):

- regels inzake inhoud en kwaliteit van het onderwijs
- regels inzake personeelsformatie en rechtspositie
- regels inzake huisvesting, materiële voorzieningen en bekostiging
- regels inzake de schoolorganisatie en medezeggenschap

Door deze regelingen van de overheid wordt de beleidsruimte van scholen begrensd. De ruimte die overblijft voor scholen om eigen beleid te voeren wordt vaak uitgedrukt als de relatieve autonomie van een school (Giesbers, 1987). Door deregulering zal deze relatieve autonomie toenemen. De belangrijkste consequentie daarvan is dat scholen gestimuleerd worden eigen (bestuurlijk) beleid te voeren.

Uit de genoemde nota "De school op weg naar 2000" blijkt dat de beslissingsvrijheid van scholen vooral zal groeien ten aanzien van inzet van personele en financiële middelen en de realisering van een maximaal onderwijsresultaat, gericht op het behalen van de door de overheid vastgestelde kerndoelen (Giesbers, 1990; Toenders, 1989). De school zal daartoe door middel van invoering van het formatie-budgetstelsel een naar eigen inzicht te besteden budget krijgen. Daarmee kan ze zelf de personeelsformatie en de daarin gewenste functies vastleggen, waardoor functie- en taakdifferentiatie tot de mogelijkheden gaan behoren evenals aandacht voor loopbaanontwikkeling. Verder krijgen scholen de mogelijkheid extra financiën te verwerven door onderwijs aan derden te leveren tegen betaling. Tenslotte krijgen scholen in het voortgezet onderwijs door de introductie van de basisvorming meer keuzevrijheid wat betreft hun onderwijsaanbod. De overheid zal bij dit alles de kwaliteit bewaken door middel van toetsing achteraf, waarbij geformuleerde eindtermen een belangrijke rol zullen spelen als maatstaf (Giesbers, 1990). Overige instrumenten van kwaliteitszorg in dit kader zijn de eisen voor aanwezigheid van een systeem voor kwaliteitszorg en de verplichting van een jaarverslag.

In algemene zin kan gesteld worden dat met deze maatregelen een voorzichtig begin gemaakt wordt met de introductie van (quasi-) marktmechanismen in het onderwijs. Een gevolg van deregulering is immers dat concurrentie tussen scholen zal toenemen, vooral vanwege het feit dat de verschillen tussen scholen zullen toenemen. Dat de overheid nog verder wil gaan met de introductie van marktmechanismen moge blijken uit de discussies omtrent de wenselijkheid van bekostiging op basis van rendement en de openbaarmaking van schoolresultaten (Amelsvoort, Bosker & Brandsma, 1992). Als assumpties gelden bij deze vormen van outputcontrole dat scholen de kwaliteit van het onderwijs daadwerkelijk kunnen beïnvloeden en dat deze maatregelen scholen zullen stimuleren kwaliteitgericht te werken. Introductie van deze maatregelen zou het begin zijn van werkelijke marktmechanismen, met name omdat scholen voor een belangrijk gedeelte op basis van het gerealiseerde rendement beloofd zullen worden. In beperkte mate is overigens een vorm van rendementsbekostiging al ingevoerd in het MBO.

Het proces van deregulering en decentralisatie verkeert nog in een overgangsfase. Vooral in het MBO stuiten de instellingen bij de uitvoering van het beleid op problemen, die dreigen te leiden tot kritiek en weigering tot medewerking (Van Leijenhorst, 1993). De problemen hebben te maken met de beperkte financiële middelen, de tijdelijke regelgeving voor het verplicht in dienst nemen van (dure) wachtgelders en het ervaren contrast tussen de beperkte handelingsvrijheid en de grote verantwoordingsplicht enerzijds en de wens tot autonomieverhoging anderzijds (Visscher & Van Amelsvoort, 1993). Hoewel de overheid ernaar streeft om de beleidsruimte van instellingen te vergroten geeft zij zelf ook aan dat deze autonomie nog beperkt is en dat ze

de komende jaren onder meer via nieuwe wetgeving zal werken aan verdere autonomievergroting (Ritzen, 1993).

### VRAAGSTELLINGEN

Duidelijk is dat de nieuwe besturingsfilosofie van de overheid scholen voor een aanzienlijke overgang plaatst en een beroep doet op de veranderingscapaciteit van scholen, terwijl deze over het algemeen niet overschat mag worden (Van Gennip, 1991; Commissie Evaluatie Basisonderwijs, 1994). Van uitvoerders van onderwijsbeleid worden scholen in toenemende mate beleidsbepalers. Het is onoverkomelijk dat de toename in bestuurlijke bewegingsvrijheid om een eigen invulling vraagt. De school zal een eigen (bestuurlijk) beleid moeten gaan voeren gericht op haar eigen specifieke situatie.

Het omgaan met autonomie en het creëren van eigen beleid stelt de nodige eisen aan een organisatie. Ten eerste zal de school zich een toekomstgerichte werkwijze moeten aanmeten, tot uiting komend in strategisch management (Van de Berg, 1992; Toenders, 1989). Ze zal zelf in wisselwerking met de omgeving haar onderwijsaanbod moeten bepalen en tevens tijdig kunnen bijstellen. Ook zal ze een meer zelfstandige invulling moeten geven aan haar personeelsbeleid, waarbij afwegingen tussen belangen van personeel enerzijds en belangen van de school en leerlingen anderzijds, centraal zullen staan. Dit alles vereist een goede informatievoorziening over de omgeving, de markt, de kwaliteit van de docenten en het onderwijs.

Ten tweede zal de toename in beleidsvrijheid scholen stimuleren tot een marktgerichte houding (Gielen, 1987; Marx, 1987; Toenders, 1989). Dat wil zeggen dat de school een beleid voert waarin serieus rekening gehouden wordt met de wensen van doelgroepen. De schoolleiding dient daarbij op zijn minst veel aandacht te geven aan openheid van de school ten opzichte van de doelgroepen: de (potentiële) leerlingen en hun ouders, de instellingen die vervolgonderwijs geven, toeleverende scholen en de mogelijke werkgevers van afgestudeerde leerlingen. In concreto betekent dit dat de school een op haar eigen situatie afgestemd beleid dient te voeren, waarmee ze invulling kan geven aan het bieden van een alternatief voor onderwijs aan concurrerende instellingen (Giesbers, 1990). Om te kunnen overleven zal daarbij een marktgerichte houding noodzakelijk zijn, zeker als andere factoren als het overheidsbeleid inzake schaalvergroting, de daling van het aantal leerlingen (S.C.P., 1990) en de veranderende schoolkeuzemotieven van ouders en leerlingen (Boef-van der Meulen, Bronnenman-Helmers, Konings-van der Snoek, 1983; Pelkmans, 1983; Pieters, 1992; Versloot, 1990; De Vries, 1987) in ogenschouwing genomen worden.

Uit het bovenstaande kan geconcludeerd worden dat schoolorganisaties in toenemende mate gedwongen worden om in wisselwerking met de omgeving een eigen beleid te voeren. Een belangrijke vraag in dit kader is of en in hoeverre scholen daartoe in staat zijn en welke problemen zij daarbij ondervinden. Wanneer scholen op dergelijke punten onderzocht worden kan enig licht worden geworpen op de vragen of scholen (kunnen) voldoen aan eisen die vanuit het dereguleringsbeleid aan scholen worden gesteld.

Een andere relevante vraag vloeit voort uit het feit dat het beleid inzake deregulering een marktgerichte houding bij scholen wil bevorderen. Als zodanig heeft dit aspect geleid tot de vraag hoe het met deze marktgerichte houding binnen scholen is gesteld. In enge zin kan hier bijvoorbeeld gedacht worden aan de wenselijkheid om het onderwijs af te stemmen op relevante doelgroepen. In bredere zin aan de houding van scholen ten opzichte van (voorgenomen) maatregelen in het kader van deregulering.

Van belang zijn deze houdingen met name vanwege het feit dat in de literatuur veelal naar voren komt dat de implementatie van beleid sterk afhankelijk is van de mate waarin beleidssubjecten zich kunnen vinden in het (voorgestane) beleid (De Graaf & Hoppe, 1991). Dat hier specifiek ten aanzien van het dereguleringsbeleid mogelijk een probleem ligt, kan ontleend worden aan theorieën die het non-responsieve karakter van schoolorganisaties benadrukken (Boyd & Crowson, 1983; Mintzberg, 1979). Vanuit deze optiek is het succes van de implemen-



tatie van een beleid juist gericht op het vergroten van een marktgerichte houding zeker niet vanzelfsprekend.

Samenvattend gaat het om een tweetal vraagstellingen:

- a) komen scholen tot een marktgericht beleid zoals bedoeld door het dereguleringsbeleid en welke knelpunten ondervinden zij hierbij ?
- b) welke houdingen bestaan er in het onderwijs aangaande de wenselijkheid van het marktgericht functioneren van scholen en een overheidsbeleid gericht op het stimuleren hiervan ?

Deze twee vraagstellingen zullen beantwoord worden aan de hand van een tweetal inhoudelijk aan elkaar gerelateerde onderzoeken. In het eerste onderzoek (SVO-1124) stond de vraag naar het beleid dat scholen voeren om hun marktpositie te verstevigen centraal. Daarnaast werd de houding van betrokkenen bij de onderzochte scholen ten opzichte van marktgerichtheid gemeenten, evenals hun houding ten opzichte van een overheidsbeleid gericht op het bevorderen hiervan. Het tweede onderzoek (SVO-0524) betrof een onderzoek naar de houdingen van scholen ten opzichte van rendementsbekostiging in het algemeen voortgezet onderwijs.

## HET EERSTE ONDERZOEK: SCHOLEN EN HUN MARKTPOSITIE

### Inleiding

Zoals bij de vraagstelling naar voren is gekomen heeft het eerste onderzoek betrekking op de vraag of scholen in staat zijn een beleid te voeren zoals bedoeld door het dereguleringsbeleid van de overheid. Tevens werd de houding van betrokkenen ten opzichte van de wenselijkheid van marktgericht functioneren van scholen gepeild, evenals de houding ten opzichte van het overheidsbeleid gericht op het bevorderen hiervan.

### Onderzoeksopzet en steekproef

De onderzoeksopzet van dit onderzoek kan als een vergelijkende case-study getypeerd worden. Bij het onderzoek zijn 17 scholen verdeeld over drie schooltypen (zes scholen in het basisonderwijs, zes scholen in het algemeen voortgezet onderwijs en vijf scholen in het middelbaar beroepsonderwijs). Dit aantal is vooral bepaald door de financiële randvoorwaarden en specifieke eisen die de opdrachtgever aan het onderzoek stelde. In dit opzicht vormde dit aantal voor de onderzoekers een gegeven. Wel binnen de vrijheidsmarges van de onderzoekers lagen de vragen welke sector voor middelbaar onderwijs voor nader onderzoek aan de orde zou komen en uit welke gebieden de scholen geselecteerd zouden worden. Van belang is daarbij de notie geweest dat voor het onderzoek vooral die scholen interessant zouden zijn die zich aantoonbaar in een situatie bevinden waarin marktgerichtheid van belang is. Voor het middelbaar beroepsonderwijs heeft deze overweging er toe geleid dat het onderzoek zich alleen op de sector economie heeft gericht. De prognoses wijzen uit dat juist deze sector sterk te maken heeft met een daling in leerlingaantallen (S.C.P., 1990). Voor de selectie van de te onderzoeken gebieden heeft bovengenoemde notie geleid tot een keuze van een vijftal nodale gebieden<sup>1</sup> waar als gevolg van demografische ontwikkelingen sprake is van een daling in het aanbod van leerlingen. Daartoe is gebruik gemaakt van gegevens van het C.B.S. (1990). Bij de keuze voor deze nodale gebieden speelden naast de demografische ontwikkelingen factoren een rol als een zekere spreiding over Nederland, de urbanisatiegraad, de afstand in kilometers tussen gemeenten binnen deze nodale gebieden en de hoeveelheid scholen die in potentie in aanmerking komen om deel te nemen aan het onderzoek.

Om scholen van hetzelfde schooltype (basisscholen, AVO/VWO, MBO) zo zuiver mogelijk met elkaar te kunnen vergelijken, lag het voor de hand er zoveel mogelijk voor te zorgen dat ze qua externe factoren in vergelijkbare omstandigheden verkeren. Daarom zijn binnen elk gebied een paar grotere gemeenten uitgezocht waarvan meerdere scholen van hetzelfde typen aangeschreven werden om mee te werken aan het onderzoek zodat -demografisch gezien- de context van deze scholen hetzelfde zou zijn. De resultaten van de gevolgde steekproefprocedure staan vermeld in Tabel 1.



Tabel 1. Resultaten van de steekproef.

---

|   |
|---|
| nodaal a: drie MBO – opleidingen                          |
| nodaal b: drie basisscholen en twee scholengemeenschappen |
| nodaal c: twee basisscholen en een MBO-opleiding          |
| nodaal d: twee MAVO's, een MBO-opleiding en basisschool   |
| nodaal e: een scholengemeenschap                          |

---

Over ongeveer de helft van deze deelnemende scholen moet tenslotte opgemerkt worden dat het in die zin uitzonderlijke scholen zijn, dat zij ondanks negatieve demografische ontwikkelingen in hun omgeving er in toch in slagen hun leerlingaantal te doen stijgen. Dit is in zoverre een gevolg van de steekproefprocedure dat vooraf op basis van informatie van gemeenten en het C.B.S. (1990) gepoogd is per nodaal die scholen te selecteren waar over de jaren heen sprake is van een stijgende dan wel dalende tendens wat betreft het leerlingaantal. Het doel hiervan was verschillen op het spoor te komen tussen scholen die succesvol zijn in het op peil houden van het leerlingbestand en scholen die in dit opzicht minder succesvol zijn.

Samenvattend kan gezegd worden dat wat betreft regio, schooltype en kenmerken van scholen er steeds gezocht is naar de 'extreme' gevallen. Dat dit (mogelijkerwijs) tot een vertekening van de resultaten heeft geleid ligt voor de hand. Zo zal in een situatie waar scholen veel moeite moeten doen om hun leerlingbestand op peil te houden, men anders handelen en denken.

### Instrumentontwikkeling

Zoals hierboven gesteld kan de onderzoeksopzet als een vergelijkende case study worden beschouwd.

Bij de vormgeving van de instrumenten is gebruik gemaakt van de (school)marketing-theorie (Dulmers, 1985; 1987; Faes & Van Tilborgh, 1983; Kotler & Fox, 1985; Mahieu, 1985). In algemene zin houdt deze theorie zich bezig met de relatie schoolorganisatie – omgeving. De veronderstelling is dat naarmate de school meer kennis heeft over de eisen en wensen die leven in de omgeving en deze beter weet te vertalen in het uit te voeren beleid, de marktpositie van de school steviger zal zijn. Meer specifiek geeft deze theorie aan welke middelen een school kan gebruiken om zijn marktpositie te verbeteren. Deze middelen worden veelal onder een aantal categorieën samengevat.

In dit onderzoek is gebruik gemaakt van de indeling van Mahieu (1985). Hij onderscheidt de categorieën produkt (bijvoorbeeld onderwijsaanbod en -aanpak), promotie (bijvoorbeeld reclame en voorlichting), plaats (bijvoorbeeld bereikbaarheid van de school), prijs (bijvoorbeeld ouderbijdrage), profijt (bijvoorbeeld de waarde van het diploma en/of beroepsperspectieven) en personeel (bijvoorbeeld kwaliteit van het personeel).

Deze categorieën hebben in de eerste plaats als leidraad gediend voor de vormgeving voor het interview. Hiervoor is met betrekking tot elke categorie een serie vragen ontwikkeld. Deze vragen hadden als doel te onderzoeken welk beleid scholen op dit punt voerden en welke wijzigingen zich hierbij de afgelopen jaren ten gevolge van de in de inleiding genoemde ontwikkelingen hebben voorgedaan. Daarnaast zijn vragen geformuleerd over de context van de school.

Verder vormden deze categorieën de basis voor een output-lijst. Deze lijst bestond uit een reeks van (mogelijke) activiteiten die een school onderneemt om haar marktpositie te verbeteren. Op deze lijst kon aangegeven worden of men deze activiteiten wel of niet binnen de school ten uitvoer brengt en of hier in vergelijking met een aantal jaren geleden veranderingen waren opgetreden. Daartoe werd gevraagd of men deze activiteiten ook in 1987 ten uitvoer bracht. In dit opzicht is er dus sprake van een retrospectieve meting met mogelijk kwalijke gevolgen voor de betrouwbaarheid vandien.

Om de houdingen van betrokkenen bij de onderzochte scholen te meten is een vragenlijst ontwikkeld bestaande uit oorspronkelijk 43 items. Op basis van een factor-analyse, die als

Tabel 2. Schaalgegevens.

|                   | aantal items | Cronbach's $\alpha$ |
|-------------------|--------------|---------------------|
| responsiviteit    | 12           | 0.83                |
| deregulering      | 16           | 0.79                |
| kwaliteitsoordeel | 4            | 0.79                |

heuristische methode gebruikt is om te kijken of er een bepaalde structuur aan de items ten grondslag ligt, een betrouwbaarheidsanalyse en op basis van inhoudelijke gronden zijn uiteindelijk een drietal schalen geconstrueerd. De eerste schaal (responsiviteit) heeft betrekking op de bereidheid om het onderwijs af te stemmen op relevante groepen in de omgeving van de school zoals het bedrijfsleven, ouders en leerlingen. Naarmate de score op deze schaal hoger is, is de bereidheid om rekening te houden met relevante groepen in de omgeving groter. De tweede schaal (deregulering) meet de houding ten opzichte van deregulering, of, met andere woorden, de mate waarin de betrokkenen al dan niet voorstander zijn van een strakke regelgeving aangaande het functioneren van scholen. Een hoge score houdt in dat men het functioneren van scholen zo min mogelijk aan banden gelegd wil zien. De derde schaal (kwaliteitsoordeel) ten slotte geeft inzicht in de verwachte gevolgen van een beleid gericht op het verbeteren van de marktpositie van scholen. Een hoge score op deze schaal betekent dat men de gevolgen voor het functioneren van scholen positief inschat. De relevante gegevens over de ontwikkelde schalen staan vermeld in Tabel 2.

### Onderzoeksprocedure

Op elk van de deelnemende scholen werd een diepte-interview gehouden met een leidinggevende. Bij de basisscholen betrof dit altijd de (adjunct-) directeur, op de scholen voor AVO/VWO de rector en op de MBO-opleidingen economie de directeur van de betreffende opleiding of een lid van de centrale directie. Indien tijdens het interview bleek dat er op de school een speciale groep of persoon verantwoordelijk was voor een bepaald aspect van de marktpositie van de school, werd een nader gesprek geregeld. Daarbij werden in ieder geval de vragen voorgelegd die ook in het oorspronkelijke interview met de leidinggevende aan de orde waren geweest. In totaal zijn 25 personen geïnterviewd. Na het gesprek is de outputlijst aan de betrokkenen overhandigd en is gevraagd om relevante documenten.

De vragenlijst betreffende de houdingen is met behulp van de leidinggevende verspreid onder de sterkst betrokkenen bij de school, dat wil zeggen onder ouders die lid zijn van de ouderraad of oudervereniging, bestuursleden, docenten die lid zijn van de MR en de schoolleiding. Per school werden drie ouders, drie docenten, drie bestuursleden en een lid van de schoolleiding verzocht een vragenlijst in te vullen. Per school leverde dat maximaal tien respondenten op, in totaal 170 voor het complete scholenbestand. Van deze 170 personen hebben 122 de vragenlijst geretourneerd.

### Data-analyse

De drie bovengenoemde data-bronnen (interviews, document-analyse, outputlijst) zijn verwerkt tot een schoolspecifiek verslag. Dit verslag is ter correctie opgestuurd naar elke school. Vervolgens zijn de gecorrigeerde verslagen op kwalitatieve wijze geanalyseerd. Dit betekent dat vooral werd gezocht naar de gemeenschappelijke kenmerken van elk schooltype en de verschillen tussen de schooltypen. Om de conclusies zo goed mogelijk te kunnen verantwoorden is bij de analyses ook gebruik gemaakt van kwantificering. Vooral de outputlijst en de document-analyse (vooral van het PR-materiaal) hebben hier een rol gespeeld. Zo is voor de outputlijst het aantal activiteiten dat een school onderneemt per onderscheiden categorie geteld. Vervolgens is, voorzover vergelijkbaar, gekeken naar de verschillen tussen de verschillende schooltypen. Eenzelfde analyse is toegepast voor de analyse van de PR-materialen.

Om verschillen in houdingen tussen verschillende groepen vast te stellen is gebruik gemaakt van een eenwegs-variantie-analyse. Als factoren dienden respectievelijk het schooltype waarbij de respondenten betrokken waren en de geleding waartoe zij behoorden. Als afhankelijke variabelen dienden de gemiddelde scores op de drie geconstrueerde schalen (deregulering, responsiviteit, kwaliteitsoordeel).

## Resultaten

### het beleid van scholen

Zoals bij de vraagstelling naar voren is gekomen, is voor de marktpositie van scholen van belang dat zij in staat moeten zijn een toekomstgerichte werkwijze te ontwikkelen tot uiting komend in een op de situatie van de school gericht onderwijs- en personeelsbeleid. Van belang daarbij zijn informatie over de omgeving, de leerlingmarkt, de kwaliteit van docenten en onderwijs en aandacht voor relevante doelgroepen als (potentiële) leerlingen, toeleverende en afnemende scholen en werkgevers. Het zijn deze punten die hieronder aan bod zullen komen.

Duidelijk is dat het moeilijk is gezien de kwalitatieve aard van de gegevens deze in tabelvorm weer te geven. Niettemin kan toch enig inzicht in de materie worden verkregen door tabel 3. Deze tabel is tot stand gekomen op basis van de output-vragenlijst en geeft met name een eerste inzicht in de veranderingen die scholen de afgelopen jaren op een aantal van de bovengenoemde terreinen hebben doorgevoerd.

Uit de resultaten uit Tabel 3 kan afgeleid worden dat binnen vrijwel alle schooltypen op bovengenoemde aspecten de afgelopen jaren de nodige veranderingen zijn opgetreden. In algemene zin kan gezegd worden dat scholen een meer responsieve houding ten opzichte van hun omge-

Tabel 3. Gemiddeld aantal activiteiten per deelgebied.

|   | BO   |      | AVO  |      | MBO  |       |
|---|------|------|------|------|------|-------|
|   | 1987 | nu   | 1987 | nu   | 1987 | nu    |
| <b>promotie</b>                             |      |      |      |      |      |       |
| voorlichting                                | 1.17 | 1.33 | 6.00 | 5.83 | 6.00 | 5.80  |
| reclame / p.r.                              | 1.00 | 2.00 | 1.00 | 2.33 | 1.40 | 3.60  |
| herkenbaarheid                              | 1.33 | 3.50 | 4.00 | 5.17 | 5.40 | 8.40  |
| <b>externe contacten</b>                    |      |      |      |      |      |       |
| verbeteren contacten met ouders/leerlingen  | 3.17 | 3.67 | 5.00 | 5.50 | 2.60 | 3.20  |
| vergroten medezeggenschap ouders/leerlingen | 1.50 | 2.33 | 2.00 | 2.50 | 2.20 | 3.00  |
| contact toeleverende scholen                | -    | -    | 3.83 | 9.00 | 5.00 | 16.20 |
| contact vervolgonderwijs                    | 3.17 | 4.50 | 4.17 | 5.33 | 3.60 | 5.40  |
| contact bedrijfsleven                       | -    | -    | -    | -    | 3.60 | 5.40  |
| <b>beleid</b>                               |      |      |      |      |      |       |
| personeel                                   | 2.33 | 2.30 | 4.83 | 5.50 | 6.60 | 8.00  |
| systematiek gegevens                        | 2.00 | 2.83 | 3.67 | 4.67 | 2.80 | 4.60  |
| beleidsplannen                              | 1.50 | 3.33 | 1.17 | 3.83 | 1.40 | 4.40  |
| <b>onderwijs</b>                            |      |      |      |      |      |       |
| onderwijsaanpassing                         | 2.00 | 4.67 | 2.83 | 4.17 | 4.40 | 5.80  |
| andere veranderingen m.b.t. onderwijs       | 1.67 | 2.33 | 3.17 | 3.83 | 2.80 | 3.80  |

ving innemen, meer informatie over hun omgeving tot hun beschikking hebben en meer beleidsmatig te werk gaan.

Relevant in het kader van de responsiviteit zijn vooral de toegenomen contacten van scholen met doelgroepen als ouders, bedrijfsleven en toeleverende scholen. Vooral de toename in contacten van MBO-scholen met toeleverende scholen springt in het oog. Specifiek moet hier niet alleen gedacht worden aan het intensiveren van contacten met leerkrachten en schoolhoofden, maar vooral aan de toename van het fenomeen terugrapportage aan toeleverende scholen over de prestatie van ex-leerlingen.

Dit laatste wijst er al op dat MBO-scholen ook meer dan voorheen de beschikking hebben over relevante informatie. In de eerste plaats heeft deze betrekking op gegevens omtrent leerlingen zoals gegevens over de leerprestaties, rendement en vakken- en richtingkeuze binnen de school. In de tweede plaats weet men in toenemende mate ook waar leerlingen na het volgen van de opleiding terecht komen. Dergelijke tendensen zijn ook bij de andere schooltypen te traceren (zie tabel 3 onder het kopje systematische gegevens). Daarnaast hebben vrijwel alle scholen in het onderzoek de beschikking over gegevens inzake andere relevante omgevingsontwikkelingen. Zo hebben vrijwel alle scholen prognoses over de te verwachten leerlingaantallen en de gebieden/buurtten van afkomst tot hun beschikking. Verder zijn er bij de scholen in het middelbaar beroepsonderwijs aanzetten tot bredere omgevingsanalyses naspeurbaar waarbij systematisch geïnventariseerd wordt wat bijvoorbeeld decanen, werkgevers en/of ex-leerlingen van de opleiding vinden. Als zodanig zijn de onderzochte scholen dus in toenemende mate in staat op basis van deze kennis hun marktpositie te verstevigen en hun beleid er op af te stemmen.

Feitelijk kan hier gewezen worden op de toename van het aantal relevante beleidsplannen en van het aantal personen binnen de school dat specifiek belast is met bepaalde taken in het kader van het versterken van de marktpositie. Specifiek gaat het bij het laatste bijvoorbeeld om het institutionaliseren van contacten met het bedrijfsleven, het onderhouden van contacten met ouders, potentiële leerlingen, toeleverende scholen en de pers.

Een laatste tendens is tenslotte dat met name de scholen in het AVO en het MBO hun marktpositie proberen te verstevigen door veranderingen in te voeren in hun onderwijs. Bij scholen in het AVO worden daarbij vooral wegen gezocht om qua aanpak aan te sluiten bij een beoogde doelgroep, terwijl in het MBO veelvuldig gezocht wordt naar de mogelijkheid om nieuwe vakken/richtingen aan te bieden. Niet alleen om het onderwijs in toenemende mate af te stemmen op het bedrijfsleven, maar ook om het keuzeaanbod voor leerlingen zo groot mogelijk te maken met als doel het potentieel aan geïnteresseerde leerlingen te verhogen.

Tenslotte, geen eenduidige verschillen zijn er wat betreft het huidige functioneren op bovengenoemde aspecten gevonden tussen scholen met dalende dan wel stijgende leerlingaantallen. Wel kan geconcludeerd worden dat scholen met dalende leerlingaantallen hun activiteiten om hun marktpositie te verbeteren de afgelopen jaren meer geïntensiveerd hebben dan scholen met stijgende leerlingaantallen. Dit blijkt met name uit de resultaten op de output-lijst; de verschilscore tussen het aantal activiteiten die scholen in 1987 ondernamen en die ze nu ondernemen is steeds groter bij scholen in een bedreigde positie dan bij scholen met een constante of stijgende leerlinginstroom.

### *houdingen*

De resultaten met betrekking tot de houdingen zijn wel eenduidig in tabelvorm weer te geven. In tabel 4 staan de gemiddelde scores van de verschillende geledingen op de drie ontwikkelde schalen vermeld.

Uit deze scores blijkt dat met name ouders een responsieve houding van de school het meest waarderen en docenten het minst. Dit verschil wordt vooral bepaald door de scores op de items die betrekking hebben op de betrokkenheid van ouders bij het onderwijs. Met name ouders zijn van mening dat scholen op dit punt beter moeten functioneren, waarbij zij vaak gesteund worden door bestuursleden. Specifiek kan hier gedacht worden aan het presenteren van jaarcijfers zodat ouders beter tussen scholen kunnen kiezen en een grotere betrokkenheid van ouders bij het onderwijs. Met name docenten denken hier vaak minder positief over. Een dergelijk verschil



Tabel 4. Houdingen van verschillende betrokkenen ten opzichte van het streven naar meer responsiviteit, deregulering en de verwachte gevolgen van deregulering voor het functioneren van scholen.

|                                      | ouders<br>(n=30) | school-<br>leiding<br>(n=18) | docenten<br>(n=41) | bestuur<br>(n=31) | F-ratio | p-waarde |
|--------------------------------------|------------------|------------------------------|--------------------|-------------------|---------|----------|
| responsiviteit                       | 4.20             | 3.91                         | 3.76               | 4.06              | 4.23    | 0.00     |
| deregulering                         | 2.44             | 2.86                         | 2.56               | 2.90              | 3.85    | 0.01     |
| kwaliteitsoordeel<br>schaalrange 1-5 | 2.97             | 3.76                         | 3.17               | 3.64              | 4.88    | 0.00     |

doet zich ook voor als gekeken wordt naar de verhouding schoolleiding – docenten. Zo zien docenten wel de noodzaak en wenselijkheid van profilering in, maar willen niet dat dit ten koste gaat van hun autonomie; een grotere invloed van schoolleiding op het onderwijs en het nascholingsbeleid -als middel om te komen tot een meer duidelijke profilering van de school-, wordt door weinig docenten als positief ervaren. Op de dereguleringschaal zijn ook verschillen tussen de verschillende geledingen te onderkennen. Hier zijn ouders de meest 'conservatieve' groep en voorstander van een sterke regulering, terwijl met name de bestuursleden daar iets soepeler over denken. Zonder overigens overtuigend voorstander van dereguleringsmaatregelen te zijn.

Hetzelfde geldt min of meer voor de ingeschatte gevolgen (kwaliteitsoordeel). Ouders denken negatiever over de gevolgen dan de schoolleiders en bestuursleden. Deze laatste twee groepen zijn zelfs opmerkelijk positief over de verwachte gevolgen; door het stimuleren van marktgerichtheid verwachten zij kwaliteitsverbetering van het onderwijs.

Uit Tabel 5 kunnen de verschillen in houdingen tussen betrokkenen bij de verschillende schooltypen afgelezen worden.

Uit deze tabel blijkt dat de verschillen tussen schooltypen met name zitten in de mate waarin deregulering gewenst wordt. Binnen het basisonderwijs en AVO zijn weinig voorstanders van deregulering, terwijl deregulering meer gewaardeerd wordt binnen het MBO. Deze verschillen worden vooral verklaard door aspecten als de mening over contractonderwijs, sponsoring, mogelijkheden tot een flexibel personeelsbeleid (met name differentiatie in salaris tussen goede en/of schaarse docenten en anderen binnen de school) en de vrijheid om meer dan nu het geval is het onderwijsaanbod te bepalen. Het zijn met name betrokkenen bij het MBO die hier voorstander van zijn, terwijl betrokkenen bij het basisonderwijs hier de nodige bezwaren tegen koesteren.

Tabel 5. Houdingen binnen verschillende schooltypen ten opzichte van het streven naar responsiviteit, deregulering en de verwachte gevolgen voor het functioneren van scholen.

|                                      | Basisschool<br>(n=49) | AVO<br>(n=35) | MBO<br>(n=38) | F-ratio | p-waarde |
|--------------------------------------|-----------------------|---------------|---------------|---------|----------|
| responsiviteit                       | 3.88                  | 4.03          | 4.01          | 0.85    | 0.42     |
| deregulering                         | 2.35                  | 2.55          | 3.16          | 2.5     | 0.00     |
| kwaliteitsoordeel<br>schaalrange 1-5 | 3.19                  | 3.20          | 3.59          | 2.47    | 0.00     |



## HET TWEEDE ONDERZOEK: RENDEMENT

**Inleiding**

De overheid overweegt in het kader van het dereguleringsbeleid invoering van een vorm van rendementsbekostiging in de bovenbouw van het HAVO- en VWO als middel om scholen te stimuleren hun rendement te verbeteren. Met een financiering op basis van rendement verplaatst de aandacht van de overheid zich per definitie van de input en procesbewaking naar de output-kant van het onderwijs. Voordat deze overheidsplannen verder uitgewerkt worden is eerst het draagvlak voor deze maatregel in het onderwijsveld gepolst en zijn de mogelijke beleidsreacties van scholen geïnventariseerd. Dit onderzoek naar het oordeel van de schoolleiders over rendementsbekostiging wordt hieronder kort beschreven.

**Selectie van scholen en respons**

Voor de selectie van scholen is gebruik gemaakt van de integrale leerlingtellingen. Alle HAVO- en VWO-scholen zijn ingedeeld op de kenmerken denominatie (openbaar of bijzonder), schoolsoort (HAVO of VWO), urbanisatiegraad ( $\leq 100.000$  inwoners of  $> 100.000$  inwoners), schooltype (HAVO/VWO-scholengemeenschap of anderszins samengestelde scholengemeenschap) en rendement (laag, middelmatig of hoog rendement). Het rendement is berekend door per school de ratio van percentage bevorderden in de bovenbouw plus het percentage geslaagden op het totaal aantal leerlingen in de bovenbouw te berekenen. Het rendementscriterium is van belang omdat verwacht kan worden dat de reactie van scholen op rendementsbekostiging afhankelijk is van het eigen, huidige rendement. De combinatie van de vijf selectiecriteria resulteert in een tabel met 48 cellen ( $2 \times 2 \times 2 \times 2 \times 3$ ). Er is gestreefd naar een bezetting van één school per cel. In eerste instantie werd een school benaderd. Bij weigering tot deelname werd een nieuwe school benaderd. Naast deze scholen zijn acht scholen met een hoge concentratie culturele minderheden benaderd. Het bleek niet mogelijk voor iedere combinatie een school te vinden. Zo komen bijzondere scholen met een zeer laag rendement nauwelijks voor. In totaal zijn 94 scholen benaderd, waarvan er 53 hun medewerking hebben verleend. Dit betekent een responspercentage van 56%. Hoewel de selectie niet representatief kan worden genoemd voor de populatie HAVO- en VWO-scholen in Nederland, heeft de gekozen gestratificeerde benadering het voordeel dat zinvol vergelijkingen kunnen worden gemaakt tussen sub-groepen van scholen.

**Onderzoeksprocedure**

De gegevens zijn verzameld met behulp van een schriftelijke vragenlijst en een interview. De schriftelijke vragenlijst diende voor het verzamelen van gegevens over leerlingaantallen, slaagen- en doublurepercentages en over het huidige beleid van scholen inzake rendementsbewaking. Gegevens over het huidige beleid zijn verzameld om een vergelijking te kunnen maken met de voorgenoemde maatregelen indien een vorm van rendementsbekostiging zou worden ingevoerd. In de interviews met de correctoren is ingegaan op het oordeel van over rendementsbekostiging, alsmede op beleidsmaatregelen die de school zou treffen bij invoering van rendementsbekostiging. Om de respondent in de gelegenheid te stellen intern overleg te voeren over de mogelijke reacties van de school op invoering van rendementsbekostiging zijn de interviewonderwerpen vooraf aan de respondent meegedeeld. De respondenten zijn geconfronteerd met de financiële consequenties voor de eigen school van invoering van vier optionele bekostigingsmodellen, met de vraag daaraan gekoppeld wat zij zouden doen bij invoering van één van deze modellen. Een probleem bij deze wijze van werken is het gevaar van sociale wenselijkheid. Met andere woorden, durven respondenten ongewenste maatregelen te berde te brengen. Om dit te voorkomen zijn een aantal maatregelen genomen. In de eerste plaats is een glijdene schaal geconstrueerd (zie volgende paragraaf) met daarin opgenomen vele gradaties van niet wenselijke antwoorden. Daarnaast is de vraagstelling aangepast door niet te vragen of de school de maatregel zou nemen, maar hoe groot men de kans op invoering van een dergelijke maatregel inschat. Vervolgens is gebruikt gemaakt van 'dubbele bevraging'. Aan respondenten is eerst op algemeen niveau en vervolgens op concreet niveau dezelfde vraag gesteld. Hun antwoorden bleken con-

sistent. Hoewel hiermee de kans op sociale wenselijkheid niet volledig weggenomen is, mag op basis van de literatuur (Swanborn, 1987) wel verwacht worden dat de kans hierop sterk vermindert is.

Tenslotte is de kans op sociaal wenselijke antwoorden ondervangen door de helft van de respondenten eerst met het meest ingrijpende bekostigingsmodel (50%-model) te confronteren en vervolgens met de minder ingrijpende modellen. De andere helft respondenten werd eerst met het minst ingrijpende bekostigingsmodel (5%-model) geconfronteerd en vervolgens met de meer ingrijpende. De antwoordpatronen van beide groepen schoolleiders bleken niet te verschillen.

### Instrumentontwikkeling en data-analyse

Het voorgenomen beleid van scholen is getypeerd met de termen strategisch beleid en strategisch gedrag. *Strategisch beleid* betreft de beleidsvoering die enerzijds gericht is op de verbetering van het rendement van de eigen school en anderzijds bijdraagt aan de realisatie van de overheidsdoelen, namelijk het in samenhang vergroten van het kwalitatief en het kwantitatief rendement van de bovenbouw van het voortgezet onderwijs. Een belangrijk criterium is dat deze maatregelen zodanig ten goede komen aan de leerlingen dat uiteindelijk het niveau van de leerprestaties wordt verhoogd. Strategisch beleid verwijst aldus naar maatregelen die leiden tot de gewenste effecten voor het voortgezet onderwijs. *Strategisch gedrag* kan daarentegen meer getypeerd worden als ontwikkelingsgedrag, gericht op budgetmaximalisatie, waarbij het belang van de leerling niet voorop staat. Het betreft de beleidsvoering die weliswaar gericht is op rendementsverbetering van de eigen school, doch niet bijdraagt aan de realisatie van de overheidsdoelen ten aanzien van de verbetering van zowel het kwantitatieve als het kwalitatieve rendement van de bovenbouw van het voortgezet onderwijs. Strategisch gedrag verwijst aldus naar beleidsmaatregelen die leiden tot ongewenste neveneffecten. We kunnen hier te maken hebben met ad hoc beslissingen als reactie op een wellicht tegenvallende bekostiging. Voorbeelden hiervan zijn ingrepen in het schoolonderzoek (verlagen van niveau door vereenvoudiging van opgaven, het laten leren voor de toets of het achteraf aanpassen van de norm) of een tijdelijke verscherping van de doorstroming van de onderbouw naar de bovenbouw (bijstellen overgangsnormen, 'bespreekzone' voor twijfelgevallen versmallen of verzwaren proefwerken). Het kan ook een weloverwogen lange-termijnbeleid, gericht op regulatie van de instroom betreffen. Om de mogelijke beleidsreacties op één van de optionele bekostigingsmodellen te kunnen weergeven in termen van strategisch beleid en strategisch gedrag zijn twee schalen geconstrueerd. De constructie van de eerste schaal leverde de volgende statistische gegevens op (Tabel 6).

Tabel 6. Ontvouwingsgegevens.

|                       |  | H(I) | T(H) |
|-----------------------|--|------|------|
| strategisch<br>beleid | ▲  |      |      |
|                       | d – optimaliseren functioneren organisatie   | .62  | 3.04 |
|                       | a – optimaliseren leerlingbegeleiding        | .41  | 4.21 |
|                       | e – stimuleren (vak)inhoud. verbeteringen    | .42  | 5.34 |
|                       | g – veranderen schoolwerkplan                | .40  | 5.51 |
|                       | c – optimal. selectie- en determinatiebeleid | .31  | 4.29 |
|                       | j – verscherpen aannamebeleid                | .34  | 3.73 |
|                       | i – verscherpen dublerebeleid                | .39  | 3.20 |
|                       | h – verscherpen doorstroombeleid             | .22  | 1.39 |
| strategisch<br>gedrag | ▼  |      |      |
| Totale schaal         |  | .38  | 4.79 |

De H(I)-coëfficiënt geeft aan in hoeverre een maatregel schaalbaar is. De coëfficiënt loopt van 0 (niet schaalbaar) tot 1 (perfect schaalbaar). In het algemeen duidt een waarde van .30 op voldoende en .40 op redelijke schaalbaarheid. Daarnaast is de T-waarde vermeld, die wanneer groter dan 1.96, aangeeft dat de geschatte H(I)-waarde statistisch van 0 (niet schaalbaar) afwijkt met 5% kans op een verkeerde beslissing (in werkelijkheid was de maatregel niet schaalbaar). De totale schaal is voldoende tot redelijk homogeen; het verwijderen van item h (verscherpen doorstroombeleid) leidt niet tot een verbetering, zodat dit item ondanks de matige itemkarakteristiek gehandhaafd blijft.

Naast het aangeven van beleidsprioriteiten is scholen gevraagd naar concrete maatregelen die genomen zouden worden bij invoering van rendementsbekostiging. Hiervoor is een Likert-schaal ontwikkeld voor de dimensie van strategisch beleid tot strategisch gedrag. De betrouwbaarheid van deze schaal is goed (Cronbach's  $\alpha = 0,78$ ). De verzamelde data voor het totale onderzoek zijn geanalyseerd met behulp van rechte tellingen, variantie-analyses, correlatietechnieken en schaaltechnieken.

## Resultaten

### *huidig beleid inzake bewaking van het rendement*

Scholen verschillen in hun huidige beleidsvoering ten aanzien van de aanname, de selectie en determinatie en de leerlingbegeleiding en deze verschillen hangen samen met hun rendement. HAVO-scholen met een hoog rendement baseren de aanname van de leerlingen in de onderbouw en in de bovenbouw meer op gebleken prestaties dan op criteria als voorkeur van leerlingen en/of ouders. HAVO-scholen met een laag rendement blijken bij de advisering over de vakkenpakketkeus risico's te vermijden; er wordt liever te laag dan te hoog geadviseerd. Bij VWO-scholen treden deze verbanden niet op. HAVO-scholen met een hoog rendement hebben verhoudingsgewijs minder verschillende voorzieningen voor leerlingbegeleiding, maar besteden gemiddeld wel meer uren aan die begeleiding per leerling. VWO-scholen met een hoog rendement blijken juist meer verschillende voorzieningen voor leerlingbegeleiding te hebben terwijl er geen significante verschillen qua tijdsbesteding optreden.

### *oordeel*

Er zijn geen verschillen aangetroffen tussen scholen in hun houdingen ten opzichte van rendementsbekostiging. Scholen oordelen over het algemeen tamelijk negatief over de invoering van een dergelijk bekostigingsstelsel, hetgeen geïllustreerd wordt in Tabel 7.

Respectievelijk 64,2% en 60,4% van de scholen is tegen rendementsbekostiging in het algemeen, dan wel tegen rendementsbekostiging voor alleen de bovenbouw. Er worden met name negatieve gevolgen voor de leerlingen gevreesd. Zwakke leerlingen en laatbloeiers zullen minder kansen krijgen een diploma te halen, zo oordelen de schoolleiders. Het minst negatief wordt

Tabel 7. Oordeel over rendementsbekostiging.

|   | negatief<br>% | neutraal<br>% | positief<br>% | N  |
|---|---------------|---------------|---------------|----|
| oordeel over rendements-<br>bekostiging in het algemeen                   | 64.2          | 20.8          | 15.0          | 53 |
| oordeel over rendements-<br>bekostiging voor de bovenbouw                 | 60.4          | 24.5          | 15.1          | 53 |
| oordeel over rendements-<br>bekostiging met bonus voor<br>probleemgroepen | 46.5          | 25.8          | 27.6          | 48 |

geoordeeld over een vorm van rendementsbekostiging waarbij een bonus voor geslaagden uit achterstandsgroepen, zoals culturele minderheden, wordt uitgekeerd (ruim 27% oordeelt positief en circa 26% heeft een neutraal oordeel). Het begeleiden van leerlingen in achterstandssituaties vraagt een extra prestatie van scholen en deze mag beloofd worden, zo vindt men.

Schoolleiders gaven ook hun oordeel over mogelijke andere consequenties voor de school van invoering van rendementsbekostiging. Met name het pedagogisch/didactisch klimaat en de omgang met leerlingen zullen volgens hen leiden onder deze maatregel. Daarnaast komen volgens de helft van de schoolleiders de vele extra-curriculaire activiteiten, zoals meerdaagse excursies en creatieve activiteiten onder druk te staan. Positieve consequenties zou de maatregel kunnen hebben voor de professionaliteit van de organisatie, de efficiëntie van het werk en de marketing en de Public Relations.

Om enig idee te krijgen van het rendementsbewustzijn van de schoolleider is hierover een aantal vragen gesteld. Op de vraag of men de slaagpercentages wel eens vergelijkt, antwoordt 96% met 'ja' en 4% met 'nee' (N=47). Op de vraag of men het landelijk gemiddelde slaagpercentage 1989/1990 kent antwoordt 75% van de HAVO-scholen (N=24) en 87% van de VWO-scholen (N=23) bevestigend. Vervolgens is aan de schoolleiders eerst het rendementscijfer van de school gepresenteerd en daarna gevraagd aan te geven hoe men denkt dat het rendement van de school zich tot het landelijk gemiddelde verhoudt. De antwoorden van de respondenten geven aan dat het rendementsbewustzijn niet erg groot is; slechts 50% van de betrokkenen heeft een juist inzicht van het rendement van de school in verhouding tot het landelijk gemiddelde. Er is verder gekeken of dit rendementsbewustzijn samenhangt met de aard van het oordeel over rendementsbekostiging. Dit blijkt niet het geval te zijn.

#### *voorgenomen beleidsmaatregelen*

De beleidsreacties van scholen zijn onderscheiden naar algemene en concrete maatregelen en kunnen worden afgebeeld op de dimensie van strategisch gedrag naar strategisch beleid, respectievelijk met behulp van de ontvouwingschaal en de Likertschaal. De schaa scores voor de algemene beleidsmaatregelen lopen van 0 (strategisch gedrag) tot 7 (strategisch beleid). De schaa scores voor de concrete maatregelen lopen van 1 (strategisch gedrag) tot 5 (strategisch beleid). Er zijn variantie-analyses uitgevoerd van de scores op de ontvouwingsschaal (algemeen beleid) en van de scores op de Likertschaal (concrete maatregelen) als afhankelijke variabelen met de factoren denominatie (openbaar of bijzonder), urbanisatiegraad (<100.000 inwoners of >100.000 inwoners), rendement (laag, middelmatig of hoog rendement), schooltype (HAVO/VWO-scholengemeenschap of anderszins samengestelde scholengemeenschap), onderwijssoort (HAVO of VWOs) en oordeel over rendementsbekostiging (negatief, neutraal of positief). Vier analyses leverden significante verbanden op. De gevonden significante verbanden staan vermeld in Tabel 8.

Scholen met een hoog rendement zijn relatief meer geneigd te kiezen voor algemene beleidsmaatregelen die tot het strategisch beleid kunnen worden gerekend, dan scholen met een laag rendement. Hetzelfde patroon blijkt uit de positie van de scholen op de Likertschaal, die is geconstrueerd op basis van de concrete maatregelen. Bijzondere scholen, scholen met een hoog rendement en de HAVO-scholen zijn iets meer geneigd tot strategisch beleid dan de openbare scholen, de scholen met een laag rendement en de VWO-scholen. Echter, scholen neigen over het algemeen meer naar strategisch beleid dan naar strategisch gedrag. Extreme vormen van het algemeen meer naar strategisch beleid van de bewaking van het kwalitatief niveau (bijvoorbeeld strategisch gedrag, zoals versoepeling van de schoolonderzoeken), worden amper of niet als beleidsoplossing genoemd. Het probleem van de sociale wenselijkheid bij het geven van antwoorden is waarschijnlijk, ondanks de aanpassingen in de vraagstelling, niet geheel opgelost. Vermeld dient echter te worden dat scholen beseften dat hun antwoorden van belang zijn voor de verdere beleidsvorming rond rendementsbekostiging. Gezien hun negatieve oordeel over deze vorm van outputcontrole zou men ook kunnen verwachten dat scholen juist extreme vormen van strategisch gedrag noemen, teneinde de voorgenomen maatregelen inzake rendementsbekostiging af



Tabel 8. Strategisch beleid en strategisch gedrag op ontvouwingsschaal (algemene maatregelen) en Likert-schaal (concrete maatregelen) naar denominatie, rendement en onderwijssoort.

|                             | Algemene maatregelen |      | Concrete maatregelen |      | N  |
|-----------------------------|----------------------|------|----------------------|------|----|
|                             | Gem.                 | SD   | Gem.                 | SD   |    |
| Denominatie <sup>1</sup>    |                      |      |                      |      |    |
| Bijzonder                   |                      |      | 3.17                 | 0.47 | 23 |
| Openbaar                    |                      |      | 2.97                 | 0.41 | 30 |
| Rendement <sup>2</sup>      |                      |      |                      |      |    |
| laag                        | 3.82                 | 1.33 | 2.90                 | 0.43 | 17 |
| midden                      | 4.23                 | 1.60 | 2.98                 | 0.42 | 17 |
| hoog                        | 5.18                 | 1.69 | 3.26                 | 0.41 | 19 |
| Onderwijssoort <sup>3</sup> |                      |      |                      |      |    |
| HAVO                        |                      |      | 3.17                 | 0.36 | 28 |
| VWO                         |                      |      | 2.93                 | 0.50 | 25 |

<sup>1</sup>  $p = 0.10$ ;  $Eta^2 = 0.05$ <sup>2</sup>  $p = 0.03$ ;  $Eta^2 = 0.13$ <sup>3</sup>  $p = 0.04$ ;  $Eta^2 = 0.08$ 

te wentelen. Een empirische indicatie voor het feit dat extreme vormen van strategisch gedrag niet direct te verwachten zijn, betreft de in eerder onderzoek (Bosker, 1991) aangetroffen (licht) positieve correlatie tussen kwalitatief en kwantitatief rendement. Hogere prestaties gaan in de praktijk samen met minder zittenblijven, minder drop-out en minder afstroom.

Wat betreft de concrete maatregelen, blijkt dat meer dan 50% van de scholen het zeer waarschijnlijk acht dat bij invoering van een vorm van rendementsbekostiging de volgende maatregelen zullen worden getroffen:

1. creëren van meer bijscholingsmogelijkheden voor docenten;
2. invoeren dan wel automatiseren van een leerlingvolgsysteem;
3. stoppen van meer tijd per leerling in studiebegeleiding;
4. voorzichtiger adviseren bij twijfel over schooltypekeuze;
5. vaker doorstroming van HAVO-3 naar het MBO adviseren;
6. sneller de afstroom naar een lager schooltype adviseren.

#### *veranderingen in beleid*

De maatregelen die de scholen zeggen te gaan nemen bij invoering van een vorm van rendementsbekostiging, zijn vergeleken met het huidige beleid dat de school voert, teneinde inzicht te krijgen in de veranderingen in beleid. Tabel 9 toont dat scholen meer geneigd zijn strategische beleidsmaatregelen ter verbetering van hun rendement te treffen naarmate het organisatie- en onderwijskundige beleid op de corresponderende terreinen in mindere mate ontwikkeld is. Dit komt meer specifiek tot uitdrukking bij aspecten van de leerlingbegeleiding, de leerlingbeoordeling en de schoolorganisatie. Hoe minder faciliteiten men in de onderbouw voor de leerlingbegeleiding kent, des te eerder kiest men voor beleidsterreinen en concrete maatregelen die kunnen worden gerekend tot strategisch beleid. Verder geldt dat scholen verhoudingsgewijs meer kiezen voor de strategische beleidsterreinen, naarmate ze in de huidige situatie minder maatregelen kennen om ongeoorloofd verzuim van leerlingen tegen te gaan en naarmate hun beoordelingssysteem minder systematisch is.



Tabel 9. Samenhang tussen aspecten van huidige en toekomstige schoolsituatie.

| Huidig beleid                      | Voorgenomen<br>algemene maatregelen | Voorgenomen<br>concrete maatregelen |
|------------------------------------|-------------------------------------|-------------------------------------|
| <b>AANNAME</b>                     |                                     |                                     |
| toelatingscriteria onderbouw       | -.10                                | .16                                 |
| toelatingscriteria bovenbouw HAVO  | -.18                                | -.01                                |
| toelatingscriteria bovenbouw VWO   | -.17                                | -.27                                |
| <b>LEERLINGBEGELEIDING</b>         |                                     |                                     |
| voorzieningen onderbouw            | -.27                                | -.23                                |
| voorzieningen bovenbouw            | -.09                                | -.01                                |
| aantal uren onderbouw              | .02                                 | .16                                 |
| aantal uren bovenbouw              | -.02                                | .12                                 |
| <b>LEERLINGBEOORDELING</b>         |                                     |                                     |
| beoordelingssysteem                | -.25                                | -.08                                |
| <b>DOORSTROMING</b>                |                                     |                                     |
| criteria advies na brugklas        | -.13                                | -.15                                |
| <b>SCHOOLORGANISATIE</b>           |                                     |                                     |
| aantal functioneringsgesprekken    | -.03                                | -.06                                |
| verzuimmaatregelen                 | -.33                                | -.15                                |
| <b>SELECTIE/DETERMINATIE</b>       |                                     |                                     |
| advisering vakkenpakketkeuze       | .01                                 | .13                                 |
| <b>BESTUURLIJKE BETROKKENHEID:</b> |                                     |                                     |
| organisatiebeleid                  | .01                                 | .11                                 |
| onderwijskundig beleid             | -.13                                | -.13                                |
| <b>PRESTATIEGERICHTHEID</b>        |                                     |                                     |
|                                    | -.10                                | .02                                 |

vet gedrukte correlaties zijn significant bij  $\alpha < .10$

cursief gedrukte correlaties zijn significant bij  $\alpha < .20$

## CONCLUSIES

In dit artikel staan twee vragen centraal. De eerste vraag luidt of scholen komen tot een marktgericht beleid zoals bedoeld door het dereguleringsbeleid en welke problemen zij daarbij ondervinden. De tweede vraag heeft betrekking op de houdingen aangaande de wenselijkheid van marktgericht functioneren van scholen en een overheidsbeleid gericht op het stimuleren hiervan. Deze vragen zullen achtereenvolgens besproken worden.

Vanuit het eerste onderzoek gesteld kan geconcludeerd worden dat scholen door het dereguleringsbeleid inderdaad gestimuleerd worden beleidsmatiger te werken. Dit gegeven kan aan de hand van een aantal ontwikkelingen nader geïllustreerd worden. In het basisonderwijs heeft het formatiebudgetsysteem er volgens de betrokkenen toe geleid dat scholen steeds meer beschikken over prognoses over de verwachte leerlinginstroom. In de onderzochte VWO-scholen is men druk bezig met het vormgeven van de vrije ruimte in het onderwijsaanbod. In de onderzochte MBO-scholen heeft volgens de betrokkenen de lumpsum-systematiek geleid tot het creëren van nieuwe opleidingen, flexibilisering in de doorstroommogelijkheden en het creëren van nieuwe functies dan wel het differentiëren van functies. Een voorbeeld hiervan is dat alle bezochte MBO-scholen in meer of mindere mate taakuren hebben vrijgemaakt om een beter

beleid aangaande aspecten die van groot belang zijn voor de marktpositie (bijvoorbeeld leerlingwerving) mogelijk te maken. Tenslotte valt bij alle schooltypen te wijzen op de toename van het aantal contacten met specifieke doelgroepen.

Deze conclusie wil niet zeggen dat scholen geen problemen ondervinden. In algemene zin kan gesteld worden dat het proces van deregulering en decentralisatie nog in een overgangsfase verkeert. Dit levert volgens de betrokkenen een aantal problemen op. In de eerste plaats is er de spanning tussen het stimuleren van de marktgerichtheid van scholen aan de ene kant en aan de andere kant de belemmeringen die scholen ondervinden om aan deze marktgerichtheid handen en voeten te geven. Een voorbeeld hiervan is dat de MBO-scholen in dit onderzoek willen concurreren met betrekking tot het onderwijsaanbod, maar daarbij in de praktijk allerlei belemmeringen ondervinden, met name omdat bij de verdeling van vakken/richtingen over scholen vanuit de overheid overwegingen ten aanzien van de denominatie een belangrijke rol spelen. Een ander problematisch punt in deze is dat, ondanks het formatiebudgetstelsel, de mogelijkheden om te komen tot een flexibele personeelsbeleid nog altijd als gering ingeschat worden. Met name de verplichting om bij vacatures wachtgelders in dienst te nemen is hier de boosdoener.

In de tweede plaats vloeien er knelpunten voort uit het feit dat er in veel scholen sprake is van een cultuur die niet toegesneden is op de veranderende omstandigheden. Uit de interviews met de schoolleiders en uit de vragenlijst blijkt dat de houding van docenten in veel gevallen niet in de pas loopt met een marktgerichte filosofie. Niet alleen zijn zij weinig geneigd om deel te nemen aan de beleidsvoering, maar tevens houden zij veelal vast aan hun opvattingen aangaande autonomie. Duidelijk is dat in een dergelijke constellatie de vorming en uitvoering van instellingsbeleid niet eenvoudig is.

De laatste conclusies hebben betrekking op de houdingen binnen de onderzochte scholen. Met betrekking tot de responsiviteit binnen scholen kan geconcludeerd worden dat deze in het algemeen groter is dan de theorievorming op dit terrein doet vermoeden. Zo levert een sterke gerichtheid van scholen in het beroepsonderwijs op het bedrijfsleven voor de meeste respondenten geen problemen op. In dit opzicht zijn opvattingen die vasthouden aan de 'pedagogische provincie' blijkbaar aan het verdwijnen. Problematisch is wel de verhouding tot ouders. Niet alleen schatten betrokkenen de deskundigheid van ouders met betrekking tot het onderwijs laag in, maar tevens wijzen veel docenten een grote betrokkenheid van ouders bij het onderwijs af. Voorzover het dereguleringsbeleid als doel heeft de afstemming tussen het onderwijsaanbod van scholen en de wensen van onderwijsconsumenten te vergroten, moeten de kansen hierop als klein ingeschat worden.

Wat betreft de houdingen aangaande het overheidsbeleid kan geconcludeerd worden dat er tussen de verschillende schooltypen grote verschillen bestaan over de wenselijkheid van deregulering. In het basisonderwijs is men in grote lijnen geen sterk voorstander hiervan, in het MBO wel. Dit zal betekenen dat maatregelen in het kader van het dereguleringsbeleid in de verschillende schooltypen verschillend zullen uitwerken. Een voorbeeld hiervan is de mogelijkheid om te differentiëren naar salaris. In het basisonderwijs is hier weinig behoefte aan, in het MBO veel meer. Dat dit consequenties heeft voor de uitwerking van dit beleid in de verschillende schooltypen ligt voor de hand. Getuige de grote verschillen tussen schooltypen ligt het evenzeer in de rede om bij een nadere uitwerking van het dereguleringsbeleid met deze verschillen in oriëntatie rekening te houden.

Specifieke inzichten omtrent de houdingen kunnen ontleend worden aan het onderzoek naar de wenselijkheid van rendementsbekostiging. Een belangrijke conclusie uit dit onderzoek is dat scholen afwijzend staan tegenover de invoering van een dergelijk systeem. Met name de notie dat vooral kenmerken van de leerlinginstroom bepalend zijn voor het rendement van de school speelt hier een belangrijke rol. Ondanks dit negatieve oordeel over de voorgestelde vormen van rendementsbekostiging hebben scholen echter wel vertrouwen in de mogelijkheden om hun rendement te manipuleren. Uit de genoemde te verwachten beleidsreacties van HAVO- en VWO scholen blijkt dat scholen gevoelig zijn voor bekostigingsprikkelers en bij de invoering van een dergelijk systeem hun gedrag zullen aanpassen zonder te vervallen in extreme vormen van strategisch gedrag. Zoals gezegd kan sociale wenselijkheid hier een rol spelen. Daar staat echter

tegenover dat scholen in het MBO als reactie op rendementsbekostiging inderdaad wenselijk geachte maatregelen nemen gericht op optimalisering van de leerlingbegeleiding, -beoordeling en -selectie (Brandsma & Thijssen, 1994; Pelkmans & De Vries, 1991). Aangezien er geen reden is om aan te nemen dat de onderzochte scholen op dit punt wezenlijk van MBO-scholen verschillen, mag aangenomen worden dat scholen inderdaad conform de boven weergegeven reactie zullen handelen.

Een probleem inzake het verbinden van financiële consequenties aan een systeem van outputcontrole is echter wel dat het een technisch complexe maatregel is. Te wijzen valt hier op de noodzaak om met uitermate verfijnde scenario's te werken en op het bestaan van vele valkuilen die bij een gebrekkige implementatie kunnen optreden. In dit opzicht verdient het aanbeveling om eerst een wat terughoudender middel in te zetten. Te denken valt hier aan de verplichting tot (openbare) jaarverslagen. Getuige het eerste onderzoek bestaan tegen een dergelijke maatregel in het onderwijs veel minder bezwaren. Bovendien wordt zo aan een sterk bij ouders levende wens tegemoet gekomen worden.

## NOOT

1. Een nodaal is een gebied waarvan de bewoners in het maatschappelijk verkeer, waaronder de deelneming van onderwijs, geïntereerd zijn op bepaalde kernen in dit gebied. Een onderwijskern is een gemeente waar alle vormen van onderwijs die direct aansluiten op het lager onderwijs, aanwezig zijn (C.B.S., 1990).

## LITERATUUR

- Amelvoort, H.W.C.H. van, Bosker, R.J., & Brandsma, T.F. (1992). *Outputfinanciering: pro's en contra's*. Enschede: Universiteit Twente, OCTO.
- Berg, R. van den (1992). Baas in eigen school; innovatie als een strategisch vraagstuk binnen school. *Meso*, 61, 2-5.
- Boef-van der Meulen, S., Bronneman-Helmers, R. & Konings-van der Snoek, M. (1983). *Schoolkeuzemotieven en meningen over onderwijs*. Rijswijk: Sociaal en Cultureel Planbureau.
- Boyd, W.L. & Crowson, W.L. (1985). The changing conception and practice of public school school administration. *Review of Research in Education*, 9, 311-373.
- Bosker, R.J., Guldemond, H., Hofman, R.H., & Hofman, W.H.A. (1988). *Kwaliteit in het voortgezet onderwijs*. Groningen: RION.
- Bosker, R.J. (1991). De keuze van een onderwijsindicator; criterium, meetniveau en statistisch model. *Tijdschrift voor Onderwijswetenschappen*, 21, 191-205.
- Brandsma, T.F. & Thijssen, J.A. (1994). *Flexibiliteit in de kwalificatievoorziening. Een onderzoek naar organisatorische en onderwijskundige kenmerken van een flexibiliseringsstrategie in het kader van de herstructurering van het beroepsonderwijs*. Enschede: OCTO (in voorbereiding).
- C.B.S. (1989). *Loop van de bevolking per gemeente*. 's Gravenhage: SDU/CBS publicaties.
- C.B.S. (1990). *Het voortgezet onderwijs regionaal bezien ('88-'89)*. 's Gravenhage: SDU/CBS publicaties.
- Chubb, J.E. & Moe, T.M. (1990). *Politics, markets and American schools*. Washington D.C.: Brookings Institution.
- Commissie Evaluatie Basisonderwijs (1994). *Zicht op kwaliteit*. 's Gravenhage: SDU.
- Dulmers, R.J. (1987). Het belang van onderwijsmarketing. In R.J. Dulmers & C. Dietvorst. *Marketing voor scholen*. Alphen aan de Rijn/Brussel: Samson Uitgeverij.
- Faes, W. & Tilborgh, C. van (1983). *Marketing van diensten, een handboek voor praktisch management*. Deventer: Kluwer.
- Gennip, J. van (1991). *Veranderingscapaciteiten van basisscholen*. Nijmegen: ITS.
- Geurts, J. (1989). *Van Niemandland naar beroepsstructuur. Een studie over de aansluiting tussen onderwijs en arbeid op het niveau van aankomend vakmanschap* (proefschrift), Nijmegen: ITS.
- Gielen, K.J.M. (1987). Autonomie van scholen en schoolmanagement-ontwikkeling. In L. Genemans, (ed.). *Autonomie van scholen en deregulering: verslag van een studieconferentie, Nijmegen, 6 februari 1987* (pp. 99-107) Nijmegen: ITS.

- Giesbers, J.H.G.I. (1987). Autonomie en deregulering; ook een kwestie van kunnen en/of willen. In L. Genemans, (ed.). *Autonomie van scholen en deregulering: verslag van een studieconferentie, Nijmegen, 6 februari 1987*. (pp. 153-169) Nijmegen: ITS.
- Giesbers, J.H.G.I. (1990). Autonomievergroting: lust of last? In J.H.G.I. Giesbers & P.M.E.F. Adriaens [et al.]. *De noodzaak tot verantwoording: beschouwingen over autonomie en evaluatie in het onderwijs*. (pp. 8-20) Algemeen Bureau Katholiek Onderwijs, Thema 1. Den Haag.
- Graaf, H. v.d. & Hoppe, R. (1991). *Beleid en politiek*. Muiderberg: Coutinho.
- Hoof, J.J. van (1986). *Aansluiting tussen onderwijs en arbeidsmarkt: recente ontwikkelingen en uitgangspunten voor beleid*. 's Gravenhage: Organisatie voor Strategisch Arbeidsmarktonderzoek.
- Jakobs, J.J. (1987) Opvattingen in het onderwijsveld over grotere autonomie van scholen. In L. Genemans (ed.). *Autonomie van scholen en deregulering: verslag van een studieconferentie, Nijmegen, 6 februari 1987*. (pp. 81-91) Nijmegen: ITS.
- Kotler, P. & Fox, K.F.A. (1985). *Strategic marketing for educational institutions*. Englewood Cliff/ New Jersey: Prentice Hall.
- Leijenhorst, H. van (1993). *Kamerstuk 19790, nr 55*. (Verslag van een mondeling overleg). 's Gravenhage: SDU.
- Lieberman, M. (1989). *Privatization and Educational Choice*. New York: St. Martin's Press.
- Mahieu, P. (1985). *Schoolbank te koop: over schoolreclame en onderwijsmarketing*. Leuven:Acco.
- Marx, E.C.H. (1987). Vermogen van scholen tot het voeren van bestuurlijk beleid. In L. Genemans (ed.). *Autonomie van scholen en deregulering: verslag van een studieconferentie, Nijmegen, 6 februari 1987*. (pp. 7-27) Nijmegen: ITS.
- Ministerie van Onderwijs en Wetenschappen (1988). *De school op weg naar 2000*. Zoetermeer: Ministerie van Onderwijs en Wetenschappen.
- Mintzberg, H. (1979). *The structuring of organizations*. Englewood Cliffs: Prentice Hall.
- Pelkmans, A.H.W.M., van Kessel, N., van Viersen, D., Hahn, A., & Buurke, W. (1983). *Wensen omtrent scholen en de onderwijsplanning; vijf deelonderzoeken over kwalitatieve onderwijsbehoeften, 'vertalend onderwijs' en de planning van het voortgezet onderwijs*. Nijmegen: ITS.
- Pelkmans, A.H.W.M. & Vries, B. de (1991). *Modulering/flexibilisering in de tweedefase SVM*. Nijmegen: ITS. Voordracht gehouden op de conferentie 'Flexibiliteit in de kwalificatievoorziening', in het kader van het netwerk Beleidsonderzoekers Beroeps onderwijs en Volwasseneducatie, Utrecht, 24 oktober 1991.
- Pieters, L. (1992). Trends in schoolkeuze na de basisschool. *Meso*, 64, 25 -30.
- Ritzen, J.M.M. (1993). *Kamerstuk 19790, nr 54*. (Brief van de Minister van Onderwijs en Wetenschappen). 's Gravenhage: SDU.
- Schuur, W.H. van (1984). *Structure in political beliefs, a new unfolding model application to European party activists*. Amsterdam: CT Press.
- Smets, P. (1985). De school als organisatie. In J.L. Peschar & A.A. Wesselingh. *Onderwijs sociologie*. Groningen: Wolters Noordhoff.
- Sociaal-Cultureel Planbureau (1990). *School en schaal*. Rijswijk: CPB.
- Swanborn, P.G. (1987). *Methoden van sociaal-wetenschappelijk onderzoek*. Meppel:Boom
- Thoolen, E. (1987). Afvloeiingsregelingen in het onderwijs. *Meso*, 35 (7), 7-18.
- Toenders, L. (1989). Scholen in het informatietijdperk. In R. Van Hoewijk (ed.). *Schoolleider van vandaag, klaar voor morgen?: Liber amicorum voor Kees Gielen*. (pp. 15-47) Arnhem: Hogeschool Interstudie, Sector Management.
- Versloot, A.M. (1990). *Ouders en vrijheid van onderwijs: Schoolkeuze in de provincie Utrecht*. Utrecht (dissertatie).
- Visscher, A.J., & Amelsvoort, H.W.C.H. van (1993). *Evaluatie strategische factor personeelsbeleid BVE-sector (subsidieaanvraag SVO 93125)*. Enschede: Universiteit Twente, OCTO.
- Vries, A. de (1987). Kiezen voor een school. Schoolkeuze en marketing; een analyse van de 'afnemers' in het onderwijs. *Meso*, 22-26.
- Wansink, H. (1992). *Een school om voor te kiezen, naar een actuele onderwijspolitiek*. Amsterdam: Bert Bakker.
- Wieringen, A.M.L. van (1987). Deregulering en kritieken op onderwijsbeleid: vermogen van scholen tot voeren van bestuurlijk beleid. In L. Genemans (red.). *Autonomie van scholen en deregulering* (pp. 41-56). Nijmegen: ITS.



Tijdschrift voor Onderwijsresearch 14 (1989), nr. 4, pp. 237-258

## Onderwijs en Maatschappelijke Ongelijkheid, een Overzicht van 25 jaar theorievorming en onderzoek in Nederland\*

Jan C.C. Rupp en Anton A. Wesselingh  
*Universiteit van Utrecht en Nijmegen*

### ABSTRACT

Education and Social Inequality. A Review of 25 Years of Dutch Research.

This paper argues that the developments of the last two and a half decades in Dutch research on education and social inequality are characterized by a loss of importance of the debates on the various dimensions of the (essentially contested) concept of inequality, which were dominant in the first decade, and by a growing interest in the development of more specific theories and in doing historical and other empirical investigations. The three conceptual debates are resp. on deficit versus difference, on intelligence: heredity or environment, and on reproduction versus emancipation. These debates are important insofar as they urge the social scientist to define clearly what the problem to be investigated is and what his or her standards are. Fruitful they became when the research was directed at cultural differences between families, differences between and within schools, bilingual education and negotiations in the classroom, related with social class, gender and ethnic background. The theories of Bourdieu at the one hand and of Boudon at the other became very influential in the course of the eighties. A strong tradition was built up in research on schoolcareers following Coleman a.o. with implications mainly in the sphere of the political-arithmetic tradition. A characteristic feature of the Dutch research climate is a total disagreement on the question if and how a form of integration of secondary education must be realized. The reviewers plea for more comparative research in the way Ringer did.

### 1. INLEIDING EN KORTE SAMENVATTING

Wie gevraagd wordt een balans op te maken van vijftig jaar theorievorming en onderzoek in Nederland op het gebied van onderwijs en maatschappelijke ongelijkheid, is genoodzaakt de Nederlandse activiteiten in een internationaal kader te plaatsen. Buitenlandse ontwikkelingen oefenen invloed uit op de wetenschappelijke activiteiten in ons land en - zij het in mindere mate - omgekeerd. Een vergelijkend overzicht maakt tevens duidelijk wat typisch Nederlands is aan het overzicht, hetzij doordat er plekken wit zijn gebleven op de Nederlandse onderzoekskaart, hetzij omdat de betreffende wetenschappelijke activiteit voornamelijk in ons land wordt aangetroffen. De Nederlandse situatie behoeft op zich weer een verklaring. De discussie ligt in ons land bij een onderwerp als ongelijkheid bijvoorbeeld, zo gevoelig dat het niet zonder meer legitiem wordt gevonden er onderzoek naar te doen; organisatie van onderzoek en van onderzoekers in verbanden als



de Vereniging voor de bevordering van het Onderzoek naar de relatie tussen Onderwijs en Maatschappelijke Ongelijkheid (OOMO) en werkgroepen van SISWO, zijn noodzakelijk om legitimatie te verkrijgen en te behouden. De verticale verdeling in 'zuilen' en de verhouding tussen de zuilen is een vraagstuk dat veel belangrijker en eigener voor de samenleving wordt gevonden dan de horizontale verdeling in rangen en standen, in seksen en in etnische groeperingen. Het begrip eenheid wordt eerder geassocieerd met de pacificatie tussen de zuilen dan die tussen de sociale lagen (cf Vos e.a., 1985).

Een review heeft altijd een beoordelend karakter. Sommige ontwikkelingen worden als positief beoordeeld, andere minder. Richtinggevend bij zo'n beoordeling is de wetenschapsopvatting van de beoordelaar. De onze sluit aan bij die van Lakatos (Chalmers, 1981). Wij gaan er van uit dat er tegelijk meerdere onderzoeksprogramma's lopen, ieder met eigen domeinvooronderstellingen (mens- en maatschappijopvattingen) en harde kern en met eigen positieve heuristieken, die globaal richting geven aan het onderzoek, onder meer op het hier besproken terrein. Een voorbeeld is de positief heuristische waarde van de metatheorie van Olson (*The logic of collective action*) voor het onderzoek van Boudon (1973) naar de classespecifieke afweging van de kosten en baten van eventuele sociale stijging van de kinderen. Een voorbeeld uit een ander programma is de positieve heuristiek van de kapitaaltheorie van Bourdieu (1979) voor de theorie over stratificatie en sociale klassen, waarin nu niet langer alleen economische hulpbronnen, maar ook culturele worden verdisconteerd (Sixma en Ultee, 1983).

In Nederland heeft het Leidse Talentenproject (Van Heek e.a., 1986) een belangrijke rol gespeeld. Het 'Talentenproject' is de verzamelnaam voor een serie onderzoekingen naar de maatschappelijke factoren die de doorstroming van geschikte leerlingen ('talenten') naar het voortgezet onderwijs beïnvloeden. Uitgangspunt was de idee dat zich in de qua status lagere sociale lagen in Nederland een 'reservoir' aan niet-gebruikt talent zou bevinden, een reservoir dat Nederland hard nodig had voor zijn industrieel-technologische ontwikkeling. Ook in andere Westerse landen werden dergelijke projecten uitgevoerd.

De discussie binnen en rondom dit project maakte duidelijk hoezeer het begrip onderwijsongelijkheid een 'essentially contested concept' (Connolly, 1983) is, dat wil zeggen een begrip waaraan meerdere, soms onderlinge tegenstrijdige dimensies te onderscheiden zijn. Die dimensies hangen samen met brede theoretische orientaties, met onderzoeksprogramma's derhalve. Ook in ons land bestaan meerdere onderzoeksprogramma's (vgl. Harbers, 1986).<sup>1</sup>

Lange tijd stonden in het wetenschappelijk debat over onderwijsongelijkheid drie kwesties in het brandpunt van de belangstelling: 'difference-deficit'; intelligentie, erfelijkheid en omgeving; en reproductie versus reductie van ongelijkheid. Deze debatten speelden zich in de jaren zeventig grotendeels af op het vlak van de domeinvooronderstellingen van de diverse onderzoeksprogramma's en niet op theoretisch niveau. Het ging daarbij om de vraag van welke aard het probleem van onderwijs en maatschappelijke ongelijkheid nu precies is, of het eigenlijk wel een probleem is en zo ja, of het onderwijs wel iets aan de maatschappelijke ongelijkheid kan veranderen. Het is belangrijk de discussie over domeinvooronderstellingen, normen en waarden te onderscheiden van die over theorieën die er op gericht zijn een verschijnsel, een geïdentificeerd deel van de werkelijkheid te

1. Ten onrechte stelt Harbers overigens dat de leiders van het Talentenproject zich bij hun onderzoek hebben laten leiden door hun meritocratische voorkeuren. De uitkomsten van het onderzoek waren, juist tegen de verwachtingen van de onderzoekers in, immers dat er van een reserve aan talenten onder de arbeidersbevolking geen sprake was. Toen andere onderzoekers aantoonde dat er wel een reserve was, gaven de leiders van het project dit niet zonder moeite toe.

verklaren. Een discussie over mens- en maatschappijopvattingen kan ook met argumenten, op rationele wijze worden gevoerd (Habermas, 1983). Maar een discussie op het morele niveau heeft geen verklarende kracht. Een voorbeeld uit het debat over onderwijsongelijkheid kan dat verduidelijken. De meeste onderzoekers - maar niet alle - vatten gelijkheid op in termen van gelijke schoolloopbanen bij gelijke prestaties. Deze meritocratische definitie fungeert als de ideale norm waartegen de feitelijke ontwikkeling kan worden afgezet. Die norm is arbitrair, maar niet willekeurig want één uit een beperkt aantal mogelijkheden. Een andere opvatting van gelijkheid is een gelijke verdeling van schoolloopbanen over scholen van diverse typen tussen kinderen van verschillende sociale herkomst, sekse en etnische herkomst. Een derde definitie, ideale norm is één schoolloopbaan voor alle kinderen. Een dergelijke norm kan gebruikt worden als maat voor het democratisch gehalte van een land. Bij internationale vergelijkingen gaat concurrentie tussen landen een rol spelen. Voorzover het internationale gezelschap van onderzoekers er al in slaagt consensus te bereiken over een norm voor (on)gelijkheid (bijvoorbeeld in termen van sociale mobiliteit), worden de resultaten van die vergelijkingen altijd sterk aangevochten (Ganzeboom, 1983). Van dergelijke definities zijn er meerdere in omloop (Matthijssen, 1971, 1975; Meijers, 1983). Er zijn ook onderzoekers die alle tot nu toe genoemde definities liberaal vinden omdat het probleem gedefinieerd wordt binnen het bestaande schoolsysteem terwijl dit systeem zelf bekritiseerd zou moeten worden in het belang van met name kinderen uit de arbeidersklasse (Van Calcar e.a., 1968; Brands e.a., 1977), van vrouwen (Veneberg, 1982) en van etnische groepen (Mullard, 1986). Hoe men hierover ook denkt, dergelijke debatten gaan over normen en verklaren niets, omdat zij nog geen heuristische waarde hebben. Het is belangrijk waar mogelijk resultaten van onderzoek bij de discussie over normen en waarden te betrekken. Het kan veel misverstanden voorkomen. Zo hoeft gelijkheid in het onderwijs nog niet gepaard te gaan met mobiliteit in de maatschappij: in bepaalde historische perioden blijkt een grotere mobiliteit in het onderwijs zelfs ten koste te gaan van de sociale mobiliteit (Ringer, 1979, p. 29; Dronkers, 1986, p. 59). Een ander voorbeeld is de bevinding dat in het onderwijsstelsel meerdere hoofdwegen en zijwegen bestaan (op weg naar de universiteit of één van de in een land voorkomende soorten universiteiten; op weg naar een leidinggevende positie in de samenleving) en dat de meningen over wat onder kwaliteit van het onderwijs te verstaan, ook onder de diverse elites, sterk uiteenlopen. Dit pleit tegen de aanname dat in het onderwijs sprake zou zijn van één ideologie, of van een dominante ideologie (Abercrombie, e.a. 1981; *British Journal of Sociology of Education*, 1986, 2; *Oxford Review of Education* 13, 1987; Jungbluth, 1985).

Gaandeweg is het accent verschoven van discussies over mens- en maatschappijopvattingen (de domeinvooronderstellingen van de onderzoeksprogramma's) naar discussies over meer specifieke uit de heuristieken van de programma's afgeleide theorieën als mogelijke verklaringen van het ongelijkheidsprobleem. Een rol daarbij hebben ook gespeeld de teleurstellende ervaringen opgedaan met de compensatie- en activeringsprogramma's. Deze vormden voor enige van de bij de stimuleringsprojecten betrokken sociale wetenschappers aanleiding zich hernieuwd en op een meer fundamentele wijze dan voorheen bezig te gaan houden met de vraag naar de achtergronden van onderwijsongelijkheid. Het debat over 'difference-deficit' is theoretisch productief geworden enerzijds in het onderzoek naar verschillen tussen en binnen scholen, anderzijds in het onderzoek naar tweetalige opvoeding (Jungbluth, 1985; Rupp, 1988). Het debat over de intelligentie is door Vroon (1980) voorlopig afgesloten met de conclusie dat het vooralsnog niet mogelijk is de normen van de onderzoeker en de gevonden feiten uit elkaar te halen. Het reproductiedebat is theoretisch productief geworden toen men, geïnspireerd door de interpretatieve sociologie, verklaringen voor de ongelijkheid in en door het onderwijs mede is gaan zoeken in de richting van onderhandelingen en strijd in de klas (Wesselingh, 1985b, 96/110).

Een balans over Nederlandse theorievorming en onderzoek opgemaakt, geeft dan het volgende beeld te zien. Positief is de gestage groei van het *historisch onderzoek* op het probleemgebied. Kennis van de geschiedenis van een probleem en van de pogingen dat probleem op te lossen, is onontbeerlijk voor een goed begrip ervan. Historisch onderzoek wordt in Nederland niet verricht bij gebrek aan theorie, integendeel het onderzoek rust op een theoretische basis: Van Tijn (1977) plaatst het onderwijs in een sociaaleconomische kontekst, Frijhoff (1981) in een sociaal-culturele. Positief is het historisch onderzoek ook omdat er comparatief te werk wordt gegaan, een werkwijze waarvoor Idenburg (1961, 1962, 1964, 1967 en 1971) al sterk pleitte en die hij zelf met succes heeft toegepast. De Belgisch-Nederlandse Vereniging voor de Geschiedenis van Opvoeding en Onderwijs speelt de laatste jaren een stimulerende rol.

Een tweede positieve ontwikkeling situeren wij binnen het onderzoek naar *schoolloopbanen*. Ontwikkelingen in de sfeer van de statistische methoden en van modellen, vooral in de USA, worden hier productief gebruikt. Vooral de jongste discussie over de diverse niveaus van analyse (individueel niveau, schoolniveau, regionaal niveau, landelijk niveau) is van grote importantie. Wil dergelijk onderzoek echter niet in empiristische en ad-hoc verklaringen blijven steken, dan is het wel nodig theorieën te ontwikkelen over de aard en werking van het sociale zowel als over de aard en werking van het individuele. Productief is het schoolloopbaanonderzoek tot nu toe vooral geweest in het pogen ontwikkelingen in de tijd aan te geven rond het probleem van onderwijsongelijkheid en mogelijke invloeden van beleidsmaatregelen te traceren.

Een derde positieve ontwikkeling valt aan te wijzen op het vlak van de *theorievorming*. De rationele keuzetheorie van Boudon en de kapitaal-theorie van Bourdieu (cultureel-, sociaal- en economisch kapitaal) zijn daarvan de belangrijkste voorbeelden. Deze theorieën hebben vele Nederlandse onderzoekers geïnspireerd. Daarnaast heeft een organisatietheorie als die van Marx (1975) vruchtbaar gewerkt op het onderzoek naar verschillen in effectiviteit tussen scholen.

Minder positieve ontwikkelingen zijn er ook te melden. Vergeleken met de gang van zaken in andere landen dient de *universiteit* in Nederland vrijwel niet tot voorwerp van onderzoek. In andere landen is de universiteit van oudsher het meest belangrijke onderwerp van onderzoek en theorievorming, ook als het gaat om vraagstukken van ongelijkheid. In Nederland domineert de opvatting dat het basisonderwijs het belangrijkste is omdat dat het onderwijs voor allen is. Men kan met evenveel recht stellen dat de universiteit in de samenleving als het belangrijkste onderwijs geldt en dat daarom - als ideaal - iedereen een universitaire studie moet kunnen volgen. Het weinige onderzoek dat tot nu toe wel is verricht naar de universiteit en naar nieuwe vormen van intellectualiteit (o.a. Frijhoff, 1981, 1983; Wendrich, 1981), verkeerde lange tijd in een marginale positie in ons land. Hopelijk kan de in 1982 opgerichte wetenschappelijke vereniging Batavia Academica daarin verandering aanbrengen, en zal deze vereniging ook het ongelijkheidsvraagstuk in zijn onderzoeksprojecten meenemen.

Een tweede minder positieve ontwikkeling is de impasse in de discussie over de vormgeving van het *onderwijsvoorrangsbeleid* en van de integratie van (de eerste fase van) het *voortgezet onderwijs*. De overmatige belangstelling voor evaluatie-onderzoek duidt op een angstig gebrek aan inhoudelijke denkbeelden. Het is een te onderzoeken probleem op zich waarom in Nederland, in tegenstelling tot bijna alle andere Westerse landen, nog steeds nauwelijks geïntegreerd voortgezet onderwijs voorkomt en de bestaande middenscholen de bestrijding van ongelijkheid niet in hun doelstellingen hebben opgenomen<sup>2</sup>. De

2. Zie Rupp (1983). Tijdens een internationale conferentie over de middenschool besteedde Lagerweij (1986) geen enkele aandacht aan dit probleem, in tegenstelling tot de sprekers uit de andere landen. In de plannen van de W.R.R. voor de basisvorming was in eerste instantie bestrijding van ongelijkheid niet als doelstelling opgenomen.

V.S. begonnen met de high school reeds aan het einde van de vorige eeuw; in Engeland richtte men direkt na de tweede wereldoorlog comprehensive schools op; in de Bondsrepubliek bestaan in sommige deelstaten wel Integrierte Gesamtschulen en in andere niet.

Een derde negatief punt op de balans is de stagnatie in de *theorievorming ten aanzien van de ongelijkheid naar sekse en naar land van herkomst*. De snelle, positieve ontwikkelingen in de sekse-ongelijkheid kunnen niet verklaard worden met de beschikbare theorieën (cf Jungbluth, 1982), die weliswaar er op gericht zijn de ongelijkheid te verklaren, maar niet de condities aangeven waaronder ongelijkheid zou kunnen worden verminderd en historisch verminderd is. Op het gebied van ongelijkheid en etniciteit vindt de door geen sociale wetenschapper voorziene ontwikkeling plaats dat ouders Islamitische scholen gaan oprichten. Wellicht is een veelbelovende wetenschappelijke ontwikkeling het zoeken naar per etnische groep verschillende emancipatiewegen (Koot en Rath, 1987). Voor het overige worden er voornamelijk discussies op het niveau van domeinvooronderstellingen gevoerd over de betekenis van het behoud van taal en cultuur van het land van herkomst, over de ernst van discriminatie en dergelijke.

Moet een review-artikel over onderzoek op het gebied van onderwijsongelijkheid strikt genomen alleen gaan over de voortgang in de kennis over de rol die het onderwijs speelt ter verklaring van verschillen in schoolloopbanen die samenhangen met sociaal milieu, sekse en etnische herkomst? Wij vinden van niet. Ook factoren die buiten het onderwijs liggen zijn van groot belang. Zo is dan ook inzicht nodig in de vraag wat onder 'sociaal milieu' verstaan moet worden. Anders gezegd: welke verschillen in opvoedingspraktijken hangen samen met verschillen in schoolloopbanen?<sup>3</sup> Hetzelfde geldt voor de invloed van de regio en van het land van herkomst. Maar de rol van het onderwijs zelf blijft natuurlijk een belangrijk studieobject. Na een langdurige impasse als gevolg van de resultaten van het onderzoek van Coleman (1966) en van Jencks (1972) is de kennis over de rol van het onderwijs goed op gang gekomen. Coleman vond indertijd geen specifieke bijdrage van (verschillen tussen) scholen aan het probleem van de ongelijkheid. Hij onderzocht vooral organisatorische kenmerken van scholen. Fijnere statistische analysetechnieken maken het intussen ook mogelijk de invloeden van het sociaal milieu medierende (reproducerende en versterkende) rol van scholen in kaart te brengen. Schoolloopbanen, ongelijkheid in stand houdende of veranderende kenmerken van schoolstelsels als geheel, van ontwikkelingen in die stelsels, van verschillen tussen scholen en verschillen in scholen gemaakt tussen leerlingen van verschillende herkomst, worden langzamerhand ontdaan van het stof dat hen onzichtbaar maakte. Dat geldt ook en al eerder voor de rol van de gezinsopvoeding, van de taalontwikkeling in relatie tot de sociale omgeving en van het onderzoek naar levenskansen in onderwijs, arbeid en maatschappij in het algemeen. We zullen van de stand van zaken op deze terreinen melding maken in dit overzicht. Voor verschillen tussen onderwijsgeveenden en hun aktiviteiten en tussen leerlingen van dezelfde herkomst bestaat nog maar weinig belangstelling, althans vanuit het gezichtspunt van onderwijsongelijkheid.

## 2. OVER HET SCHOOLLOOPBAANONDERZOEK

Het schoolloopbaanonderzoek kent een bredere doelstelling dan het leveren van een bijdrage aan de analyse en verklaring van de ongelijkheidsrelaties in het onderwijs (al is het met dat thema wel sterk begaan). Het loopbaanonderzoek opereert over een breder front

3. Ringer (1979) wijst op de sterk uiteenlopende opvattingen over het verschijnsel 'stratificatie' in de diverse landen.



van een algemene thematiek: het verklaren van empirische verschillen in schoolloopbanen in het algemeen en niet speciaal de sekse- of milieuspecifieke verschillen tussen groepen leerlingen. Een goed voorbeeld van deze brede aanpak biedt het werk van Blok en Saris (1980). Ook 'levenskansen' hoort dus bij dit terrein van onderzoek gelet op het belang van onderwijs voor de positietoewijzing in de maatschappij. Het is dan ook niet onze bedoeling de opkomst en de bloei van het loopbaanonderzoek uitputtend te beschrijven; daarvoor kan men elders terecht<sup>4</sup>.

Ons interesseert de bijdrage die het schoolloopbaanonderzoek heeft geleverd aan de verheldering van de ongelijkheidsproblematiek.

In Nederland loopt de ontwikkeling van het schoolloopbanenonderzoek chronologisch gezien als volgt: na een aantal jaren van betrekkelijke stilte na het Talentenproject komt een nieuwe stroom onderzoekingen op gang. Dat onderzoek bouwde voort op de nogal aangevochten resultaten van de eerste reeks onderzoekingen uit de jaren zestig (o.a. Van Heek e.a., 1968) en op nieuwe resultaten van buitenlands onderzoek (o.a. Jencks, 1972, Boudon, 1973 en Bowles en Gintis, 1976). Van Heek dacht, zoals later bleek ten onrechte (Boon van Ostade, 1972; Van Kemenade en Kropman, 1972; Van Heek e.a., 1973), dat er van een reserve aan onbenut talent niet echt sprake was. Jencks c.s. meenden dat de rol van de school betrekkelijk secundair is als het gaat om reproductie van maatschappelijke ongelijkheid. Bowles en Gintis stellen dat de school er wel degelijk toe doet en juist bijdraagt aan het handhaven van ongelijkheid. Men blijft ook nu vooral geconcentreerd op de overgang van het lager naar het voortgezet onderwijs en de opvoedingsprocessen in de voorschoolse periode.

In 1975 verschijnt de dissertatie van Peschar, waarin school- en beroepsloopbanen van leerlingen van uiteenlopende sociale herkomst worden vergeleken, onder controle van de gemeten intelligentie o.a. De discussie die op dit interessante project volgt, wordt voor een deel in dit tijdschrift gevoerd (TOR 1975). Dronkers en De Jong (1978) introduceren de padanalyse in het loopbaanonderzoek. Hun replicatie van Jencks (1972) met Nederlandse data werkt inspirerend maar blijkt uiteindelijk toch niet te voldoen: de auteurs stappen over naar het zogenaamde Wisconsin-model (Sewell and Hauser, 1980). De hausse van padanalyse en andere regressievergelijkingen die daarna volgt wordt door Tesser (1986) beschreven, evenals de opkomst van de Lisrel-techniek en de toepassing van enkele niet-metrische technieken, waaronder log-lineaire analyse. Tessers conclusie 'dat de ontwikkeling van data-analysetechnieken' (1986, 11) is juist. Inderdaad is de methodische geavanceerdheid van het schoolloopbaanonderzoek geen punt van discussie, die is algemeen erkend en het loopbaanonderzoek dankt zijn erkenning en reputatie vooral daaraan. Niet in de laatste plaats doordat dank zij het invoeren van nieuwe variabelen, het uitputtend analyseren van bestaande data en het invoeren van multi-level analyses het empirisch fundament onder het ongelijkheidsonderzoek is verstevigd.

Methodisch geavanceerd en ook beleidsgericht? Voor een deel zeker. Niet voor niets plaatst Tesser (1986, p.16) het Nederlandse schoolloopbanenonderzoek in de Engelse traditie van het zogeheten 'political-arithmetic' onderzoek. Dat impliceert "een gerichtheid op het beschrijven en dokumenteren van de toestand van de samenleving in verband met centrale sociale en politieke kwesties. Een waarde bepaalde keuze van onderwerpen gaat daarbij gepaard met relatief objectieve en geavanceerde methoden van data-analyse."<sup>5</sup>

4. Bij Dronkers (1986) bijvoorbeeld. Overigens is de huidige bloei van het schoolloopbaan-onderzoek ook niet vanzelf tot stand gekomen. Vooral dank zij het onvermoeibaar activeren, publiceren en congresseren van enkelen is veel bereikt.

5. Tesser verwijst hierbij naar Halsey, Heath and Ridge (1980).



Een dergelijke karakterisering lijkt hoogstens voor een deel van toepassing op het Nederlandse onderzoek. We komen daar nog op terug. Dronkers, sprekend over het complexe karakter van de materie en over de moeite die het kost de ingewikkelde verbanden te analyseren: 'Deze ontrafeling van effecten is van groot praktisch belang' (1986, p.62), daarmee expliciet wijzend op de ongefundeerdheid van talloze ideeën en politieke uitspraken. Maar, Tessers omschrijving van hiervoor van het loopbaanonderzoek en de 'sociale en politieke kwesties' waarmee dat onderzoek zich bezig houdt, is slechts ten dele correct en wijkt enigszins af van Dronkers' uitspraak die meer wijst op koele distantie, het belang van 'getting the facts', het leveren van een solide empirische basis op grond waarvan (door anderen) politieke uitspraken kunnen worden gedaan en beleidsbeslissingen genomen. De boeiende analyse van Dronkers bijvoorbeeld, over veranderingen in de samenleving (met name in de economie) en de groei van de onderwijsdeelname, laat wel iets vermoeden over de maatschappelijke consequenties van een en ander maar biedt geen voorspellingen (Dronkers, 1989a). We wijzen hier ook op het onderzoek van Faasse e.a. (1986). Veel schoolloopbaanonderzoek lijkt juist sterk naar binnen gekeerd, geobsedeerd door de complexe relaties, het ontrafelen daarvan en de discussie daarover. De behoefte om nieuwe grenzen te verkennen op het terrein van de methoden en technieken, leidt dan ook tot steeds andere secundaire analyses op bestaande (gerepareerde) data van soms lang geleden afgesloten cohorten. Nieuwe dataverzameling vindt plaats.

Over de stand van zaken met betrekking tot de theorievorming moeten we enkele kanttekeningen maken. Op dit terrein worden aan de ene kant enige vorderingen gemaakt vooral rond het thema sociale stratificatie (Sixma & Ultee, 1983; Luykx e.a., 1986) en ook met betrekking tot een theorie van de bijdrage van cultuur en culturele hulpbronnen in het onderwijs (Ganzeboom & De Graaf, 1983; Ganzeboom e.a., 1987; De Graaf, 1987). Ook de verworvenheden op methodisch terrein die hebben geleid tot inzichten in de problemen van multi-level analyse, betekenen een grote vooruitgang in het sociaalwetenschappelijk onderzoek op dit terrein (vgl. Krefst, 1987).

Aan de andere kant is van een theorie over het proces van de schoolloopbaan en haar determinanten toch niet echt sprake en is de oogst bepaald mager. Het pionierswerk van Boudon over de keuzemomenten die mensen in hun school- en beroepsloopbaan tegenkomen, heeft de Nederlandse onderwijssociologen (en onder hen de loopbaanonderzoekers) wel geïnspireerd maar heeft hen niet kunnen helpen een 'loopbaantheorie' te ontwikkelen. Tesser is heel duidelijk op dit punt: 'Discussie over de relatie tussen analysetechnieken en toelaatbare empirische uitspraken op basis van de onderzoeksresultaten die met die technieken worden gegenereerd, heeft inhoudelijk geleid tot een zekere bescheidenheid.' (ibid.). Zo kun je het ook zeggen. Een voorbeeld: in het schoolloopbaanonderzoek is de aandacht terecht uitgegaan naar enkele variabelen die al na het Talentenproject leken 'afgeschreven', variabelen zoals regio en richting. Maar ondanks voorwerk van bijvoorbeeld Meester en De Leeuw (1984), zijn wij er nog steeds niet in geslaagd systematische vergelijkingen te maken tussen het gewicht van de variabelen SES, sekse en regio. Is er een relatie gelegd tussen regio en 'cultureel klimaat' in de zin van het arsenaal aan vormen waarover inwoners van een bepaalde regio kunnen beschikken om hun bestaan vorm te geven (vgl. Vliegen en De Jong, 1981)? En is het loopbaanonderzoek theoretisch in staat geweest te verklaren waaraan de forse inhaal-manoeuvre van vrouwen in het onderwijs (gelet op de deelnamecijfers aan v.o. en w.o.) moet worden toegeschreven?

Ook de volwasseneneducatie, van groot belang voor de levenskansen van mensen, blijft tot nu toe buiten beeld, zoals trouwens ook enkele andere deeltijdse vormen van onderwijs. Schoolloopbaanonderzoek beperkt zich grotendeels tot het reguliere dag-onderwijs en daarbinnen vooral tot het algemeen voortgezet onderwijs. Zo is tot nu toe ook geen aandacht besteed aan bedrijfsopleidingen, om een voorbeeld te noemen van deze beperking.

De vraag is ook of het totaal ontbreken van de variabele etniciteit in de theorievorming binnen het loopbaanonderzoek tot nu toe, ook moet worden gezien als een gevolg van de deficieten van de traditionele bestanden. Voorlopig houden we het maar bij deze pragmatische veronderstelling, immers de brede en algemene doelstelling van het schoolloopbaanonderzoek wordt nogal eens doorkruist door de beperkingen van het materiaal: gedateerde en onvolledige gegevens, niet te herstellen lacunes, en - niet te vergeten - overwegingen van bescherming en privacy van individuen, zoals die door de overheid (bijv. het CBS) de laatste jaren worden gehanteerd.

Ontwikkelingen op iets langere termijn zijn moeilijk in het onderzoek te vatten, ook in het schoolloopbaanonderzoek. Dronkers wijst daar terecht op (1986, p. 120 ev.). Dat impliceert dat school- en stelsel factoren in het lopende loopbaanonderzoek maar ten dele kunnen worden betrokken. Te denken valt daarbij aan vernieuwingen in het recente verleden als de invoering van de Mammoetwet, het dalende leerlingenaantal, de verlenging van de leerplicht en voor de nabije toekomst zaken als de integratie kleuter- en lager onderwijs en de forse groei van het speciaal onderwijs. Wat de invloed van beleidsmaatregelen voorstelt in hun effecten op de schoolloopbaan, is moeilijk aan te geven. Vermoedelijk wrekt zich hier het gebrek aan internationaal vergelijkend onderzoek waarbij Nederland betrokken is en waarbij stelsel factoren in vergelijkend opzicht beschreven kunnen worden en hun specifieke betekenis voor de Nederlandse situatie duidelijk kan worden. Uitspraken voor Nederland als geheel krijgen dan meer reliëf en zijn verantwoord (vgl. Meijnen, 1988). Ook past hier tenslotte de kanteekening dat het onderzoek een enigszins 'provinciaals' karakter heeft: op enkele uitzonderingen na (Peschar, 1987; Ganzeboom e.a., 1987) houdt het loopbaanonderzoek zich vooral bezig met het analyseren van data van Nederlandse kohorten van plaatselijke of regionale oorsprong.

### 3. KENMERKEN VAN SCHOOLSTELSELS, SCHOLEN, ONDERWIJSGEVENDEN EN LEERLINGEN

#### Schoolstelsels

In de jaren zestig waren de vergelijkende studies van Idenburg (1961, 1962, 1964 en 1971) naar de schoolstelsels van diverse landen in de wereld toonaangevend. Hij betrok bij deze studies heel nadrukkelijk het ongelijkheidsvraagstuk. Idenburgs belangstelling ging vooral uit naar het voortgezet onderwijs en hij vergeleek in dat verband de high school in de V.S. met vormen van geïntegreerd voortgezet onderwijs in andere landen, waaronder Nederland met zijn mammoetwet. Bij latere onderzoekers als Bourdieu (1970, 1979) en Ringer (1979) en in ons land bij Frijhoff (1981, 1983), gaat de belangstelling primair uit naar de universiteit en naar het secundair onderwijs dat op de universiteit voorbereidt. Onderzocht worden de houding van de sociale klassen in een bepaalde historische periode tegenover de universiteit en de ontwikkelingen in deze verhoudingen door onder meer de opkomst van nieuwe sociale klassen en door technologische ontwikkelingen. Een belangrijke uitkomst van dit onderzoek is dat in ons land, evenals in Engeland, geen consensus tot stand gekomen is over de betekenis van onderwijs in het algemeen en van de universiteit in het bijzonder tussen de adel en het oude patriciaat enerzijds en de opkomende industriële bourgeoisie anderzijds. In ons land en in Engeland, hoewel zij ook verschillen in tijdstip van aanvang en in tempo van industrialisering, heeft de industriële bourgeoisie zich afgekeerd van de universiteit en is de bedrijfsopleiding, de 'praktijk' de belangrijkste eigen leerschool geworden. Dit in tegenstelling tot een land als Frankrijk waar de diverse elites (de culturele, de economisch-administratieve en de technische) elk

hun eigen vorm van universiteit hebben gerealiseerd. Onderzoek van Beekenkamp en Dronkers (1984) is te zien als een bevestiging van deze bevindingen voor de actuele situatie in ons land<sup>6</sup>.

Deze per land verschillende situatie heeft ook gevolgen voor de houding van andere sociale klassen tegenover onderwijs in het algemeen en het vwo (in de V.S. binnen de high school de 'academic track') in het bijzonder. Er zijn in ons land twee zeer van elkaar verschillende manieren om - het schaarse goed van - een leidinggevende positie te bereiken, afhankelijk van het maatschappelijk veld. Eén loopt via de universiteit, de ander loopt via het bedrijf. Ringer (1979) bekijkt bij zijn onderzoek naar de historische betekenis van de universiteit en het VWO in diverse landen, tevens het democratisch gehalte, de toegankelijkheid van de universiteiten. Nederland behoort met West-Duitsland tot de landen die van oudsher een relatief laag percentage handarbeiderskinderen en vrouwen op de universiteiten kennen (Girls and Women, 1986). Een probleem dat verklaring behoeft.

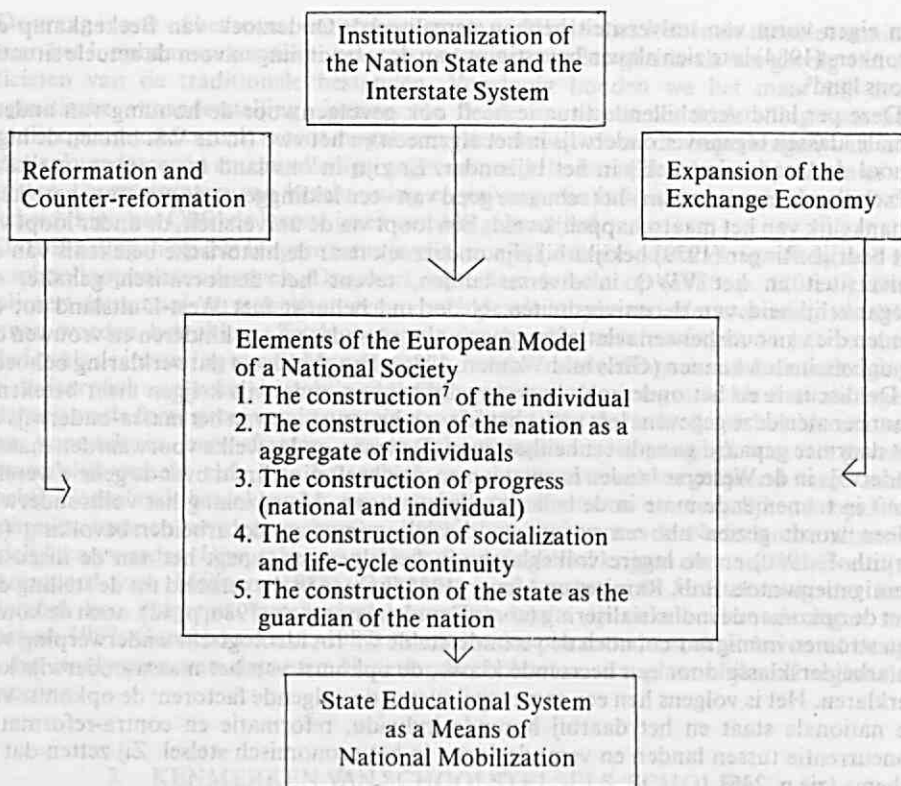
De discussie en het onderzoek naar ongelijkheid en onderwijs krijgen meer betekenis wanneer men deze gegevens legt naast het feit van het ontstaan van het massa-onderwijs en het daarmee gepaard gaande eenheidsstreven. De vraag onder welke voorwaarden massa-onderwijs in de Westerse landen is ontstaan en zich heeft uitgebreid over de gehele wereld, heeft in toenemende mate in de belangstelling gestaan. Maar zolang het volksonderwijs alleen wordt gezien als een vorm van disciplineren van de arbeidersbevolking (cf. Kruithof, 1980) en de lagere volksklassen in het algemeen, voegt het aan de discussie weinig nieuws toe. Boli, Ramirez en Meyer (1985) komen in dit verband tot de stelling dat niet de opkomende industrialisering (zoals Dronkers nog stelt, 1986, p. 42), noch de komst van stromen immigranten, noch de veronderstelde wil tot ideologische onderwerping van de arbeidersklasse door een heersende klasse, de opkomst van het massa-onderwijs kan verklaren. Het is volgens hen een samenspel tussen de volgende factoren: de opkomst van de nationale staat en het daarbij horende individu, reformatie en contra-reformatie, concurrentie tussen landen en veranderingen in het economisch stelsel. Zij zetten dat in schema (zie p. 246).

De constructie van het 'individu' en van de natie als een verzameling individuen, het individuele en nationale vooruitgangsgeloof, de denkbeelden over de opvoedbaarheid van de mens en over de staat als hoedster van de natie; al deze denkbeelden ontstaan tegen de achtergrond van de Reformatie en de Contra-reformatie enerzijds en de ontwikkeling van de kapitalistische (vrije markt)economie anderzijds. De Reformatie betekent dat de overheersende positie van één kerk - en daarmee van het samenbindend beginsel - in Europa vervalt. De geloofsverdeeldheid schept de behoefte aan een nationale staat.

Het tijdperk van de Verlichting in Europa geeft steun aan een (op rationalistische leest geschoeid) geloof in de vooruitgang van de menselijke samenleving onder de hoede van de nationale staat. Met behulp van het volksonderwijs zou het voor de staat mogelijk zijn het welzijn van de kinderen te dienen, de armoedecyclus te doorbreken waarin veel gezinnen in de loop van de achttiende eeuw terecht waren gekomen en actieve productieve staatsburgers te vormen. Onderwijs wordt van cruciaal belang, ook in de concurrentie met andere staten.

Onderzoek in ons taalgebied van Van Setten (1982) en Van Damme (1987) en ook wel van Lenders (1988) spoort met deze these. Deze theorie lijkt uiteindelijk veelbelovender dan die van Archer (1979, 1982; Rupert en Wolthuis, 1985), die ten onrechte de industrialisatie als grondvoorwaarde aanwijst. Massa-onderwijs wordt historisch gezien niet als tegengesteld aan standenonderwijs opgevat. Het gaat in het onderwijs voor allen

6. De onderzoekers beweren zelf ten onrechte dat hun onderzoek kritiek inhoudt op de theorie van Bourdieu.



Uit: F.O. Ramirez and J. Boli, 1987, p. 10)

om wat - aangepast aan rang en stand - de burger moet weten om zich als burger te kunnen gedragen. Het onderwijs heeft op deze manier ook een sterk opvoedend karakter gekregen. Daar de eisen aan de voor de burger noodzakelijk geachte kennis steeds opgeschroefd worden, gaat het algemene kennisniveau van de bevolking omhoog, maar blijven de relatieve verschillen tussen de diverse sociale milieus bestaan. Historici als Boekholt (1978), De Booy (1980), Knippenberg (1985) en Boekholt en De Booy (1987) hebben een schat aan materiaal bijeengebracht over de inhoud van het massa-onderwijs en de deelname eraan in de negentiende eeuw in Nederland. Hun analysekader beperkt zich jammer genoeg tot de schoolstrijd. Het zou zeer de moeite waard zijn hun materiaal en ook materiaal uit de moderne geschiedenis te heranalyseren vanuit het referentiekader van Ramirez en Boli op het punt van de opvoeding tot staatsburgerschap in een naast in zuilen ook in standen verdeelde en als zodanig te handhaven staatsvorm.

#### Verschillen tussen scholen

Het onderzoek van Coleman (1966) heeft lange tijd de toon gezet. Diens conclusie dat de

7. Ramirez en Boli spreken van 'The myth of...' in plaats van 'the construction of...' Wij prefereren 'construction' omdat dit een algemener begrip is dat ook de ideologische functie omvat.



school er niet of weinig toe doet en dat het sociaal milieu doorslaggevend is, is pas sinds de jaren tachtig genuanceerd. In navolging van onderzoek in Engeland en de USA is hier onderzoek gedaan naar de effectiviteit van scholen, onder andere door Meijnen (1984) en door Van der Wolf (1984). Het gaat hier om de vraag of scholen, bij gelijk gehouden leerlingsamenstelling, verschillen in schoolvorderingen bij leerlingen realiseren. Een vraag die ook bij de evaluatie van het onderwijsvoorrangsbeleid aan de orde komt. Bij dit onderzoek worden vooral aspecten van de organisatie van de school en van het onderwijs betrokken, afkomstig uit organisatiekundige theorieën. Jammer genoeg is er sinds Marx (1975) nog relatief weinig organisatie-theoretisch werk gedaan met betrekking tot het onderwijs. Kritiek op de onderzoeksmethode is o.a. geleverd door Kalmijn en Swanborn (1987). Zij wijzen er terecht op dat in het schooleffectiviteitsonderzoek veel te weinig aandacht wordt besteed aan de vraag naar de stabiliteit van de schoolvariabelen.

Een tweede vrij recente benadering is gebaseerd op het werk van Bourdieu. Uitgangspunt is dat ook voor scholen het marktbeginnsel geldt. Op elk niveau van het onderwijs zoekt een school zich een positie op de markt van scholen en tracht deze positie te handhaven. In ons land is deze markt voor wat betreft het voortgezet onderwijs evident (van VWO tot LBO) en binnen deze totaalmarkt bestaan ook weer deelmarkten, zoals binnen het VWO een Montessorilyceum tegenover een traditioneel lyceum (Kalmijn en Batenburg, 1986), en openbaar tegenover bijzonder onderwijs (Coleman e.a., 1982; vgl. ook Van Laarhoven e.a., 1986 en 1987, en Dronkers, 1989b). Gegeven hun marktpositie recruteren scholen leerlingen van bepaalde sociale groepen; bepaalde sociale groepen zoeken een school uit die past bij hun toekomstperspectief. Dit perspectief zal volgens Bourdieu per klasse verschillen en binnen elke klasse tussen fracties van die klasse met een meer economische dan wel een meer culturele oriëntatie (vgl. ook Bernstein, 1977; Collins, 1971). Volgens onderzoek van Rupp e.a. (1989) bestaat de marktpositie van basisscholen uit het opleidingsdomein van de school, dat zijn de scholen voor voortgezet onderwijs waarvoor de basisschool specifiek wil opleiden (de een voor VWO en HAVO; de ander voor MAVO en LBO enz.). Het opleidingsdomein is het belangrijkste aspect van wat genoemd wordt de opleidingsstatus van een basisschool. Het opleidingsdomein van een school is een zaak van voortdurende onderhandelingen tussen schoolleiding en ouders.

### Verschillen binnen scholen

Scholen kunnen behalve van elkaar verschillen, ook intern verschil maken in het onderwijs dat zij kinderen geven, afhankelijk van de herkomst en de sekse van die leerlingen. Gezien de lange traditie die de USA met de High School hebben, bestaat daar ook de langste onderzoekstraditie op het gebied van milieugebonden interne differentiatie binnen de High School (Oakes, 1985; vgl. diverse artikelen in het tijdschrift *Sociology of Education*, Vol. 1986 en 1987).

De laatste jaren is dit type onderzoek weer sterk in opkomst. In ons land heeft vooral Jungbluth (1985) de aandacht gevestigd op deze kant van de onderwijsongelijkheid. Zijn stelling is dat in elke school verschil wordt gemaakt in geboden onderwijs, afhankelijk van het leerlingbeeld dat onderwijsgevendenden hanteren, een beeld dat weer samenhangt met sekse en sociaal milieu van de leerlingen. Het belangrijkste verschil in het leerlingbeeld is de verwachting die men heeft van wat de leerling kan presteren, resp. welke vorm van voortgezet onderwijs de leerling zal kunnen bereiken. In het basisonderwijs, in elke basisschool tekenen zich in de loop van de jaren diverse stromen af voor het voortgezet onderwijs en deze stroomverdeling is klasse- en seksegebonden. Op deze manier wordt verschil in sociaal milieu vertaald in onderwijsongelijkheid (Zijlmans en Van der Kley, 1985).



**Verschillen tussen onderwijsgeevenden**

Na de studie van Renes (1969) over de recruitering van onderwijsgeevenden voor het lager onderwijs is er, in tegenstelling tot wat andere landen laten zien, geen onderzoek verricht naar de sociale achtergronden van onderwijsgeevenden van welk niveau van onderwijs dan ook. De achterliggende gedachte bij dergelijk onderzoek is dat een open recruitering van docenten, in het bijzonder aan de universiteit, een teken van democratisering is; verder dat verschil in sociale achtergrond van leerling en docent de emancipatie van de leerling zal belemmeren. De laatste tijd is wel onderzoek gedaan naar het kennisniveau van aanstaande onderwijzers (Koster 1987) in het kader van een algemeen onderzoek naar de kwaliteit van het onderwijs. Dergelijk onderzoek komt in Nederland moeizaam van de grond omdat de meningen over wat goed onderwijs is hier sterk verschillen en dan met name onder de diverse elites. Een teken hiervan is de grote belangstelling in deze bevolkingsgroepen voor de Vrije scholen, Montessorischolen, de Algemeen Byzondere Schoolverenigingen en dergelijke. Een dergelijke differentiatie kent men in andere landen niet. Ook in andere landen lopen de meningen uiteen maar het beeld is daar minder gefragmenteerd. In Engeland gaat het debat grofweg gesteld tussen twee stromingen: de liberaal-rationalistische ('core-curriculum') en de 'popular culture' traditie ('really useful knowledge') (zie Veugelers en Wendrich, 1986). In de USA gaat de discussie tussen klassieke vorming en modernisering (Bloom, 1987).

Van diverse kanten (Meijnen, 1988) is gewezen op de waarschijnlijkheid dat de verdeling van de meest bekwame docenten over de scholen ongelijk gebeurt, maar onderzocht is dit nog niet.

Opmerkelijk is de belangstelling voor de geschiedenis van het beroep van onderwijzer, in het bijzonder van de 'volksonderwijzers' en hun organisaties (DuBois-Reymond, 1981; Wolthuis, 1982; Karsten, 1986; Veld, 1987). In het voortgezet onderwijs wijzen we op onderzoek naar de invloed van de bonden op het onderwijsbeleid (Leune, 1976). Hopelijk groeit deze interesse uit tot een traditie. Het onderwijzersberoep is in ons land altijd een belangrijk middel geweest tot sociale stijging; het was de enige weg die naast het VWO toegang gaf tot de universitaire studie (in de pedagogiek). De historische studies geven een steeds nauwkeuriger beeld van de dilemma's waarvoor de 'volksonderwijzer' zich gesteld ziet. Een van die dilemma's is de vraag of werken aan de 'verheffing van het volk' niet altijd gepaard gaat met disciplineren (Lenders, 1988).

**Verschillen tussen leerlingen**

In veel van de gangbare theorieën en van het lopende onderzoek is het 'individu' impliciet het passieve resultaat van omgevingsinvloeden. Dat geldt niet voor de rationele keuzetheorie, waarin het individu als actor gezien wordt die voor- en nadelen van bepaalde keuzen rationeel afweegt en waarin het 'sociale' niet een entiteit is onafhankelijk van het individuele, maar het (soms onverwachte) gevolg van individuele handelingen. Deze theorie kan echter niet de bestaande normen verklaren noch de bestaande veranderingen daarin, en geeft ook geen inzicht in onderwijsstelsels en de veranderingen daarin.

Er is nog een theoretische stroming waarin aandacht wordt gevraagd voor het handelen van individuen. In die stroming wordt de ongelijkheid in en door het onderwijs opgevat als de resultante van strijd (onderhandelingen) in de klas over de definitie van de situatie. Tegenover de dominante definitie van de omstandigheden waaronder geleerd kan worden, wat belangrijke kennis is en wat een goede leerling, worden soms door leerlingen en soms door docenten andere definities gesteld (vgl. Van der Kley, 1983; Zijlmans en Van der Kley, 1985). De mogelijkheden om kollektief in verzet te komen tegen de dominante definitie worden op steeds ingenieuzere manieren verkleind, manieren die ontleend zijn aan de organisatie van de arbeid in fabriek en kantoor. Het is met name in de Amerikaanse

(Apple, 1979, 1982; Wexler, 1976, 1987; Giroux, 1983) en de Engelse onderwijssociologie (Willis, 1977) dat voor deze processen aandacht gevraagd wordt. In Nederland noemen we het onderzoek naar voortijdige schoolverlaters (Bijlsma, 1981) en naar de diverse definities van 'school' 'leren' en 'werken' zoals die onder leerlingen van LBO en MAVO worden aangetroffen (Geurts en Tesser, 1976).

#### 4. KENMERKEN VAN GEZINNEN

In onderzoek naar de samenhang tussen sociaal milieu en schoolloopbaan wordt als maat voor 'sociaal milieu' bijna altijd het beroep van de ouder(s) - meestal van de vader - en/of het opleidingsniveau en/of het inkomen genomen. Met andere woorden: de sociaal-economische status wordt gemeten. Stratificaties verschillen sterk per land en veranderen in de loop van de geschiedenis. Zo wordt er op het ogenblik in ons land gediscussieerd over een mogelijke (nieuwe) tweedeling van onze samenleving: tussen betaald werkenden en uitkeringsgerechtigden. De consequenties van een dergelijke tweedeling voor het onderwijs zijn nog onvoldoende helder (Bosman en Louwes, 1988). Behalve misschien in het geval van samenhang tussen de hoogte van het inkomen en gevolgd, niet-kosteloos onderwijs, verklaart een gevonden samenhang nog niets, want waarom heeft de opleiding of het beroep van de ouders invloed? In ons land heeft deze vraag zich altijd in de belangstelling van de onderzoekers mogen verheugen. Theoretisch en empirisch is door een reeks van onderzoekers (o.a. Vervoort, 1968; Rupp, 1969; Claeys, 1972; Meijnen, 1977; De Graaf, 1987), gezocht naar verschillen in opvoedingspatronen en -praktijken tussen de diverse sociale milieus, die mede verschillen in schoolloopbanen kunnen helpen verklaren. Onderzoek is gedaan naar de volgende aspecten (zie ook het overzicht bij Meijnen, 1984):

- a. de betekenis van de sociaal-economische omstandigheden, resp. hulpbronnen van de ouders;
- b. de betekenis van de sociaal-culturele omstandigheden van het gezin;  
Vervoort spreekt in dit verband van gezinsculturele factoren, Rupp van het cultuur-pedagogisch patroon van de opvoeding, Tesser noemt het intellectueel thuis-klimaat en De Graaf spreekt in navolging van Bourdieu van culturele hulpbronnen;
- c. verschillen in normen- en waardenpatronen; meer of minder prestatiegericht; een conformistische versus een positionele rolstructuur; meer of minder conformistisch op sociaal gebied; hoger of lager zelfbeeld. De algemene stelling van de onderzoekers komt er op neer dat een succesvolle schoolloopbaan een 'moderner' gezinsleven vraagt en dat moderne gezinnen vooral in de midden- en hogere klassen worden aangetroffen (vgl. Van Setten, 1987).

Geleidelijk aan is er een duidelijk consistent beeld ontstaan over de vraag welke opvoedingspatronen de samenhang tussen sociaal milieu en schoolloopbaan kunnen helpen verklaren. Het zijn hoofdzakelijk de culturele opvoedingspatronen door welke het milieu invloed uitoefent. De financiële hulpbronnen zijn in betekenis afgenomen. Bourdieu heeft laten zien dat per sociaal milieu onderscheid gemaakt moet worden tussen meer cultureel georiënteerde en meer economisch georiënteerde gezinnen. Deze verschillende oriëntaties leiden tot verschillende schoolkeuzen en tot andere opleidingsroutes. De opleiding van de ouders lijkt ook belangrijker te worden dan het beroep. Ook dit wijst in de richting van de onder b. genoemde factoren.

- Het betreft dan de volgende activiteiten in en rond het gezin:
- a. de culturele participatie van de ouders: deelname aan concerten, museumbezoek, leesgewoonten enzovoort, activiteiten die belangrijker zijn voor de schoolloopbaan van de kinderen naar de mate waarin zij een avant-gardistisch karakter dragen: het

luisteren naar een uitvoering van een strijkkwartet van Bartok is belangrijker dan het luisteren naar een pianoconcert van Beethoven enz.

- b. het inwijden van de kinderen in de sfeer van het culturele erfgoed: hen voorlezen en later bekend maken met de literatuur; hen meenemen naar concerten en musea.
- c. het direkt ondersteunen van de schoolloopbaan door het bezoeken van ouderavonden, het tonen van belangstelling voor het schoolwerk, zorgen voor de begeleiding van huiswerk zonodig enz.

Naast bovengenoemd onderzoeksprogramma dat de verklaring zoekt in hulpbronnen, is er het programma van de 'rational choice'. Boudon (1973) heeft dit uitgewerkt voor het probleemgebied onderwijsongelijkheid. Hij is voor de belangrijkste beslismomenten in de schoolloopbaan van kinderen nagegaan hoe elke ouder een afweging maakt van baten en kosten bij elk van de mogelijke keuzen en ook hoe deze afweging beïnvloed wordt door het sociaal milieu van de ouders. De potentiële baten van een keuze voor een VWO-opleiding kunnen variëren met de vraag naar academisch opgeleiden op de arbeidsmarkt. Daarbij kan ook prestige sterk meetellen, speciaal bij de economisch rijken zonder veel cultureel kapitaal. Kosten kunnen zijn de economische kosten of direkte kosten van de studie (hoewel daar een beurs tegenover kan staan), maar ook de indirecte kosten, de inkomsten die ouders derven doordat hun kind langer op school blijft. Maar er zijn ook geheel andere kosten. Boudon noemt hier de kosten van het arbeiderskind dat een school voor VWO zou gaan bezoeken en daarmee van zijn vriendjes en mogelijk ook van zijn ouders zou kunnen vervreemden (vgl. de kosten-baten analyse van sociale mobiliteit door Goldthorpe, 1980; Dronkers e.a., 1982). De kosten liggen hier op psychologisch vlak. Op grond van dergelijke afwegingen kunnen twee VWO-leerlingen met verschillende milieu-achtergrond verschillende onderwijsroutes gaan volgen, ook al verschillen zij niet in schoolprestaties (vgl. Peschar, 1975).

De kracht van Boudons theorie is dat zij preciese beweegredenen onderzoekt en keuzen kan verklaren. Het Nederlandse onderzoek à la Boudon heeft interessante gegevens opgeleverd, ook op individueel niveau.

## 5. MAATSCHAPPELIJKE ONGELIJKHEID, ONDERWIJS EN TAAL

In de beginperiode van het ongelijkheidsonderzoek in het onderwijs lag een van de accenten op taalontwikkeling en taalgebruik van (jonge) kinderen. De beheersing van de standaardtaal of de taal van de dominante groepen in de samenleving werd geacht een *conditio sine qua non* te zijn voor het leveren van goede schoolprestaties en daarmee voor het bereiken van een hoog opleidingsniveau. De sociale omgeving, het sociale milieu waarin men als kind opgroeit bepaalt voor een groot deel over welk taalgebruiksreservoir men uiteindelijk beschikt en daarmee de aansluiting naar de fase van de secundaire socialisatie in het onderwijs.

Onder invloed van de sociolinguïstiek (Appel e.a., 1976; Lenders e.a., 1978) die zich vooral in de jaren zestig sterk heeft ontwikkeld en zich uitgebreid heeft beziggehouden met de relatie tussen taal(gebruik) en sociale klasse, is veel aandacht besteed aan dit aspect van de relatie tussen onderwijs en maatschappelijke ongelijkheid (cf ook Huls, 1982). De theorie van Basil Bernstein (1971) is in dit verband van belang. Hij stelde dat de sociale omgeving leidt tot het al of niet beheersen van de standaardtaal en het ontwikkelen van vermogens (codes) om een bepaald taalgebruik te hanteren. Kinderen uit de hogere klassen zouden dergelijke vermogens in sterker mate 'van huis' meekrijgen dan kinderen uit de arbeidersklasse. De discussie die - voor en na Bernstein - de gemoederen bezig hield, ging

eigenlijk over de vraag of men zou moeten uitgaan van één standaardtaal en afwijkingen daarvan (deficietopvatting) of dat er in feite meerdere talen op gelijke voet naast elkaar bestaan (differentieopvatting). Bernstein zelf is er - blijkens tegenspraken in zijn eigen teksten - niet helemaal uitgekomen (vgl. Jackson, 1974; Atkinson, 1987), maar deze tegenstrijdigheden zijn inhaerent aan het multidimensionele concept onderwijsongelijkheid.

De definitie van het probleem waar het nu eigenlijk om ging, verschoof in de loop der jaren. Tot in de jaren zeventig speelde het probleem eigenlijk alleen maar voor 'inheemse' groepen met van het standaard Nederlands afwijkende talen of dialecten: de 'anderstaligen' bevonden zich binnen de grenzen van de ene nationale Nederlandse cultuur (Fries, Kerkraads enz., vgl. Wijnstra, 1976). Toen echter ook in ons land de immigratiegolven voor een multiculturele samenleving gingen zorgen, veranderde de situatie. Onze tot dan toe betrekkelijk eenvoudige oplossing van het probleem van de meertaligheid (extra lessen Fries bijvoorbeeld) werkt nu niet meer: andere oplossingen moeten worden gevonden: multicultureel of intercultureel onderwijs; onderwijs in eigen taal en cultuur. De zaak is aanzienlijk gecompliceerder geworden. Taalverwerving van leerlingen van etnische minderheden (o.a. het probleem van de tweetaligheid) is nu een belangrijk topic geworden. Scholen krijgen steeds meer te maken met leerlingen van verschillende etnische herkomst, vooral in de grote steden en worden zo multi-etnisch van samenstelling.

Over de aanpak van het probleem van de etnische groepen en hun taalontwikkeling dreigen evenals ten tijde van de sociolinguïstiek uit de jaren zestig en zeventig, twee kampen te ontstaan die op essentiële punten van elkaar verschillen. Het voert te ver die verschillen hier uitgebreid te vermelden. We volstaan met op te merken dat het lijkt alsof de scheidslijn loopt tussen diegenen die met enige nadruk toch acculturatie - zeg aanpassing - voorstaan aan de Nederlandse cultuur, opgevat als mono-cultuur ook in het belang van de bepaald niet tijdelijke immigrantenstatus (De Jong 1987 bijv.) en diegenen die de problematiek zien in termen van etnische verhoudingen in de Nederlandse samenleving (Leeman, 1985).

Sociolinguïsten houden zich nu uitgebreid bezig met het probleem van de tweetaligheid. De reeks proefschriften van de laatste jaren, waarin dat probleem wordt behandeld, (o.a. Appel, 1984, J. Teunissen, 1987) is aanzienlijk en de resultaten wijzen grotendeels in één richting: tweetaligheid zou de beste aanpak zijn voor de taalverwerving, ook van het Nederlands, van immigrantenkinderen. In deze kringen en in de publieke discussie vinden we dan ook voor- en tegenstanders van tweetaligheid (de methode van de 'onderdompeling'), maar deze standpunten blijven betrekkelijk 'technisch' zolang niet de kwestie van de monocultuur of het multiculturele karakter van de samenleving wordt gesteld inclusief de problemen rond de etnische verhoudingen die daarmee gepaard gaan. Het probleem vraagt derhalve om een zekere mate van 'sociologisering'. Een veelbelovende nieuwe stroming (Koot en Rath, 1987) gaat uit van culturele verschillen tussen etnische groepen met betrekking tot het belang dat de groep hecht aan onderwijs als middel tot emancipatie. Sommige etnische groepen hebben een traditie zich in een nieuwe samenleving via onderwijs op te werken, andere kiezen een weg in de politiek, weer anderen bouwen een zaak op. We wijzen hier op een recente ontwikkeling: de oprichting van Islamitische scholen.

## 6. ONDERWIJS EN LEVENSKANSEN (LIFE-CHANCES)

Sinds Ph.J. Idenburg in 1958 de term 'sleutelmacht der school' reserveerde voor wat later de allocatiefunctie van het onderwijs is gaan heten, heeft deze toeleveringstaak van het



onderwijs vrijwel onafgebroken in de belangstelling gestaan<sup>8</sup>. Zeker in het (onderwijs) sociologisch onderzoek bestaat op dit terrein een zekere traditie, die voortkomt uit het onderkende belang van de school en van scholing voor het bereiken van maatschappelijke posities, in het bijzonder de hogere. Onderwijs is van belang bij het toedelen van 'Lebenschancen' (Weber, 1956): leden van verschillende sociale klassen hebben ongelijke toegang tot schaarse middelen en zullen als gevolg daarvan een verschillende maatschappelijke positie innemen. Onze maatschappij is daarbij zo ingericht dat het onderwijs een 'bijna exclusieve rol heeft verkregen bij de verdeling van materiële en immateriële goederen en posities (...)', aldus Dronkers (1986, p.112) in een echo van Schelsky (1957, p.18). Het onderwijs vormt daarmee een belangrijke schakel in de continuering van maatschappelijke verschillen over generaties heen, zonder overigens die verschillen zelf te genereren. Het is kortom geen systeemkenmerk of onderwijskundig gevolg van de pedagogische structuur van het onderwijsbestel, maar een gevolg van de inrichting van de samenleving als geheel en de plaats van het onderwijsstelsel daarbinnen.

Dergelijke uitspraken over de rol van de school in de maatschappij zijn de laatste decennia zo zeer gemeengoed geworden in de onderwijssociologie dat de broodnodige relativisering meestal achterwege blijft. Toch wordt de rol van het onderwijs op die manier nogal overschat: de school heeft op dit punt bepaald geen monopoliepositie, want andere socialiserende instanties leveren hun aandeel (gezin, peergroup, media enz.). Bovendien heeft het onderwijs niet voor elke maatschappelijke groep dezelfde plaats en betekenis. De geschiedenis laat zien bijvoorbeeld dat economische elites een zeer selectief gebruik maken van de beschikbare onderwijsvoorzieningen. Dat betekent dat *de rol van het onderwijs bij de vorming van elites in de maatschappij* een belangrijk thema in dit verband kan zijn. In Nederland is dit een onderontwikkeld thema vergeleken met de situatie in het buitenland. Bourdieu (1964, 1984) is opnieuw het meest sprekende voorbeeld, maar ook in Engeland bestaat voor dit thema ruime aandacht vooral waar het gaat om de rol van de traditionele elite-scholen in dat land (vgl. *The Education of Elites*, 1977). De sleutelrol van de school en de monopoliepositie van dit instituut dienen derhalve gerelativeerd: dat verschilt naar tijd en plaats voor uiteenlopende maatschappelijke groepen.

In Nederland bestaat een duidelijke traditie op het terrein van onderzoek en theorievorming rond de relatie tussen *onderwijs en arbeid(smakt)*, een van de belangrijkste aspecten van het vraagstuk 'onderwijs en levenskansen'<sup>9</sup>. Op dit terrein is vooral werk verricht door Vervoort (1981 oa.) en door Van Hoof en Dronkers (1980; vgl. ook Van Hoof, 1987). Het werk van Van Hoof en Dronkers heeft voor een deel ook zijn sporen achter gelaten in het hiervoor besproken loopbaanonderzoek. De bedoelde studie werd verricht in het kader van het Nationaal Programma Arbeidsmarkt Onderzoek, waarbij verschillende universitaire instellingen en departementen waren betrokken. In dit verband moet ook het onderzoek van Diederer (1983 o.a.) worden genoemd, een rapportage van de 3e fase van het longitudinale project school- en beroepsloopbanen 'Van jaar tot Jaar' (zie ook Diederer, 1989). Maar ook buiten de kringen van loopbaanonderzoekers is er ontwikkeling te constateren op dit terrein. Het theoretisch werk van Vervoort (zie hiervoor) geïnspireerd door de Engelse socioloog Antony Giddens en ondersteund door Wesselingh (1985a), bevat een pleidooi de reproductietheorie theoretisch te verbreden in een meer Weberiaanse richting en uit te breiden naar de relatie onderwijsarbeid. Dat

8. Dat lijkt nu vanzelfsprekend maar dat was het toen allerminst. 'Het schoolwezen als economisch verschijnsel' (de titel van Idenburgs oratie aan de Universiteit van Amsterdam in 1967) is lange tijd verwaarloosd in de onderwijswetenschappen. Het thema is nog steeds marginaal ten opzichte van de traditionele onderwijskundige thema's.

9. Dronkers wijst in zijn overzichtsartikel ook nog op de 'huwelijksmarkt' en de relatie tussen bereikt opleidingsniveau en de kansen op deze markt (1986, 113 ev.).



betekent een verdieping van de 'sleutelmacht' van de school en een versterking van de rol van de arbeidsmarkt, omdat daar enkele vitale toewijzingsprocessen plaats vinden, die de functies van de school in die zin verhelderen dat naar de achtergronden van de 'sleutelmacht' wordt gekeken. Soortgelijke opvattingen zijn neergelegd in de rapportage van de analyse van de Volkstelling 1971, voorzover het de onderwijsgegevens betreft (Vliegen en De Jong, 1981). In deze lijn werken ook de sociologen Glebbeek en Mensen (1986). Hun eveneens Neo-Weberiaanse oriëntatie lijkt vooral geïnspireerd door Parkin (1971). Werk in deze lijn kan het herstel bevorderen van de relatie onderwijs, arbeid en stratificatie c.q. ongelijkheid. De Stichting voor Onderzoek van het Onderwijs (SVO) tenslotte, heeft prioriteit gegeven aan de uitbouw van een programmatische Onderwijs en Arbeids situatie. In een publikatie van Hövels en Peschar (1985) wordt de draad van het in 1980 afgesloten Nationaal Programma (zie hiervoor) gedeeltelijk weer opgepakt en een programma voor onderzoek geschetst.

Naast 'onderwijs en arbeid' als een thema binnen het aandachtsveld onderwijs en levenskansen, is er het thema *onderwijs en deelname aan de cultuur*. Dit thema is sterk verbonden met het werk van de Franse (cultuur)socioloog Pierre Bourdieu. Eerder in dit overzichtsartikel is al aangegeven welk soort onderzoek in Nederland in deze lijn is ontwikkeld. In ons land bestond op dit terrein trouwens een zekere traditie door de publikaties van het Sociaal en Cultureel Planbureau met beschrijvende gegevens over deelname aan culturele activiteiten in Nederland, vaak met commentaar van sociologen. Vermeldenswaard is tenslotte in dit verband het onderzoek van Gadourek naar de relatie tussen onderwijs en levensstijl (1986).

Een vierde thema dat ressorteert onder 'levenskansen' is het thema *onderwijs en gezondheid*. Ook hier is de aansluiting met de problematiek van de maatschappelijke ongelijkheid van belang. Maatschappelijke ongelijkheid kent verschillende dimensies (bezit, macht, kennis enz.) die vaak met elkaar samenhangen en zo leiden tot duurzame verschillen tussen sociale groepen in de maatschappijstructuur. Die verschillen zijn bovendien van relatief permanente aard doordat generatiegewijze steeds dezelfde sociale klassen bevoordeeld respectievelijk benadeeld worden. Dat is in het kort de theorie van de maatschappelijke reproductie. Tot voor kort bleef 'gezondheid' meestal buiten de discussie in de onderwijswetenschappen. Ten onrechte. Gezondheid en het bezit van een dergelijk schaars goed is evenzeer - en misschien nog wel meer - bepalend voor de levenskansen die mensen (soms letterlijk) geboden worden. We weten nog onvoldoende over deze verschillen in realiseringmogelijkheden op het terrein van ziekte en gezondheid. We weten ook nog onvoldoende hoe belangrijk mensen deze zaken achten, hoeveel waarde aan gezondheid wordt toegekend in vergelijking met andere maatschappelijke goederen en diensten (vgl. Suurmeyer, 1986)<sup>10</sup>. Suurmeyer (o.c.) geeft hiervan een eerste uitwerking, maar het zal duidelijk zijn dat het hier een nog praktisch onontgonnen terrein betreft, zeker de verbinding tussen onderwijs en gezondheid.

Bij de onderzoekslijn van 'onderwijs en levenskansen' past de kanttekening dat theoretisch-verklarend in het Nederlandse onderzoek nog maar weinig werk is verricht; het meeste onderzoek is beschrijvend van karakter en opzet. Over het terrein van onderwijs en levenskansen zijn, dat zal na het voorgaande duidelijk zijn, ook nauwelijks absolute uitspraken te doen. Het belang van het onderwijs op dit punt is evident, maar is tijd- en plaatsgebonden en verschilt voor diverse maatschappelijke groepen en samenlevingen. Voor onze tijd en voor onze samenleving geldt dat op de sleutelmacht van de school weliswaar het nodige valt af te dingen maar dat anderzijds de vitale positie van onderwijs als instantie die in deze maatschappij posities toewijst grotendeels intact blijft.

10. Er is bijvoorbeeld geen onderzoek gedaan naar de mogelijke relatie tussen de kwaliteit van medici (en klinieken), de sociale herkomst van hun cliënteel en de levensduur.

## LITERATUUR

- Abercrombie, N., Hill, S. & Turner, B.S. (1980). *The dominant ideology thesis*. London.
- Appel, R. (1984). *Immigrant Children Learning Dutch. Sociolinguistic and Psycholinguistic Aspects of Second-language Acquisition*. Amsterdam (diss.).
- Appel, R., Hubers, G. & Meijer, G. (1976). *Sociolinguïstiek*. Utrecht.
- Appel, R. & Wesselingh, A.A. (1987). Sociale verschillen in het onderwijs. In: P. Span et al. (red.), *Onderwijzen en Leren (pp. 77-100)*. Groningen.
- Apple, M.W. (1979). *Ideology and curriculum*. London.
- Apple, M.W. (1982). *Education and power*. London.
- Archer, M.S. (1979). *The social origins of educational systems*. London.
- Archer, M.S. (Ed.) (1982). *The sociology of educational expansion*. London.
- Atkinson, P. (1987). *Language, Structure and Reproduction. An Introduction to the Sociology of Basil Bernstein*. London.
- Beekenkamp, G.G. & Dronkers, J. (1984). De plaats van het onderwijs in de recrutering van president-directeuren. In J. Dronkers & F.N. Stokman (red.), *Nederlandse élites in beeld. Recrutering, samenhang en verandering (pp. 85-100)*. Deventer.
- Bernstein, B. (1971). *Class, Codes and Control, Vol. 1. Theoretical Studies towards a Sociology of Language*. London.
- Bernstein, B. (1977). *Class, Codes and Control, Vol. 3. Towards a Theory of Educational Transmissions (revised edition)*. London.
- Blok, H. & Saris, W.E. (1980). Relevante variabelen bij het doorverwijzen na de lagere school, een structureel onderzoek. *Tijdschrift voor Onderwijsresearch*, 5, 2, 63-79.
- Bloom, A. (1987). *The closing of the American mind*. New York.
- Boekholt, P.Th.F.M. (1978). *Het lager onderwijs in Gelderland 1795-1858*. Zutphen.
- Boekholt, P.Th.F.M. & De Booy, E.P. (1987). *Geschiedenis van de school in Nederland*. Assen.
- Bois-Reymond, M. du (1981). *Onderwijzersleven*. Nijmegen.
- Boli, J., Ramirez, F.O. & Meyer, J.W. (1985). Explaining the origins and expansion of mass education. *Comparative Education Review*, 29, 145-169.
- Boon van Ostade, A.H. (1972). Het milieu en de overgang van het basis- naar het voortgezet onderwijs in Nederland. *Sociologische Gids*, 19, 219-228.
- Booy, E.P. de (1980). *Kweekhoven der wijsheid. Basis en vervolg onderwijs in de steden van de provincie Utrecht van 1580 tot het begin der 19e eeuw*. Utrecht.
- Bosman, R. & Louwes, W. (1988). Eenouder- en tweeoudergezinnen en schoolloopbanen. *Mens en Maatschappij*, 63, 1, 5-23.
- Boudon, R. (1973). *L'inégalité des chances. La mobilité sociale dans les sociétés industrielles*. Paris.
- Bourdieu, P. (1979). *La distinction. Critique sociale du jugement*. Paris.
- Bourdieu, P. (1984). Espace sociale et genèse de 'classe'. *Actes de la recherche en sciences sociales*, 52/53.
- Bourdieu, P. & Passeron, J.Cl. (1964). *Les Héritiers. Les Etudiants et la Culture*. Paris.
- Bourdieu, P. & Passeron, J.Cl. (1970). *La reproduction. Eléments pour une théorie du système d'enseignement*. Paris.
- Bowles, S. and Gintis, H. (1976). *Schooling in Capitalist America. Educational Reform and the Contradictions of Economic Life*. New York.
- Brands, J., Egas, G., Karsten, J. & Wendrich, E. (1977). *Andere wijs over onderwijs*. Nijmegen.
- British Journal of Sociology of Education* (1986), 7, 2. Special Number 'Whatever happened to inequality?'
- Bijlsma, K.M. et al. (1981). Leerlingculturen: verzet tegen een etiket. In *Jeugdcultuur en onderwijs*. Amsterdam.
- Calcar, C. van, Soutendijk, S. & Tellegen, B. (1968). *School, milieu, prestatie*. Enschede.
- Chalmers, A. (1981). *Wat heet wetenschap?* Meppel.
- Claeys, P.C. (1972). *Gezin en schoolsucces*. Antwerpen.
- Coleman, J.S. & Campbell, E.S. (1966). *Equality of educational opportunity*. Washington.
- Coleman, J.S., Hoffer, T. & Kilgore, S. (1982). *High school achievement; public catholic and private schools compared*. New York.

- Collins, R. (1971). Functional and conflict theories of educational stratification. *American Sociological Review*, 36, 1002-1119.
- Connolly, W.E. (1983). *The terms of political discourse*. New York.
- Damme, D. van (1987). Contradicties van alfabetisering. *Comenius*, 25, 3-26.
- Diederer, J. (1983). *De keuze van een beroep. Van Jaar tot Jaar derde fase*. Nijmegen.
- Diederer, J. (1989). *Meer of minder gelijk. Over sexe-verschillen in schoolloopbaan en beroepskeuze en over de gedeeltelijke konditionering van die sexe-verschillen door sociaal milieu en schoolprestatie*. Nijmegen.
- Dronkers, J. (1986). Onderwijs en sociale ongelijkheid. In J.A. van Kemenade et al. (red.) *Onderwijs: bestel en beleid*. Deel 2A, pp.42-151. Groningen.
- Dronkers, J. (1989a). Veranderingen in samenleving en economie en de groei van de onderwijsdeelname sinds 1960. *Comenius*, 9, 1, 47-72.
- Dronkers, J. (1989b). Schoolkenmerken en individuele factoren. In P. Vogel et al. (red.) *Onderwijs: keuzen en kansen*. Muiderberg (ter perse).
- Dronkers, J. & de Jong, U. (1978). Jencks en Fägerlind op zijn Hollands. Een aanzet tot de studie van de relaties tussen milieu, intelligentie, onderwijs, beroep en inkomen. *Sociologische Gids*, 25, 1, 4-30.
- Dronkers, J., de Graaf, P. de Haan, D. Peek, J. en Swanborn, P. (1982). Veranderingen in ongelijke onderwijskansen. *Comenius*, 5, 97-113.
- The Education of Elites* (1977). Block V, Culture and Class. Open University Course E 202: Schooling and Society. Milton Keynes.
- Faasse, J.H., Bakker, B., Dronkers, J. & H. Schijf (1986). Voor en na de Mammoetwet. *Mens en Maatschappij*, 61, 75-94.
- Frijhoff, W.Th.M. (1981). *La société néerlandaise et ses diplômés, 1575-1814*. Amsterdam.
- Frijhoff, W.Th.M. (1983). Van onderwijs naar opvoedend onderwijs. Ontwikkelingslijnen van opvoeding en onderwijs in Noord-Nederland in de achttiende eeuw. In *Werkgroep Achttiende Eeuw, Onderwijs en opvoeding in de achttiende eeuw* (pp. 3-40). Amsterdam.
- Gadourek, I. (1986). Verschillen in levensstijl bij de voornaamste geledingen in Nederland. In Th. Mensen & G.W. Meijnen (red.) *Sociale Stratificatie en Onderwijs* (pp. 63-74). Lisse.
- Ganzeboom, H. & de Graaf, P. (1983). Beroepsmobiliteit tussen generaties in Nederland in 1954 en 1977. *Mens en Maatschappij*, 58, 28-52.
- Ganzeboom, H. et al. (1987). Intergenerationele klassenmobiliteit in Nederland tussen 1970 en 1985. *Mens en Maatschappij*, 62, 1, 17-43.
- Geurts, J. & Tesser, P. (1976). *Werkende jongeren en hun onderwijs*. Nijmegen.
- Giroux, H. (1983). *Theory and resistance in education*. London.
- Girls and women in education. A cross-national study of sex inequalities in upbringing and in schools and colleges* (1986). Paris (OECD).
- Glebbeek, A. & Mensen, Th. (1986). Het onderwijs tussen meritokratie en reproductie. In Th. Mensen & G.W. Meijnen (red.) *Sociale Stratificatie en Onderwijs*. (pp. 111-132). Lisse.
- Goldthorpe, J.A. (1980). *Social mobility and class structure in modern Britain*. Oxford.
- Graaf, P.M. de (1987). *De invloed van financiële en culturele hulpbronnen in onderwijsloopbanen*. Nijmegen.
- Habermas, J. (1983). *Moralbewusstsein und kommunikatives Handeln*. Frankfurt.
- Halsey, A.H., Heath, A. & Ridge, J.M. (1980). *Origins and Destinations. Family, Class and Education in modern Britain*. Oxford.
- Harbers, J. (1986). *Sociale wetenschappen en hun speelruimte*. Groningen.
- Heek, F. van, et al. (1968). *Het Verborgene Talent. Milieu, Schoolkeuze en Schoolgeschiedenis*. Meppel.
- Heek, F. van, Vervoort, C.E. & Weeren, P. van (1973). Nogmaals: Verborgene talenten? *Sociologische Gids*, 4, 309-316.
- Hoof, J.J. van (1987). *De arbeidsmarkt als arena*. Amsterdam.
- Hoof, J.J. van & Dronkers, J. (1980). *Onderwijs en Arbeidsmarkt*. Deventer.
- Hövels B. & Peschar, J.L. (1985). *Tussen wal en schip. Een onderzoeksprogramma op het terrein van onderwijs en arbeids-situaties*. Nijmegen.
- Huls, H.A. (1982). *Taalgebruik in het gezin en sociale ongelijkheid. Een interaktioneel-sociolinguïstisch onderzoek*. Nijmegen.

- Idenburg, Ph.J. (1961). Onderwijs en welvaart. *Economisch-Statistische Berichten*, mei, 472-478.
- Idenburg, Ph.J. (1962). *Het ideaal van de optimale ontwikkeling der talenten en de pedagogische structuur van het schoolwezen*. Groningen.
- Idenburg, Ph.J. (1964). *Schets van het Nederlandse schoolwezen*. Groningen.
- Idenburg, Ph.J. (1967). *Het schoolwezen als economisch verschijnsel*. Amsterdam.
- Idenburg, Ph.J. (1971). *Theorie van het onderwijsbeleid*. Groningen.
- Jackson, L.A. (1974). The Myth of Elaborated and Restricted Code. *Higher Education Review (Spring 1974)*, pp.65-81.
- Jencks, Chr. et al. (1972). *Inequality. A reassessment of the effect of family and schooling in America*. New York.
- Jong, M.J. de (1987). *Herkomst, kennis en kansen. Allochtone en autochtone leerlingen tijdens de overgang van basis- naar voortgezet onderwijs*. Rotterdam.
- Jungbluth, P. (1982). *Docenten over onderwijs aan meisjes: 'positieve discriminatie' met een dubbele bodem*. Nijmegen.
- Jungbluth, P. (1985). *Verborgene differentiatie*. Nijmegen.
- Kalmijn, M. & Batenburg, R. (1986). Reproductie van cultureel en economisch kapitaal op een traditioneel en een Montessorilyceum. *Tijdschrift voor Onderwijsresearch*, 11, pp. 149-163.
- Kalmijn, M. & Swanborn, P.G. (1986). Van der Wolf's onderzoek naar schooluitval. Verslag van een secundaire analyse. *Pedagogische Studien*, 63, 435-444.
- Karsten, S. (1986). *Op het breukvlak van opvoeding en politiek. Een studie naar socialistische volksonderwijzers rond de eeuwwisseling*. Amsterdam.
- Kemenade, J.A. van & Kropman, J.A. (1972). Verborgene talenten? Kritische kanttekeningen bij een onjuiste interpretatie. *Sociologische Gids*, 19, 219-228.
- Kley, P. van der (1983). *Zeg na jij! Over regels, patronen, beurten, selectie en reproductie in het lager onderwijs*. Purmerend.
- Knippenberg, H. (1986). *Deelname aan het lager onderwijs in Nederland in de negentiende eeuw*. Amsterdam.
- Koot, W. & Rath, J. (1987). *Ethnicity and emancipation. International Migration*, 25/4, 426-440.
- Koster, K.B. (1987). Reken- en wiskunde-onderwijs in de basisschool. In J.J. van Kuyk (red.). *Basisvorming in de basisschool*. (pp.-44-52). Tilburg.
- Kreft, G.G. (1987). *Models and methods for the measurement of school effects*. Amsterdam.
- Kruithof, B. (1980). De deugdzaame natie. *Symposion*, 2, 22-37.
- Laarhoven, P. van, Bakker, B., Dronkers, J. & Schijf, H. (1986). Some aspects of school careers in public and non-public primary schools. *Tijdschrift voor Onderwijsresearch*, 11, 83-96.
- Laarhoven, P. van, Bakker, B., Dronkers, J. & Schijf, H. (1987). Richting van de school en schoolloopbanen in het voortgezet onderwijs. *Tijdschrift voor Onderwijsresearch*, 12, 23-40.
- Lagerweij, N.A.J. (1986). De dramatische strijd om de Nederlandse middenschool. In W. Wielemans (red.) *Vernieuwing in het secundair onderwijs. Een Europese situering*. (pp. 63-80). Leuven.
- Leeman, Y. (1985). Onderwijs en etnische verhoudingen. In J.L. Peschar & A.A. Wesselingh (red.). *Onderwijssociologie, een inleiding* (pp. 417-435). Groningen.
- Lenders, J. (1988). *De burger en de volkschool. Culturele en mentale achtergronden van een onderwijshervorming in Nederland 1780-1850*. Nijmegen.
- Lenders, J., Mazeland, H. & Nieuwstadt, M. van (1978). *Taal, arbeidersklasse en ongelijkheid. Een kritiek op de theoretische en praktische veronderstellingen van de sociolinguïstiek*. Amsterdam.
- Leune, J.M.G. (1976). *Onderwijsbeleid onder druk. Een historisch-sociologisch onderzoek naar het opereren van lerarenverenigingen in het Nederlandse onderwijsbestel*. Groningen.
- Luijckx, R.A., Doorne-Huiskes, A. van & Ultee, W.C. (1986). Het verband tussen onderwijs en beroepshoogte voor vrouwen en mannen in Nederland in 1979. *Mens en Maatschappij*, 61, 21-33.
- Marx, E.C.H. (1975). *De organisatie van scholengemeenschappen in onderwijskundige optiek*. Groningen.
- Matthijssen, M.A.J.M. (1971). *Klasseonderwijs. Sociologie van het onderwijs*. Deventer.
- Matthijssen, M.A.J.M. (1975). Sociale ongelijkheid van onderwijs kansen; vijf interpretaties van sociale werkelijkheid. *Sociologische Gids*, 6, 427-445.



- Meester, A.C. & De Leeuw, J. (1983). *Intelligentie, sociaal milieu en de schoolloopbaan*. Leiden.
- Meijers, F.J.M. (1983). *Van Ambachtsschool tot LTS*. Nijmegen.
- Meijnen, G.W. (1977). *Maatschappelijke achtergronden van intellectuele ontwikkeling*. Groningen.
- Meijnen, G.W. (1984). *Van zes tot twaalf*. Harlingen.
- Meijnen, G.W. (1988). Het rendement van cultureel kapitaal (Oratie Univ. van Amsterdam). *Comenius*, 31, 8, 323-342.
- Mullard, C. (1986). Pluralisme, etnicisme en ideologie: implicaties voor een transformatieve pedagogiek. *Comenius*, 6, 4, 407-431.
- Oakes, J. (1985). *Keeping Track. How schools structure inequality*. New Haven.
- Olson, M. (1971). *The logic of collective action*. Cambridge Mass.
- The Oxford Review of Education* (1987), 13. Special Issue, 'Plowden: history and prospect'.
- Parkin, F. (1971). *Class inequality and political order*. London.
- Peschar, J.L. (1975). *Milieu, school, beroep. Een achterafexperiment over de periode 1958-1973*. Groningen.
- Peschar, J.L. (1987). *Zo vader-zo zoon; zo moeder-zo dochter? Vergelijkende analyse naar de processen van statusverwerving en onderwijsmobiliteit in Nederland, Hongarije en Polen*. Lisse.
- Ramirez, F.O. & Boli, J. (1987). The political construction of mass schooling. *Sociology of Education*, 60, 1, 2-17.
- Renes, P.B. (1969). *Recruterende en selectie van leerkrachten voor het lager onderwijs in Nederland*. Groningen.
- Ringer, F. (1979). *Education in modern society*. London.
- Rupert, H.G.A. & Wolthuis, J. (1985). Ontstaan en ontwikkeling van het Nederlandse onderwijssysteem. In J.L. Peschar & A.A. Wesselingh (red.) *Onderwijssociologie, een inleiding* (pp. 279-306). Groningen.
- Rupp, J.C.C. (1969). *Opvoeding tot schoolweerbaarheid*. Groningen.
- Rupp, J.C.C. (1983). Ins and outs van de Middenschoolevaluatie-discussie. *Onderwijs en Opvoeding*, 35, 3, 150-154.
- Rupp, J.C.C. (1988). Het pedagogisch optimisme van de jaren zeventig. Terugblik op het tijdperk van de compensatie- en activeringsprogramma's, de voorlopers van het onderwijsvoorrankingsbeleid. *Comenius*, 29, 90-106.
- Rupp, J.C.C., Lange, R. de & Gernler, T. (1989). De markt van basisscholen. Het belang van de opleidingsstatus van basisscholen voor de schoolkeuze. In P. Vogel e.a. (red.), *Onderwijs: Keuzen en Kansen*. Muiderberg (ter perse).
- Schelsky, H. (1957). *Schule und Erziehung in der industriellen Gesellschaft*. Würzburg.
- Setten, H. van (1982). Opvoedend onderwijs. De vernieuwing van het Nederlandse volksonderwijs in het begin van de vorige eeuw. *Comenius*, 2, 1, 5-36.
- Setten, H. van (1987). *In de schoot van het gezin. Opvoeding in de Nederlandse gezinnen in de twintigste eeuw*. Nijmegen.
- Sewell, W.H. & Hauser, R.M. (1980). The Wisconsin longitudinal study of social and psychological factors in aspirations and achievement. In A.C. Kerckhoff (Ed.) *Research in Sociology of Education and Socialization*, Vol. I. Greenwich (Conn.).
- Sixma, H. & Ultee, W.C. (1983). Een beroepsprestigeschaal voor Nederland in de jaren tachtig. *Mens en Maatschappij*, 58, pp. 360-382.
- Suurmeijer, Th. (1986). Sociale stratificatie en gezondheid. In Th. Mensen & G.W. Meijnen (red.). *Sociale Stratificatie en Onderwijs* (pp. 49-62). Lisse.
- Swaan, A. de (1988). *In Care of the State*. Oxford.
- Tesser, P. (1986). *Sociale herkomst en schoolloopbanen in het voortgezet onderwijs*. Nijmegen.
- Teunissen, J.M.F. (1986). *Eén school, twee talen: een onderzoek naar de effecten van een tweetalig intercultureel onderwijsprogramma voor Marokkaanse en Turkse leerlingen op basisscholen te Enschede*. Utrecht.
- Tijn, Th. van (1977). Achtergronden van de ontwikkeling van het lager onderwijs en van de schoolstrijd in Nederland, 1862-1905. Een poging tot sociaal-historische analyse. In L. Box, J. Dronkers, M. Molenaar & J. de Mulder (red.), *Vrijheid van onderwijs* (pp. 25-36). Nijmegen.
- Veld, Th. (1987). *Volksonderwijs en leerplicht. Een historisch sociologisch onderzoek naar het ontstaan van de Nederlandse leerplicht 1860-1900*. Delft.



- Veneberg, A. (1982). De overleving van 'vrouwelijke' onderwijsinhouden. *Comenius*, 2, 1, 37-57.
- Vervoort, C.E. (1968). *Gezin en schoolkeuze bij handarbeiders*. Leiden.
- Vervoort, C.E. (1981). Interventiestaat, onderwijs en ongelijkheid. In P. v.d. Kley et al. (red.). *Interventiestaat en Ongelijkheid* (pp. 117-134). Nijmegen.
- Veugelaers, W. en Wendrich, E. (1986). Neokonservatisme, democratisering en pedagogiek. Recente Engelstalige publicaties over onderwijspolitiek. *Comenius*, 23, 360-375.
- Vliegen J.M. & Jong, U. de (1981). *Onderwijs in Nederland*. Monografieën Volkstelling 1971, nr. 12. Den Haag.
- Vos, J.F., Koning, P. de & Blom, S. (1985). *Onderwijs op de tweesprong*. Den Haag.
- Vroon, P. (1980). *Intelligentie. Over het meten van een mythe en de politieke, sociale en onderwijskundige gevolgen*. Baarn.
- Weber, M. (1956). *Wirtschaft und Gesellschaft*. Zweiter Halbband. Tübingen.
- Wendrich, E. (1981). De pedagogische conceptie van Antonio Gramsci. *Comenius*, 1, 101-117.
- Wesselingh, A.A. (1985a). *Onderwijs en reproductie van maatschappelijke ongelijkheid 1975-1985*. Nijmegen.
- Wesselingh, A.A. (1985b). Onderwijs tussen reproductie en emancipatie. In J.L. Peschar & A.A. Wesselingh (red.). *Onderwijssociologie, een inleiding* (pp. 343-371). Groningen.
- Wexler, Ph. (1976). *The sociology of education; beyond inequality*. Indianapolis.
- Wexler, Ph. (1987). *Social analysis of education. After the new sociology*. London.
- Willis, P. E. (1977). *Learning to labour. How working class kids get working class jobs*. Westmead.
- Wolf, J.C. van der (1984). *Schooluitval; een empirisch onderzoek naar de samenhang tussen schoolinterne factoren en schooluitval in het regulier onderwijs*. Lisse.
- Wolthuis, J. (1982). De onderwijsvernieuwing in Groningen van 1800 tot 1820 en de rol daarin van de onderwijzers en hun gezelschappen. *Info*, 13, 272-315.
- Wijnstra, J.M. (1976). *Het onderwijs aan van huis uit Friestalige kinderen: verslag van een evaluatie-onderzoek in een meertalige regio*. Den Haag.
- Zijlmans, W.C. & Kley, P. van de (1985). Een sociologie van de schoolklas. In J.L. Peschar en A.A. Wesselingh (red.). *Onderwijsociologie, een inleiding* (pp. 207-231). Groningen.

Manuscript ontvangen 28-11-1988

Definitieve versie ontvangen 3-5-1989.

## Lijst met huidige werkplek auteurs

|                           |  |
|---------------------------|--|
| H.W.C.H. van Amelsvoort   | Universiteit Twente<br>Toegepaste Onderwijskunde   |
| H. Blok                   | Universiteit van Amsterdam<br>SCO-Kohnstamm Instituut  |
| B. P.M. Creemers          | Rijksuniversiteit Groningen<br>GION  |
| M. Elshout-Mohr           | Universiteit van Amsterdam<br>SCO-Kohnstamm Instituut,<br>Instituut voor de Lerarenopleiding |
| I.E.C. Heebing-ter Woerds | Arbeidsvoorziening Nederland<br>Doetinchem   |
| F.J. van der Krogt        | Katholieke Universiteit Nijmegen<br>Pedagogische Wetenschappen<br>en Onderwijskunde          |
| J. Lowyck                 | Katholieke Universiteit Leuven<br>Departement Pedagogische Wetenschappen                     |
| J.J.G. van Merriënboer    | Open Universiteit Nederland<br>Onderwijstechnologisch Expertisecentrum                       |
| F.G.W.C. Paas             | Open Universiteit Nederland<br>Onderwijstechnologisch Expertisecentrum                       |
| J.C.C. Rupp               | Universiteit van Amsterdam<br>Amsterdamse School voor Sociaal<br>Wetenschappelijk Onderzoek  |
| W.E. Saris                | Universiteit van Amsterdam<br>Amsterdam School of Communications Reserach                    |
| J. Snippe                 | ICHTUS Hogeschool<br>Rotterdam   |
| A.A. Vermulst             | Katholieke Universiteit Nijmegen<br>Pedagogische Wetenschappen en Onderwijskunde             |
| A.A. Wesselingh           | Katholieke Universiteit Nijmegen<br>Pedagogische Wetenschappen en Onderwijskunde             |
| B. Witziers               | Universiteit Twente<br>Toegepaste Onderwijskunde   |