

---

## Verslag van het AERA-Congres 1988 te New Orleans

---

*Inleiding (W.J. Nijhof, Universiteit Twente)*

Het thema voor de jaarlijkse vergadering van de Amerikaanse Onderwijsonderzoekorganisatie AERA, "Research, Policy, and Practice", was ook dit keer weer gericht op het verbreden van haar doelgroepen. Tegelijkertijd diende dit thema om in gang gezette ontwikkelingen te continueren.

Het komt er in feite op neer dat de AERA is uitgegroeid tot een internationale jaarmarkt voor onderwijsonderzoek. Eenieder die iets te maken heeft met de programmering van onderwijsonderzoek, of met het nationale en lokale beleid inzake onderwijsonderzoek of de consequenties van onderzoek voor de onderwijspraktijk, tracht iets van zijn gading te vinden.

Het beleid van de AERA heeft er dan ook toe geleid dat al deze doelgroepen, beleidsmakers, politici, practici en onderzoekers, elkaar ontmoeten. (Of de keuze van de lokatie daaraan debet is, is niet duidelijk. Wel is duidelijk dat New Orleans, gelegen aan de monding van de Mississippi, als een krachtige magneet werkte).

Voor de onderwijsonderzoekers heeft dit tot gevolg dat zij veel minder geconfronteerd worden met de directe resultaten van onderzoek, en minder geïnvolveerd raken in discussies die de kern van de onderzoeksmethodologie en -technieken raken. Ook groeit er twijfel t.a.v. de kwaliteit en relevantie van het aanbod. Sceptis t.a.v. dit type ontwikkelingen valt bij veel onderzoekers te beluisteren. Anderzijds wordt de probleemgevoeligheid voor wat werkelijk onder beleidsmakers en practici leeft groter. Het is een dilemma dat elk jaar, ook bij de Nederlandse Onderwijsresearchdagen, terugkeert.

Meer dan 1000 sessies zijn er dit jaar georga-

niseerd. De AERA-organisatie spreekt dan ook van de grootste ontmoeting ooit georganiseerd. Opvallend was de grote deelname van Nederlandse onderzoekers. Het beleid van de overheid, van de Stichting voor Onderzoek van het Onderwijs en van onderwijsresearch-instituten om sterker te internationaliseren en de resultaten van Nederlands onderzoek in het buitenland te presenteren sorteert effect.

Naast de AERA zijn er tal van andere onderzoeksorganisaties en special interest groups, die de AERA-meting aangrijpen om hun eigen ervaringen, opvattingen en inzichten uit te wisselen. Door parallel aan de AERA symposia te organiseren kunnen deze organisaties zeer brede doelgroepen bedienen. Onderzoekers pendelen zo van de ene organisatie naar de andere en kunnen op doeltreffende wijze kennis opdoen van de meest recente ontwikkelingen op hun vakgebied. Een kroniek van de AERA moet daarom in het licht van deze organisatorische complexiteit worden gezien. Het tijdsbeeld dat geschetst wordt is een selectie van indrukken met een onvermijdelijk, sterk persoonlijk gekleurde belangstelling van de deelnemers.

Het is gebruikelijk dat de president van de AERA tijdens deze jaarlijkse ontmoeting een voordracht houdt over het gekozen kernthema. Dat thema lijkt bij voorbaat gedoemd te verdwijnen onder de grote hoeveelheid onderwerpen. Niettemin heeft Richard Shavelson getracht dit thema te belichten.

De relatie tussen onderzoek, beleid en praktijk is een hybrisch thema volgens Shavelson en roept verwachtingen op die vaak niet kunnen worden ingelost. Vooronderstellingen zijn vaak dat onderzoek niet of nauwelijks bijdraagt aan de oplossing van problemen in de onderwijspraktijk. Bij fysica zou het tegendeel vaak blijken. Persoonlijke ervaring zou van grotere betekenis zijn.

Zijns inziens zouden beleidsmakers beter beïnvloed dienen te worden, door beter te begrijpen wat zij doen. De verstaanbaarheid en begrijpelijkheid van wat onderwijsonderzoekers doen zou verbeterd kunnen worden door een helder taalgebruik toe te passen.

Onderzoekers menen vaak dat beleidsmakers en politici onredelijke en naïeve verwachtingen hebben: zij zoeken niet de ideale, maar de 'next best' oplossing voor hun problemen.

De beste bijdrage die onderzoekers volgens Shavelson kunnen leveren aan beleidsmakers is het veranderen van denkkaders (mind frames). Beleidsmakers denken vaak traditioneel over onderwijs en het onderwijsberoep. Zo beschouwen zij b.v. het onderwijsberoep als iets waarmee je geboren bent (skill), of als een persoonlijke karaktertrek, en ook wel als een professie. In het onderwijsberoep staat volgens Shavelson echter het reflectieve denken en het beslissen centraal. Onderwijsonderzoek zou meer probleemgeoriënteerd te werk moeten gaan. Daarmee bedoelt hij dat de context- en de situatiegebondenheid van de problemen centraal zouden moeten staan. Daarvoor zijn cognities (kennisstructuren) nodig die als basis dienen voor de interpretatie van onderwijs(systemen). Shavelson zoekt in feite naar gegevens en gegevensbestanden die geordend kunnen worden (zoals indicatoren van effectieve onderwijssystemen), die geclusterd kunnen worden en zich lenen voor nadere analyse, waardoor meer inzicht in samenhangen op gaat treden. Het gaat er om deze verkregen inzichten te toetsen aan de inzichten van beleidsmakers voor wie de onderzoeksresultaten uiteindelijk bedoeld zijn. Deze toetsing kan dan leiden tot alternatieve vormen van indicatorenstelsels of -systemen. Zijn poging is dus te voorkomen dat er een 'mismatch' (onjuiste koppeling) van denkkaders plaatsvindt. En deze mismatch komt herhaaldelijk voor tussen onderzoeksbeleidsmakers en practicus door het optreden van mentale verschillen. Zo willen onderzoekers weten en meten, practici maken en realiseren, en beleidsmakers beslissen. Het belangrijkste criterium voor de onderzoekers is volgens Shavelson waarheid en vrouwen dat hij waarmaakt door goed onderzoek te verrichten. De onderzoeker dient echter de implicaties van zijn onderzoek voor de praktijk te bezien, evenals de mogelijke politieke relevantie.

De belangrijkste boodschap is, dat wanneer de onderscheiden groepen hun denkkaders en verwachtingen helder expliceren, overtrokken verwachtingen van wat onderzoek vermogen worden gecorrigeerd. Onderzoek heeft een eigen specifieke kritische functie. De claim van beleidsmakers om aanbevelingen te krijgen uit

onderzoek dient gerelativeerd te worden. Daarmee wordt de positie van de onderscheiden groepen duidelijk en kan de bereidheid om naar elkaar te willen luisteren afgestemd worden.

In de nu volgende bijdragen wordt getracht een beeld te schetsen van een zevental onderwerpen, en wel:

- a) *The John Dewey Society*, een special interest group, gericht op de historische analyse van het werk van Dewey en de interpretatie daarvan door het huidige onderwijs. Als stamvader van de Amerikaanse onderwijskunde is Dewey nog steeds inspiratiebron, ook voor de Europese onderwijsonderzoeker. Biesta (R.U. Leiden) schetst zijn impressies.
- b) *Onderwijsleerprocessen*. De Klerk en Simons (K.U. Brabant) hebben uit 130 sessies een selectie gelicht van een tweetal naar hun oordeel belangrijke onderwerpen: motivatie en cognitie, en leren en transfer. Opvallend is de toenadering tussen het Amerikaanse en Europese onderzoek.
- c) *Onderwijsgedrag en lerarenopleiding*. J. Lowijck (K.U. Leuven) vraagt aandacht voor een drietal thema's binnen dit thema: lerencognities, het novice-expert perspectief, en de invloed van besparingen (een intrigerend onderwerp over de informatievoorziening van stagebegeleiders). Ten slotte laakt Lowijck de kwaliteit van het door hem gehoorde onderzoek. In het licht van Shavelsons rede een niet mis te verstane waarschuwing.
- d) *Schoolorganisatie en computers* (J. Remmers, Universiteit Twente). De invloed van computers in de schoolorganisatie is een niet te vermijden onderwerp in deze tijd. Remmers behandelt vooral het gebruik van de computer bij schooladministratie en -management, onderwijskundig computergebruik en gaat nader in op de invoeringsproblemen. Hij pleit voor meer aandacht voor effectonderzoek.
- e) *Bedrijfsopleidingenproblematiek* (J.N. Streumer, Universiteit Twente). Bedrijfsopleidingen is een veld van onderzoek dat een veel grotere traditie kent in de V.S. dan in Nederland, waar het goed en wel uit de startblokken is gekomen. Streumer schetst karakteristieken van bedrijfsopleidingen, aspecten van bedrijfsopleidingskunde, ont-

wikkelingsaspecten en de evaluatie van trainingen. Vooral het onderzoek naar de kosten-effectiviteit vraagt om nadere aandacht.

f) *Effectieve scholen* (W. G. R. Stoel, Universiteit Twente). Stoel behandelt de context van amerikaans effectonderzoek, de stabiliteit van effectiviteit, effectmaten en enkele kenmerken van effectief scholenonderzoek (zals het multilevel karakter) en de theorievorming daaromtrent. Zijn conclusie is dat een verantwoorde implementatie van effectiviteitsmaatregelen vooralsnog niet in het verschiet ligt bij gebrek aan grondige evaluatiestudies.

g) *Peilingsonderzoek* (Th. Eggen en P. Sanders, CITO). In aansluiting op het thema effectieve scholen is peilingsonderzoek een instrument bij uitstek dat de overheden de laatste jaren willen benutten om de kwaliteit van het onderwijs internationaal te meten en vergelijken.

Eggen & Sanders schetsen hun indrukken van methodologische en meettechnische problemen zoals ervaren tijdens de AERA en de meeting van de National Council on Measurement in Education (NCME). Hun conclusie is dat de produktie van werkbare procedures voorschans de overhand heeft boven de theoretische fundering daarvan. Wellicht dat Nederlands onderzoek daaraan een stevige bijdrage kan leveren in de nabije toekomst.

*The John Dewey Society (G. J. J. Biesta, Rijksuniversiteit Leiden)*

The John Dewey Society, die sinds 1987 als zogenaamde 'Special Interest Group' participeert aan het AERA-congres, verzorgde dit jaar een vijftal sessies, toegankelijk voor alle congresgangers. De bijdragen aan deze sessies, voornamelijk verzorgd door 'Dewey-scholars' en derhalve veelal handelend over Dewey of duidelijk geïnspireerd op zijn ideeën, waren overigens niet de enige indicatie dat Dewey heden ten dage in Amerika nog steeds 'leeft'. Ook in lezingen die niet door de Society waren georganiseerd, viel zijn naam geregeld. We bespreken een aantal lezingen om een indruk te geven van de rol die Dewey op dit moment in het pedagogische debat speelt dan wel wordt toebedeeld.

In een bijeenkomst met als titel 'John De-

wey and General Education' werd door D. B. Gowin vanuit een Deweyaans standpunt kritiek geleverd op de veelal geïsoleerde positie van de sociaal-wetenschappelijke onderzoeker. Dewey's nadruk op het sociale karakter van ons kennen, en zijn opvatting dat ervaring en betekenis sociaal van oorsprong zijn, worden door hem vertaald in de eis dat (sociaal-) wetenschappelijk onderzoek moet aansluiten bij de concrete problemen in de samenleving. Op basis van deze 'sociale epistemologie' beargumenteerde Gowin dat adequaat onderzoek niet alleen een 'socialisatie' van de onderzoeker in een wetenschapsgebied en in de methoden en technieken van onderzoek vereist maar dat daarnaast een 'socialisatie' in het gebied van onderzoek van het grootste belang is. Dewey's visie op wat nu exact 'concrete problemen in de samenleving' zijn, werd vervolgens toegelicht door R. Pratt. Dewey beargumenteert zijn opvatting dat 'there can be no conflict between the individual and the social' door te stellen dat *beide*, individu en samenleving op deze wijze tegenover elkaar gezet, abstracties zijn. Dit soort abstracte dualismen versluiert volgens Dewey het zicht op de werkelijke problemen tussen 'groups and other groups and individuals'. In aansluiting hierop stelde Pratt dat naar zijn mening ook het onderscheid tussen 'general education' en 'special education' een abstract dualisme is dat ons afleidt van de werkelijke problemen.

In het licht van het actuele debat over de kwaliteit van het Amerikaanse onderwijs beargumenteerde C. Collins waarom een van de discussiestukken in dat debat, M. J. Adlers 'Paidcia Proposal' (1982), niet aan Dewey opgedragen had mogen worden. Ofschoon Adler 'Deweyaans' is in zijn opvatting dat kinderen in de cultuur ingevoerd moeten worden en dat specialisatie van het onderwijs en het opdelen van kinderen in groepen niet tot een culturele eenheid maar juist tot een versnippering van de cultuur leidt, verschilt hij sterk met Dewey in de wijze waarop hij meent dat dit probleem van versnippering opgelost moet worden. Adlers visie komt er volgens Collins op neer dat allereerst bepaald moet (en ook kan!) worden welke cultuurprodukten waardevol zijn en dat invoering in de cultuur vervolgens betekent dat in het onderwijs die produkten bestudeerd moeten worden. De student die daar niet voldoende in slaagt, valt helaas buiten de boot. Het onderwijs moet dus het bestaande cul-

tuurgoed doorgeven en heeft daarmee tevens een selectiemiddel in handen. Dewey bewandelt precies op dit punt de omgekeerde weg. De cultuur bestaat niet uit vaste inhouden die overgedragen moeten worden; cultuur is het resultaat van de wijze waarop de traditie op dit moment, door alle mensen, wordt gereconstrueerd (vgl. de hierboven gemaakte opmerkingen over Dewey's 'sociale epistemologie'). Bij Dewey staat juist de gedachte voorop dat alle leden van de maatschappij mee moeten doen en mee moeten *blijven* doen. En dit staat haaks op Adlers ideeën. Collins beargumenteerde vervolgens dat de tegenstelling Adler – Dewey teruggaat op de tegenstelling tussen klassieke godsdienstige ideeën enerzijds en protestants liberalisme anderzijds. Uit de discussie die volgde op Collins' opmerking dat Dewey wel degelijk 'religious' was (en dan niet in seculier-humanistische zin), bleek overigens dat deze interpretatie van Dewey – in ieder geval bij de aanwezige toehoorders – controversieel was.

Dat er veel – en verschillend – over Dewey wordt gedacht en geschreven, werd naar voren gebracht door J. Chr. Eisele en C. Landvoigt (hoofd van een lagere school gebaseerd op Dewey's pedagogische ideeën!), onder meer met het volgende citaat: "A curious aspect of the North American educational scene is that most educationists seem to have a rather different view of what Dewey was about, along with a conviction of their privileged insight, supported irrefutably with textual evidence." (Egan, 1987). Eisele en Landvoigt demonstreerden dit met een overzicht van publikaties over Dewey uit de afgelopen twintig jaar over de rol van de school, de plaats van het individu, Dewey's opvattingen over vrouwen en zijn 'ecologische filosofie' waaruit onder meer opgemaakt moest worden dat hij de rol van de school zowel onderschat als overschat, de vrijheid van het individu zowel benadrukt als ontkent en zowel een feminist als een seksist was. Eisele en Landvoigt stelden dat veel van deze verwarring voorkomen kan worden, wanneer de diverse auteurs inzien dat zij eigenlijk niet zozeer 'over' Dewey schrijven als wel 'in antwoord op' hem. De oorspronkelijke titel van hun bijdrage ('An Analysis of the scholarly Use and Abuse of Dewey') hadden zij dienovereenkomstig dan ook gewijzigd in 'Responding to Dewey'. De veelheid van publikaties liet naar hun mening wel zien dat men bij

veel onderwerpen niet om Dewey heen kan en dat Dewey op dit moment nog steeds fungeert als inspiratiebron en gesprekspartner (overigens ook de uitgangsgedachte van het Leidse Dewey-onderzoek). Dat Dewey zo een (centrale) positie inneemt blijkt ook uit Frankels opmerking dat "to know where we stand toward Dewey's ideas is to find out, at least in part, where we stand with ourselves" (Frankel in: Westbrook, 1980, p.345).

Dat Dewey uiteraard niet op alle gebieden uitkomst kan bieden in deze tijd werd naar voren gebracht door W. en S. R. Warnick. Zij leverden kritiek op Dewey's uitgangspunt dat ieder kind een wil om te leren en zich te ontwikkelen, om te groeien en om sociaal te zijn in zich heeft. Volgens de Warnicks ontbreekt dit tegenwoordig bij kinderen. Er is geen sprake meer van een 'vonk' die eenvoudigweg tot een 'vlam' gebracht kan worden. De docent zal zelf een vonk moeten veroorzaken. Dientengevolge pleitten zij voor meer nadruk op de taak van de docent dan dat zij bij Dewey menen aan te treffen. In hun lezing lieten ze een minder bekende kant van Dewey zien, door gebruik te maken van zijn gedichten.

De 'John Dewey Lecture' van 1988, getiteld 'The Two Cultures Revisited: A Deweyan Statement', werd verzorgd door F. Crews en vormde de afsluiting van de bijdrage van de 'John Dewey Society' aan het Aera-congres. Centraal stond de vraag of het inderdaad zo is dat, zoals C. P. Snow in 1959 had beweerd, de 'natuurwetenschappelijke cultuur' en de 'literaire cultuur' twee voor altijd gescheiden werelden zijn waarin de 'natuurwetenschappelijke cultuur' veruit superieur is. Een belangrijk punt van kritiek op Snows these is dat hij lijkt te vergeten dat ook de natuurwetenschapper onderwijs en opvoeding moet genieten en dat dat toch niet geheel kan en moet bestaan uit natuurwetenschap (waarbij de actuele vraag natuurlijk is welke literaire 'klassiekers' tot de opvoeding zouden moeten behoren). Ofschoon Crew met name in de sociale wetenschappen een tendens tot het naar elkaar toegroeien van beide culturen meende te kunnen signaleren, meende Crew toch dat Dewey's optimisme – Dewey pleit voor een ongedeelde cultuur – zeker op te korte termijn onterecht is. Volgens Crew is het probleem op dit moment eigenlijk ook niet zozeer hoe beide culturen één kunnen worden, als wel hoe de literaire cultuur behouden kan blijven.

## Literatuur

- Egan, K., On Learning. A Respons to Floden, Buchmann and Schwillle. *Teachers College Record*, 1987, 88.
- Westbrook, R., Dewey's Truth. *History of Education Quarterly*, 1980, 20.

*Onderwijsleerprocessen (L. F. W. de Klerk en P. R. J. Simons, Katholieke Universiteit Brabant)*

In het kader van divisie C (die betrekking heeft op leren en instructie) hebben ruim 130 papieren postersessies plaatsgevonden. Wij volstaan hier met het noemen van de meest opmerkelijke zaken die zich binnen deze divisie hebben afgespeeld.

Evenals vorig jaar is veel informatie gepresenteerd in de vorm van posters. Nu deze vorm van presenteren – naast de al langer bestaande ronde tafel bijeenkomsten – wat meer ingeburgerd raakt, lijkt de belangstelling ervoor toegenomen te zijn. Voor zowel de onderzoeker als de congresganger wordt zo een gelegenheid geschapen om met elkaar van gedachten te wisselen over het onderzoek. Helaas moet worden geconstateerd dat niet alleen de kwaliteit van de paper-presentaties maar ook die van de posters (zowel qua lay out als qua inhoud) zeer sterk varieerde. Aan onleesbare en onbegrijpelijke posters ontbrak het bepaald niet.

Een ander, meer inhoudelijk opmerkelijk feit is dat de belangstelling voor computer ondersteund onderwijs als zodanig wat lijkt af te nemen. De ontwikkelingen binnen dit terrein gaan vooral in de richting van het ontwerpen van intelligente tutors (waarbij veelbelovende plannen maar nog weinig praktische resultaten getoond werden). Ook is er veel belangstelling voor de gebruiksmogelijkheden van de interactieve video (en beeldplaat). In diverse voordrachten werd aannemelijk gemaakt (en soms gedemonstreerd) hoe groot de mogelijkheden van dit medium zijn. De kracht van dit medium schuilt vooral in het feit dat daardoor zowel aan cognitieve als aan motivationele aspecten van leren aandacht wordt besteed.

### *Motivatie en cognitie*

De gedachte dat motivatie, de wil om doelen te bereiken, en emotie enerzijds en cognitie en

zelfregulatie anderzijds ten nauwste met elkaar samen hangen kon in menig voordracht worden beluisterd. Zo besteedde Scott Paris in zijn 'invited address' veel aandacht aan de fusie tussen 'will and skill'. De kern van zijn verhaal was dat het effect van leerstrategieën valt of staat met de vraag of de leerling het nut ervan inziet. Het is van belang te weten welke doelen de leerlingen voor zichzelf nastreven, zowel op de korte als op de lange termijn, welke opvattingen zij hebben over hun eigen mogelijkheden en waaraan zij hun prestaties toeschrijven. De Amerikanen toonden veel waardering voor het door Paris geboden overzicht van het onderzoek. Menig Europeaan daarentegen was van mening dat deze bijdrage weinig 'nieuws' bevatte en constateerde een absoluut gebrek aan belangstelling voor het – inmiddels uitgebreide – Europese onderzoek op dit terrein.

De hierboven gesignaleerde interesse voor 'persoonlijke' doelen bleek ook uit bijdragen van de groep van Bereiter (Ontario). Uit het onderzoek kwam o.a. naar voren dat er opmerkelijke individuele verschillen bestaan ten aanzien van het niveau waarop leerlingen zichzelf doelen stellen. Sommige leerlingen willen uitsluitend een opgedragen taak voltooien (task completion goals), terwijl andere ernaar streven om de leerstof te begrijpen met het doel om het geleerde onmiddellijk te kunnen toepassen (immediate problem solving). Ook zijn er leerlingen die iets leren in de verwachting of hoop dat het geleerde later eens van pas kan komen (extended problem solving). Deze doel-niveaus blijken nauw samen te hangen met de leerprestaties.

Ook Ames presenteerde onderzoek naar de effecten van verschillende typen van 'persoonlijke' doelstellingen. Zij maakte daarbij een onderscheid tussen prestatiegerichte en stofgerichte doelstellingen. Tot de eerste, zogenaamde 'performance goals' behoort bijvoorbeeld de wil om hoge cijfers te halen. Een voorbeeld van de tweede, zogenaamde 'mastery goals', is de wil om bepaalde stof te begrijpen. Niet alleen bleek ook hier dat 'will and skill' met elkaar samenhangen, maar tevens bleek dat de persoonlijke doelen via instructie te beïnvloeden zijn, waarbij de resulterende leerstrategie in belangrijke mate de kwaliteit van het leerproces bepaalt. Een interessant symposium was dat van Corno e.a. op het gebied van 'conatieve' strategieën. Hieronder

worden verstaan die strategieën die gericht zijn op de regulatie van de wil. Via regulatie van deze conatieve factoren (i.e., via regulatie van de emotie, van de aandacht, van de persoonlijke doelstellingen, van plannings-activiteiten e.d.) kan het leerproces worden beïnvloed.

Verskillende onderzoekers hebben gewezen op het belang van contextualisering. Het aanleren van leerstrategieën dient zoveel mogelijk te gebeuren binnen de context waarin ze ook gebruikt moeten worden. Afhankelijk van de gekozen strategie dient dan een keuze gemaakt te worden met betrekking tot de instructiemethode. Hiervoor komen o.a. in aanmerking: (a) model-leren in combinatie met hardop denken; (b) het zelf laten uitleggen en verklaren van de te leren strategie(ën); (c) coöperatief leren en (d) een vorm van 'reciprocal teaching' waarbij de verantwoordelijkheid voor het leren geleidelijk verschuift van de docent naar de leerling. Via onderzoek moet worden nagegaan welke methode in een gegeven situatie het meest geschikt is.

### *Transfer*

Een ander onderwerp dat – evenals vorig jaar – betrekkelijk veel aandacht trok was 'transfer'. In verschillende bijdragen kon men een optimistische toon beluisteren wanneer het ging over de mogelijkheden om transfer-effecten te bereiken. Salomon heeft getracht aannemelijk te maken dat er twee wegen zijn om dit onderwijsdoel te bereiken: de 'high road' en de 'low road'. De 'low road' heeft betrekking op het veelvuldig herhalen van taken in allerlei situaties waardoor het geleerde min of meer automatisch wordt opgeslagen in het geheugen van de leerling. De 'high road' heeft betrekking op een diepgaand verwerken van de leerstof en bij voorkeur door 'mindful abstraction'. Hiervan is bijvoorbeeld sprake als de leerling bij het lezen van een beschrijving van een situatie of verschijnsel tracht te achterhalen om welke algemene gedachte, regel of principe het in feite gaat, los van de gegeven context. Salomon heeft via onderzoek met behulp van soortgelijke 'probleem-oplos'-taken als het klassieke radiatieprobleem van Duncaker, aangetoond dat de 'low road' leidt tot veel transfer maar over een korte afstand (dat wil zeggen dat de toepassing veelal beperkt blijft tot taken die niet veel afwijken van de oorspronkelijke leertaak) en dat de 'high road'

gewoonlijk leidt tot weinig transfer, maar dat deze zeer verrijkend kan zijn. Wil men 'verre' transfer bevorderen dan is het dus volgens Salomon van belang dat de leerling leert 'mindful' bezig te zijn met leerstof.

Frederiksen is van mening dat 'bewustwording' een belangrijke conditie is voor het bevorderen van transfer. Dit zou kunnen worden bereikt door tijdens de instructie veel uitleg te geven. Resnick uitte hierover enige twijfels. Volgens haar wordt op deze wijze wel het inzicht, maar nog niet de transfer bevorderd. Transfer wordt o.a. gestimuleerd door leerlingen te confronteren met problemen. Hierdoor worden zij gedwongen te zoeken naar relevante informatie in hun geheugen en het eerder geleerde toe te passen. Het coöperatief leren oplossen van problemen is hierbij aan te bevelen. Dit leidt via discussie tot (na)denken en reflectie over de aanpak. Overigens geldt dat de keuze van de aanpak of strategie in belangrijke mate afhankelijk is van het specifieke domein waarin de taak zich bevindt en dat de relatie tussen domeinkennis en strategiekennis uitermate complex is. Dit maakt het formuleren van gebruiksaanwijzingen vooralsnog tot een hachelijke zaak.

Tot slot merken we op dat in de papers over motivatie en cognitie, leren en transfer, domein-specifieke kennis en leerstrategieën de naam van Vygotskij veelvuldig werd genoemd. De aan hem ontleende concepten lijken theorievorming en het daarop gebaseerde onderzoek op het gebied van onderwijsleerprocessen, zoals dat de laatste jaren in de Verenigde Staten plaats vindt, in belangrijke mate te ondersteunen. Wellicht vindt een soort kruisbestuiving plaats, waardoor met vrucht verder gewerkt kan worden aan het ontwerpen van optimale onderwijsarrangementen, hetgeen het uiteindelijke doel is van afdeling C van de AERA.

### *Onderwijsgedrag en lerarenopleiding* (J. Lowyck, Katholieke Universiteit Leuven)

*De dominantie van de lerarencognities*  
Wellicht beïnvloed door het thema van de AERA Conferentie 1988, met name 'Research, Policy, and Practice', is in heel wat bijdragen naar een koppeling gezocht van de studie van lerarencognities met de opleiding van onderwijsgeevenden. De uitdrukkelijke aandacht voor de genoemde cognities hangt

dan weer samen met de vigerende beleidsop-  
ties ten behoeve van de kwaliteitsverbetering  
van de leraren door het verhogen van een  
didactisch vertaalde vakinhoudelijke kennis.

Verwondering wekt dit niet, gezien de 'back  
to the basics'-beweging die volop aan de gang  
blijft. Het is in dit verband niet toevallig dat de  
'Spencer Foundation' een uitgebreid project  
aan de Stanford University heeft toegekend,  
dat door L. Shulman wordt geleid en 'de ont-  
wikkeling en de verrijking van de vakinhoudelijke  
kennis van jonge leerkrachten' tot  
voorwerp heeft.

Het door Shulman gepropageerde concept  
van de 'pedagogical content knowledge' bleek  
niet alleen in vele symposia en paper-sessies  
rond lerarengedrag en opleiding erg domi-  
nant, maar het is ook doorgedrongen tot het  
domein van de testconstructie en van de be-  
oordeling van de onderwijsloopbaan. In ande-  
re woorden: het nogal geïsoleerde studie-  
object van de lerarencognities dat in de vorige  
jaren op de AERA is gepresenteerd, neemt  
niet alleen kwantitatief toe (twaalf symposia/  
paper-sessies), maar vertoont tevens een inkt-  
vlek-effect. Blijkbaar is men de overtuiging  
toegedaan dat de in didactisch handelen ver-  
taalde vakinhoudelijke deskundigheid van de  
onderwijsgevende de voorwaarde is voor de  
competitieve kracht van het onderwijs en van-  
daar van de hele Amerikaanse maatschappij.  
De vraag is evenwel, of de proliferatie van  
deze thematiek ook gepaard gaat met een zoe-  
ken naar theoretische fundering.

#### *Het novice-expert perspectief*

De poging om het onderwijzen te verhelderen  
door expertengedrag te confronteren met dit  
van beginners, kwam aan bod in een Symposi-  
um getiteld 'Expert, novice, and postulant  
teachers: perspectives and research', en dat  
verslag uitbracht over het onderzoek dat onder  
de leiding van D. Berliner aan de University  
of Arizona wordt verricht. Opvallend in  
deze is evenwel, dat het gemeenschappelijke  
meer in de formele aanpak van de expert-novi-  
ce methodologie, dan in een achterliggende  
theorie of een communiaal onderzoeksdoel ge-  
vonden diende te worden. Bovendien vertoont  
het onderzoeksdesign de kenmerken van een  
sterke reductie van de te onderzoeken realiteit.  
Korte lesmomenten, vooraf bepaalde rolge-  
dringen, gestimuleerde herinnering op door  
de onderzoeker bepaalde punten, e.d. roepen

heel sterk de microteaching benadering van de  
jaren zestig op, zij het dan dat niet zo zeer deel-  
gedragingen, dan wel deelcognities het studie-  
object vormen. Alhoewel de titel van het sym-  
posium ook perspectieven beloofde, is dit niet  
aan de orde geweest. Er is niet ingegaan op het  
probleem van de transitie van novice naar  
expert (hoe kan dat ook met zo'n gefragmen-  
teerde aanpak, zonder longitudinale studies)  
noch op de blijvende vraag naar de waarde  
van de criteria om een expert 'expert' te noe-  
men. Het gaat meestal om onderwijsgevenden  
met een zekere ervaring, die op grond van erg  
vage indicaties ook 'goede' leerkrachten wor-  
den genoemd. Hier zouden dan meer onder-  
wijsfilosofische aspecten een rol spelen, en dit  
is blijkbaar niet de bedoeling van onderzoe-  
kers die zich exclusief op de empirie beroepen  
voor het achterhalen van de criteria voor 'ef-  
fectief' onderwijzen.

#### *De invloed van de besparingen*

Een frappant gegeven is, dat de besparingen in  
onderwijs en onderzoek, soms aanleiding ge-  
ven tot nieuwe voorstellen om de gestelde  
doelen met minder middelen te bereiken. Zo  
heeft C. Clark in zijn bijdrage 'What veteran  
teachers can learn from beginners' een plei-  
dooi gehouden om het practicum tijdens de  
opleiding te benutten als innovatiebron voor  
de begeleidende klasseleerkrachten. Hiermee  
wil hij het eenrichtingsverkeer van begeleider-  
student ombuigen in een tweerichtingsver-  
keer. Immers, de begeleider kan juist door de  
komst van een aanstaande onderwijsgevende,  
op de hoogte gebracht worden van nieuwe  
informatie en strategieën en zijn handelen  
daarop afstemmen. Clark vergeleek deze  
aanpak met de artsenopleiding, waar stage-  
mentoren blijkbaar de nieuwste medische  
informatie via hun stagiairs verkrijgen. Deze  
nogal speculatieve wijze van probleemoplos-  
sen is door D. Berliner gekritiseerd. Hij ver-  
wijst naar de ongelijke rollen, de twijfel over  
het 'nieuwe' dat een opleiding biedt, en het  
tekort aan deskundigheid van een aanstaande  
leerkracht om de nieuwe informatie in de con-  
text van de concrete praktijk te plaatsen.

#### *De kwaliteit van het onderzoek*

Het onderzoek dat in de vele symposia en  
paper-sessies is gepresenteerd, vertoont een  
aantal negatieve kenmerken, die mede de  
waarde van de uitkomsten hebben bepaald.

Het design is meestal zeer beperkt: erg kortlopend, beperkte vraagstelling, proliferatie van gekende paradigma's zonder falsificatie ervan, betwistbare interpretatietechnieken, en het tekort aan stopregels voor descriptief onderzoek.

De meeste presentaties beperken zich tot een weergave van een aantal 'feiten' van het onderzoek. Zeldzaam was het verbreden van de horizon, door inzicht te geven in de achterliggende theoretische kaders of door de verkregen uitkomsten in een breder model of een theorie te plaatsen. Ook de discussianten gingen meestal in op de interne tekortkomingen van de gepresenteerde onderzoeken, veeleer dan de uitgangspunten te bevragen.

Zoekt men de redenen voor de genoemde tekorten, dan kan de strategie voor projectverwerving en -rapportering ten dele verantwoordelijk zijn. Reeds eerder is gewezen op bepaalde dominante richtingen in het onderzoek, zoals o.m. het 'pedagogical knowledge' paradigma. Het belang ervan wordt niet zo zeer aangetoond door onderzoek, maar ontstaat door de afstemming van het thema op achterliggende, maatschappelijk bepaalde noden en verwachtingen.

Voorts ontstaat de neiging om de proefpersonen niet te selecteren vanuit methodologisch verantwoordelijke criteria, maar omwille van hun vlotte toegankelijkheid. Zo gaat het meestal om eigen 'student teachers' die tijdens de opleiding worden 'bevraagd'. Een voorbeeld hiervan is de studie van G. Morine 'Preservice teachers' conceptions of content and pedagogy: graduate vs. undergraduate'. Het gevaar hiervan is dat men de vraagstelling aanpast aan de onderzoekscontext, en niet omgekeerd.

Verder is opgevallen, dat er een typische traditie van gegevensverwerking is ontstaan: men brengt een aantal interpretaties in het materiaal aan, wat dan aansluitend wordt geïllustreerd met 'werkelijke' uitspraken van onderwijsgeevenden. Illustratief hiervoor was de presentatie van J. Shulman en J.A. Colbert 'Blue freeways: traveling the alternative route with Big-City teacher trainees'. Er wordt echter zelden aangetoond op grond waarvan de selectie van uitspraken plaatsvond, hoe men daarbij methodologisch is te werk gegaan, en hoe de verkregen informatie gevalideerd kan worden. De complexiteit van het werken met zelfrapportering blijkt in vele gevallen niet eens onderkend te zijn.

Opvallend is eveneens, dat het meestal om descriptieve studies gaat. Daarin schuilt het gevaar van een niet eindigende exploratie van alle mogelijke aspecten van onderwijzen, leraaropleiding en nascholing. De zwakte van de theorievorming, de uiteenlopendheid van de achtergrondmodellen en van de begrippen, de korte looptijd van het onderzoek en het daarmee samengaande tekort aan een prioriteitenbepaling beginnen te drukken op de kwaliteit van het onderzoeksdomein zelf. 'Onderzoek' is dikwijls een verkenning geworden, het opdoen van praktijkinspiratie door flarden uit de subjectieve ervaringswereld van (aanstaande en beginnende) leerkrachten in elkaar te weven.

Het vrije initiatief en de flexibiliteit van de Amerikaanse onderzoekers vertonen echter ook nevenverschijnselen: de verkaveling, het korte-duur effect, de nervositeit, het tekort aan historie en perspectief, het redundante, het accumulatieve en de proliferatie van de subjectieve impressies.

#### *Schoolorganisatie en computers (J. L. M. Remmers, Universiteit Twente)*

De problematiek van schoolorganisatie en computers is in twee deelgebieden te verdelen. Enerzijds is er het gebied van computertoepassingen ten behoeve van de schooladministratie en -management, anderzijds is er het gebied van het onderwijskundig computergebruik. Ik heb op beide terreinen papersessies bijgewoond, waarbij het accent op het onderwijskundig computergebruik lag. De centrale vraag was hierbij: hoe kunnen deze vormen van computergebruik het beste worden ingevoerd en wat zijn de gevolgen voor de schoolorganisatie?

#### *Schooladministratie en -management*

De papers op het gebied van schooladministratie en -management concentreren zich op drie onderwerpen. Het eerste onderwerp richt zich op beschrijvingen van opbouw en mogelijkheden van softwarepakketten. Naast de beschrijving van de administratieve onderdelen die geautomatiseerd kunnen worden en de ondersteuning van besluitvorming en planning in een schoolorganisatie, worden de voordelen met betrekking tot snelheid en volledigheid in het zonnetje gezet. Toch lijkt voor



verschillende sprekers dit alleen als argument nog niet voldoende overtuigingskracht te hebben. Zij starten dan met een informaticanalyse van scholen om op basis daarvan alsnog de computer als een krachtig instrument voor ondersteuning aan te wijzen.

De papers blijven steken in claims. Er wordt geen verslag gedaan van onderzoeken naar de mate van feitelijk gebruik of de effectiviteitsverbetering van de besluitvorming. Ik ben ook geen presentaties tegengekomen waar de effecten op de taakhoud en taakuitvoering van schoolleiders worden onderzocht. We zien hier dat het onderwijskundig onderzoek op dit terrein duidelijk achterloopt ten opzichte van het organisatie- en bedrijfskundig onderzoek. Nederlandse onderzoeken, zoals het SCHOLIS-project waarin wel onderzoek naar de effecten wordt gedaan, lijken eerder voor dan achter te lopen.

#### *Onderwijskundig computergebruik*

Op dit terrein zien we nog al wat papers die zich in feite beperken tot een reclamepraatje voor een bepaald softwareprogramma. Vervolgens blijkt dat de ontwikkelproblematiek van onderwijskundige software een ander belangrijk thema is, met vragen als: Aan welke eisen moeten programma's voldoen en hoe kan het ontwikkelproces zo gepland worden dat tijd- en budgetoverschrijdingen geminimaliseerd worden. Er werd bijvoorbeeld een presentatie van het POCO-project gehouden. Als aanbeveling wordt gewezen op prototyping als ontwikkelprocedure voor onderwijskundige software, of het belang van zo weinig mogelijk personeelwisselingen in het ontwikkelteam. De indruk ontstond dat deze taken nog wel belangrijk zijn, maar dat er een accentverschuiving optreedt in de richting van onderzoek naar de gevolgen van en invoeringsstrategieën voor een brede invoering van computers. Op de invoeringsproblematiek wordt hieronder verder ingegaan.

Met betrekking tot de gevolgen van het onderwijskundig computergebruik is een aantal onderwerpen te noemen: de hoeveelheid computergebruik door leerkrachten in de klas; de veranderende rol van de leerkracht door het gebruik van de computer in de klas; en veranderingen op leerlingniveau.

De hoeveelheid computergebruik wordt veelal door schooldistricten onderzocht. De teneur in de presentatie is dat men het gebruik

nog tegen vindt vallen, waarbij het niveau van computergebruik nog vergelijkbaar is met de resultaten van Henry J. Becker uit 1984. 'How technology affects teaching' is de vraag die onder ander gesteld wordt door het Educational Technology Center (Cambridge, Massachusetts). De belangrijkste verandering is dat de rol van de leerkracht verschuift van kennisoverdrager naar begeleider van leerprocessen. Een bijeffect van het computergebruik door leerkrachten is de taakverzwaring. Deze tendenzen zijn ook in Nederland te zien.

Op leerlingniveau wordt enerzijds onderzocht wat de effecten van het computergebruik zijn op de leerresultaten, anderzijds staat de motivatie van de leerlingen in de belangstelling. Op basis van vrij beperkte onderzoeken qua omvang (leerpsychologische experimenten) komen tegenstrijdige resultaten. In het algemeen wordt er wel een hogere motivatie bij leerlingen door computergebruik vastgesteld.

#### *De implementatieproblematiek*

Het laatste onderwerp handelt over de aspecten rondom de invoering van onderwijskundig computergebruik. Er waren twee sessies gewijd aan proefprojecten voor intensief onderwijskundig computergebruik. In beide projecten werden scholen door een computerfabrikant voorzien van een ruime hoeveelheid hardware en randapparatuur. In een van de sessies werd een video vertoont van een lagere school met een lokaal met 70 computers waarin een even grote groep leerlingen door een team van 5 leerkrachten werd begeleid. Het onderzoek verbonden aan dit project wordt geleid door Eva Baker van het UCLA uit Los Angeles en richtte zich in het eerste projectjaar op de effecten van computergebruik op de leerresultaten. In een ander project werd de effectiviteit van de invoeringsstrategie onderzocht. Verder dan een beschrijving van de activiteiten die waren uitgevoerd kwam men echter niet. De overige sessies die handelden over invoeringsstrategieën werden meestal gepresenteerd door districtsmedewerkers, waarvan sommige met vuur en vlam hun beleid aan de man probeerden te brengen. De basis voor deze plannen was eerder intuïtief dan empirisch van aard. Evaluatie van deze plannen, die vaak al in uitvoering waren, op hun effectiviteit ben ik niet tegengekomen. Een belangrijk probleem dat gesignaleerd werd in het

kader van de nascholing van docenten is de training van de opleiders. Daarvoor werd een tijd van twee jaar gerekend, terwijl er eigenlijk gisteren al met leerkrachtraining gestart had moeten worden. Deze training van leerkrachten zou zich dan met name moeten richten op de vraag hoe een leerkracht software in het curriculum kan integreren. Ten slotte was er een paper van Strudler (University of Oregon) dat zich richtte op de rol van de computer coördinator op schoolniveau. Een computer coördinator is een belangrijke 'change agent' in de school. Aanbevolen werd een minimale taakomvang van twee en een halve dag per week. Van de noodzakelijk geachte vaardigheden scoorden de sociaal-agogische en de technische het hoogst.

### *Conclusies*

Afrondend wil ik een drietal conclusies trekken. Op de eerste plaats is er nog steeds een tamelijk grote aandacht 'sec' voor de inhoud van softwareprogramma's, of deze nu onderwijskundig of administratief van aard zijn. Voorts is er een verschuiving in de aandacht aan het optreden naar de invoeringsaspecten, met name op welke wijze de invoering vorm moet worden gegeven, zodat er een breed draagvlak in de school ontstaat en/of een werkelijke integratie in de lespraktijk optreedt. De gegevens waarop de voorstellen voor invoeringsstrategieën zijn gebaseerd zijn overigens sterk intuïtief van aard. Ten slotte, onderzoek naar de effecten van computergebruik op de schoolorganisatie is nog niet of nauwelijks aanwezig.

De inhoud van de papers en de verschuiving erin laten eigenlijk een vrij logische ontwikkeling zien: eerst was de ontwikkeling van software het belangrijkste, daarin lijkt nu een redelijk niveau bereikt. Nu verschuift de aandacht naar de invoeringsproblematiek. Waarschijnlijk zullen de komende jaren de effecten op de schoolorganisatie aan aandacht winnen. In Nederland zien we hier al aanzetten toe, bijvoorbeeld de grootschalige evaluatie van het effect van het gebruik van het computerondersteunde Absentie Registratie Systeem op het absentisme.

### *Bedrijfsopleidingsproblematiek* (J. N. Streumer, Universiteit Twente)

Ook dit jaar werd door de Special Interest Group Training in Business and Industry weer een aantal sessies georganiseerd. Aangezien deze SIG betrekkelijk klein is (en informeel is georganiseerd, zo bleek mij o.a. tijdens de business meeting), kan het aanbod qua omvang niet wedijveren met dat van de AERA-divisies. Aan vier thematieken zal ik aandacht besteden: 1. karakteristieken van bedrijfsopleidingen; 2. bedrijfsopleidingskunde; 3. training, instructie en cursusontwikkeling; 4. evaluatie van trainingen.

### *Karakteristieken van bedrijfsopleidingen*

In de sessie gewijd aan de 'link between Education and Corporate Training', werd door Jones en Wall gerapporteerd over een onderzoek waarin werd nagegaan in hoeverre het Human Resources Development (HRD)-programma van Towson State University in Baltimore aansluit bij de behoeften van opleidingsafdelingen van de grootste werkgevers in de Baltimore Metropolitan Area.

Het onderzoek had het karakter van een survey en moest informatie opleveren met betrekking tot o.a. de volgende vragen: Welke trainingsvaardigheden worden het belangrijkste geacht? Hoe worden opleidingen over het algemeen opgezet en uitgevoerd? Waarvoor worden de materialen, die door de opleidingsafdeling worden geproduceerd, gebruikt?

Er bleek een grote mate van overeenstemming te bestaan tussen profit en non-profit organisaties. Er bestaat grote behoefte aan 'generic people', die onder druk kunnen werken, teamwork kunnen leveren en minimale begeleiding nodig hebben. Probleemoplosvaardigheden, communicatievaardigheden, kennis van het werktrein van de organisatie en het kunnen schrijven van rapporten scoren zeer hoog. Van groot belang worden tevens geacht het ontwerpen en uitvoeren van trainingsprogramma's en kennis van en vaardigheid in budgetteren.

Het ontwikkelen van trainingsprogramma's blijkt over het algemeen een individuele bezigheid te zijn. De meeste trainingsprogramma's zijn van het type 'on the job training'. De 'competency based approach' en 'self-instructional materials' zijn eveneens voorkomende typen.

De door de opleidingsafdelingen geproduceerde materialen worden toegepast in trainings- en communicatiesituaties. In trainingssituaties werden de materialen gebruikt voor het trainen van basisvaardigheden, managementontwikkeling en jobtraining.

De gegevens uit dit onderzoek worden gebruikt bij de revisie van het HRD-opleidingsprogramma. De in dit onderzoek gevolgde werkwijze om het opleidingsaanbod van een universitaire instelling af te stemmen op de wensen en behoeften van de potentiële werkgevers van studenten kan bijzonder inspirerend werken bij het opzetten en uitvoeren van soortgelijk arbeidsmarktonderzoek door faculteiten van Nederlandse universiteiten.

#### *Bedrijfsopleidingskunde*

In een onderzoek van Jacobs (Ohio State University) werd geconstateerd dat het inventariseren en beschrijven van 'training and development'-taken en functies op zich belangrijk is, maar dat hierbij niet te lang kan worden stil gestaan. Een groeiend aantal T&D-professionals, zowel uit de academische wereld als uit de praktijk, stelt zich de vraag of er een theoretisch T&D-netwerk bestaat en hoe zo'n netwerk de opleidingspraktijk kan ondersteunen. Om deze vraag te kunnen beantwoorden werd een zich ontwikkelend gebied van wetenschappelijke studie en praktijk, aangeduid als Human Performance Technology (HPT), in kaart gebracht. HPT tracht bestaande kennis uit verschillende disciplines, zoals de systeemtheorie, communicatiewetenschap, leerpsychologie, management en economie, te integreren en toe te passen op onderwerpen, die betrekking hebben op het functioneren van mensen in arbeidssituaties, waardoor zich langzamerhand een kennisgebied en daarmee geassocieerde procedures en processen gaan aftekenen. Jacobs pleit ervoor om geen scheiding aan te brengen tussen wetenschap, toegepaste wetenschap en technische kennis en vaardigheden uit de praktijk, aangezien dit het verkrijgen van inzicht in de samenstellende elementen van en het functioneren van effectief professioneel handelen in de opleidingspraktijk belemmert. Hij is, in navolging van Argyris, een voorstander van de 'theories of action'-benadering. Het HPT-domein bestaat uit drie belangrijke componenten: 1) functies gericht op het managen en controleren van de ontwikkeling van human

performance systems; 2) functies om human performance systems te ontwikkelen; 3) componenten van human performance systems, zoals organisatieklimaat, motivaties, doelen, (re)acties, beslissingen, consequenties, feedback.

In de discussie over de functie die onderzoek zou kunnen vervullen voor T&D (of HDR, of HPT) kan het paper van Jacobs een belangrijke rol vervullen: hij doet een waardevolle poging het terrein van dit onderzoek af te bakenen en somt een aantal onderwerpen op, waar het onderzoek zich op zou moeten richten.

#### *Training, instructie en cursusontwikkeling*

Duffy (University of New Orleans) deed verslag van een pilot study, die betrekking had op de evaluatie van een micro-computertraining, verzorgd door de universiteit van New Orleans. Gegevens werden verzameld met een zg. happiness-scale, een prestatietoets en evaluatiegegevens die door de cursisten zelf werden gerapporteerd. De presentatie had jammer genoeg grotendeels betrekking op de gegevens die werden verzameld met de happiness-scale, een instrument waartegen terecht meer en meer weerstand ontstaat, aangezien het uiterst 'softe' gegevens ('overall, I feel that I have learned a great deal in this course') oplevert.

Stone (University of Minnesota) ging in op de relatie tussen leerstijlen en de door cursisten gepercipieerde bruikbaarheid van instructiemethoden (college, rollenspel, groepsdiscussie in kleine en grote groep, individuele opdracht, demonstratie) en de mate waarin ze betrokken waren in die methoden. Daarnaast ging hij in op de vraag of de betrouwbaarheid van het door hen gebruikte Learning Style Instrument voldoende is om prescriptieve uitspraken te doen over de leerstijlen van individuele cursisten.

Ongeveer een derde van de cursisten bleek te beschikken over een 'sterk leerstijl' (strong accomodator, assimilator, diverger, converger), hetgeen implicaties kan hebben voor het opzetten van trainingen. Deze cursisten kunnen gedemotiveerd, resistent en/of gefrustreerd raken van trainingsprogramma's die wat betreft hun instructiemethoden eenzijdig georiënteerd zijn. Variëteit en flexibiliteit moeten trefwoorden zijn bij het ontwerpen van opleidingen; sensitiviteit voor de behoeften van cursisten moet een kwaliteit zijn die opleiders in hoge mate moeten bezitten.

McLinden et al. (Arthur Anderson & Co) presenterden de resultaten van een vergelijking van twee instructiehandleidingen: een gedetailleerde ('what to say and what to do') en een uitgekilde versie (zonder uitvoerige aanwijzingen voor de cursusleider). De volgende onderzoeksvragen werden beantwoord: 1. Neemt de voorbereidingstijd van de cursusleider toe bij het gebruik van de uitgekilde versie? 2. Neemt de getrouwheid in uitvoering toe als de uitgebreide versie wordt gebruikt? 3. Heeft het gebruik van de uitgekilde versie negatieve invloed op de wijze waarop de cursusleider functioneert? 4. Wat zijn de financiële implicaties?

In totaal namen 42 cursusleiders en 785 cursisten deel aan het onderzoek. Er bleek geen significant verschil te bestaan wat betreft voorbereidingstijd tussen de gebruikers van beide versies.

De veranderingen die door cursusleiders werden aangebracht in beide versies van het cursusmateriaal verschilden ook niet significant: de meeste cursusleiders (beide condities) veranderden iets aan de inhoud (72%), voegden voorbeelden toe (82%) en brachten wijzigingen aan in de tijdsverdeling van de onderwerpen (58%).

De analyse van de antwoorden van de cursisten bracht eveneens weinig verschillen aan het licht wat betreft het functioneren van de cursusleiders.

Met de uitgekilde versie werd een besparing van 59 procent bereikt in de omvang van de cursushandleiding, hetgeen een flink financieel voordeel oplevert bij de produktie van honderden cursushandleidingen per jaar (en waarschijnlijk een besparing in ontwikkel-tijd).

Deze gegevens zijn interessant in het licht van de ontwikkeling van een rationale voor het ontwerpen en produceren van cursusmateriaal.

#### *Evaluatie van trainingen*

Tot slot enige aandacht voor de evaluatie van trainingen. Interessant was een lezing van Niemiec, als opleidingsfunctionaris verbonden aan Digital Equipment Corporation. Niemiec constateert dat de termen 'cost-effectiveness' en 'cost-benefit' veelvuldig door elkaar worden gebruikt, ondanks het feit dat ze een verschillende betekenis hebben. Bij kosten-baten-analyse is zowel een specificatie van de

projectkosten als de potentiële opbrengst in geld een noodzaak. Het voordeel van kosten-effectiviteit is dat de opbrengst niet in geld hoeft te worden uitgedrukt. Uitsluitend de kosten verbonden aan de implementatie van de interventie en het relatieve effect behoeven te worden geschat.

De schatting van de interventiekosten en haar invloed kan worden gecombineerd tot een kosten effectiviteitsratio (CER). De ratio's van soortgelijke programma's kunnen vervolgens worden gerangschikt om de relatieve kosten-effectiviteit van verschillende trainingsinterventies te bepalen. Een probleem dat zich hierbij voordoet is dat verschillende instrumenten worden gebruikt om de effectiviteit van trainingsinterventies te bepalen. Een methode om dit probleem te omzeilen, is het omzetten van de resultaten van interventies in 'effect sizes'.

Deze methode wordt aangeduid als meta-analyse (afkomstig uit de experimentele psychologie). Als de kosten en effecten van trainingsprogramma's, gemeten in 'effect sizes' zijn geschat, dan kunnen ze worden gecombineerd in kosten-effectiviteitsratio's (CER) met de volgende formule:  $CER = ES/PSC \times 100$  waarbij ES de "effect size" van het trainingsprogramma is en PSC de kosten per cursist. Deze methode maakt het opleidingsmanagers mogelijk om verantwoord tussen verschillende opleidingsinterventies te kiezen.

#### *Terugblik*

Geconstateerd moet worden dat, datgene wat in onderzoek geïnteresseerde opleidingsfunctionarissen en (onderzoeks)medewerkers van HRD-afdelingen van Amerikaanse universiteiten in New Orleans presenterden, over het algemeen zeer informatief was en van een behoorlijk niveau (ook wat betreft de research-aanpak). Het loont voor Nederlandse onderzoekers, op het gebied van de bedrijfsopleidingsproblematiek, de bijeenkomsten van de SIG 'Training in Business and Industry' bij te wonen (of zich bij deze SIG aan te sluiten).

*Kenmerken van effectieve scholen*  
(W. G. R. Stoel, Universiteit Twente)

Sinds Coleman en Jencks op grond van hun onderzoek concludeerden, dat verschillen tussen scholen niet of nauwelijks samenhangen

met verschillen in leerprestaties van leerlingen afkomstig van verschillende scholen, hebben de nodige ontwikkelingen plaatsgevonden in de inzichten omtrent het bestaan en de omvang van zgn. schooleffecten.

Na de onderzoeken van bijvoorbeeld Edmonds en Rutter et al. ontstond er een euforie met betrekking tot de mogelijkheden die scholen zouden hebben om de leerprestaties van de leerlingen te beïnvloeden. Deze euforie duurt in bepaalde kringen tot op de dag van vandaag. Toch is daarvoor eigenlijk weinig reden. Door verscheidene auteurs is zware en fundamentele kritiek geuit op het uitgevoerde onderzoek. Hierbij kan verwezen worden naar publikaties van bijvoorbeeld Cuttance, Burnstein en d' Amico en in eigen land naar bijvoorbeeld Kreft, Stoel en Scheerens. De door deze auteurs gesignaleerde zwakke elementen in het onderzoek naar schooleffecten c.q. schooleffectiviteit kan men tot op de dag van vandaag aantreffen in het uitgevoerde onderzoek op dit terrein. Dit weerhoudt vele onderwijzers in met name de Verenigde Staten er echter niet van om, op grond van deze magere en omstreden basis van effectiviteitsstudies, veel geld, energie en tijd te investeren in de implementatie van de resultaten van diezelfde effectiviteitsstudies. Er werden nogal wat van dit type studies gepresenteerd. Het gemeenschappelijke in al deze presentaties was dat men tot de conclusie kwam dat al hetgeen men trachtte te implementeren ook inderdaad het gewenste effect had. Op mij maakte de wijze van rapportage en argumentatie meer de indruk dat er sprake was van wishful thinking dan van een (kritische) presentatie van onderzoeksresultaten. Deze laatste opmerking geldt overigens voor veel van de AERA-presentaties, ook daar waar het andere terreinen van onderwijsonderzoek betreft.

In iedere sessie die ik bijwoonde was er echter altijd wel een presentator die een kritische noot liet horen. Wanneer ik deze kritische noten op een rijtje probeer te zetten, wordt een, uiteraard persoonlijk gekleurde, schets verkegen van wat op dit moment de aandachtspunten zouden moeten zijn bij de opzet en uitvoering van studies naar kenmerken van effectieve scholen.

#### *Stabiliteit van effectiviteit*

Verscheidene presentatoren maken melding van het fenomeen dat scholen die het ene jaar

als zeer effectief uit de bus komen, met andere woorden de hoogste scores behaalden op landelijk genormeerde schoolvorderingstoetsen, het daarop volgende jaar lang niet altijd opnieuw deze hoge scores konden realiseren. De vraag doet zich derhalve voor of verschillen in schoolprestaties van leerlingen op verschillende scholen niet simpelweg het gevolg zijn van toevalsfluctuaties in plaats van kenmerken van die specifieke school of groep scholen. Onduidelijk is daarbij of in alle gevallen wel is gecontroleerd voor fluctuaties in de inputkenmerken van de leerlingpopulaties van de scholen. Wanneer dit niet het geval zou zijn, kan dat mede een verklaring vormen voor genoemd verschijnsel.

Een aantal jaren geleden zijn verscheidene zgn. 'outlierstudies' verricht, waarbij effectieve en minder effectieve scholen werden vergeleken ten aanzien van de manier waarop het onderwijs werd vormgegeven, de school werd georganiseerd e.d. Vaak zijn toen vergaande conclusies getrokken met betrekking tot hetgeen men als kenmerken van effectieve scholen zou kunnen karakteriseren. Vanuit methodologisch oogpunt kan sowieso kritiek worden uitgeoefend op dit type onderzoek. Wanneer blijkt dat de mate van effectiviteit van een school van jaar tot jaar fluctueert, dan moet men zich afvragen wat de waarde is van de uitkomsten van dergelijke studies. Toch hebben de resultaten van deze studies nog steeds invloed op de effectiviteitsstudies in die zin dat de selectie van schoolkenmerken die worden opgenomen in onderzoek, veelal nog steeds worden gebaseerd op de uitkomsten van deze outlier-studies. Op die manier komt men in een situatie terecht waarbij de resultaten van de huidige effectiviteitsstudies wel moeten tegenvallen, omdat men als gevolg van het ontbreken van een goed theoretisch kader, de verkeerde schoolkenmerken selecteert voor opname in onderzoek.

#### *Andere effectmaten*

Verscheidene presentatoren vroegen aandacht voor het gebruik van andere effectmaten dan gestandaardiseerde schoolvorderingstoetsen in onderzoek naar schooleffecten. De overweging hierbij was dat gestandaardiseerde schoolvorderingstoetsen te eenzijdig een beeld geven van de resultaten van de schoolloopbaan van leerlingen. Het zou te ver voeren om hier de discussie rond de keuze van effecti-

viteitscriteria te voeren, derhalve wil er erme volstaan er melding van te maken dat ook de keuze van het effectiviteitscriterium in belangrijke mate medebepalend lijkt te zijn voor de resultaten van effectiviteitsstudies.

Als alternatieven voor de traditionele effectiviteitscriteria wordt daarbij niet alleen gedacht aan niet-cognitieve effectmaten, maar ook aan weinig gebruikte maar wel schoolprestatie-indicerende effectmaten als het percentage geslaagden, zittenblijvers, drop-outs, e.d., kortom variabelen die de schoolloopbaan van leerlingen op de scholen karakteriseren.

### *Contextinvloeden*

Veel aandacht werd gevraagd voor de invloed van contextkenmerken op de mate waarin een school effectief is. Daarbij kan men denken aan cultuur- en onderwijsstelselverschillen tussen landen, die ervoor zorgen dat wat effectief is in het ene land, dit niet per definitie ook is in een ander land. Hetzelfde geldt voor verschillen tussen scholen met verschillende leerlingpopulaties. Maatregelen die effectief blijken op scholen met overwegend leerlingen uit lagere sociale milieus behoeven dat niet te zijn voor scholen met leerlingen uit overwegend hogere sociale milieus en vice versa. Ook zijn er indicaties dat bij leerlingen uit verstedelijkte gebieden een andere aanpak gewenst is dan bij leerlingen uit plattelandsgebieden. Ook ligt het voor de hand te veronderstellen dat kenmerken van effectieve basisscholen niet dezelfde zullen zijn als kenmerken van effectieve scholen voor voortgezet onderwijs. Voor deze stellingnames werd ook empirisch materiaal aangedragen.

Generaliserende uitspraken over kenmerken van effectieve scholen in het algemeen lijken dan ook uit den boze. Veel onderwijsonderzoek en verschillende innovatieprojecten op dit gebied hebben kennelijk geen boodschap aan deze voor de hand liggende conclusie, gezien de ijver en het fanatisme waarmee telkens opnieuw 'de' vijf kenmerken van Edmonds worden onderzocht en geïmplementeerd. De empirische basis hiervoor, zo is mijn indruk op grond van de bijgewoonde presentatie, is uiterst zwak.

### *Theorievorming en de causaliteit van de aange- toonde relaties*

Veel onderzoek verricht naar effectieve

schoolkenmerken is empirisch van aard. Men relateert een aantal schoolkenmerken aan effectvariabelen en constateert vervolgens een aantal (cor)relaties. Het hoe en waarom van die relaties is veelal onduidelijk. Zowel een positieve als een negatieve relatie, alsmede het totaal afwezig zijn van een relatie kan door een beetje creatieve wetenschapper worden 'verklaard'. Onderzoek op basis van een theoretisch verklaringsmodel is uiterst schaars.

### *Het multilevel karakter van dit type onderzoek*

Tot mijn verbazing werd in vrijwel geen enkele presentatie die ik bijwoonde expliciet aandacht gevraagd voor het feit dat onderzoek naar schooleffecten altijd gegevens betreft die op meerdere niveaus zijn verzameld. Zowel Burnstein als Cuttance hebben daar in het verleden al aandacht voor gevraagd. Er is inmiddels een aantal statistische pakketten beschikbaar om gegevensbestanden met een dergelijke structuur (beter) te analyseren. In Nederland hebben Van Eeden, Krefte en Bosker reeds het nodige gepubliceerd op dit terrein en onlangs nog hebben Brandsma en Knuber verslag gedaan van een onderzoek waarbij gebruik werd gemaakt van deze geavanceerde statistische analyses. Derhalve doet het wat vreemd aan dat in de Verenigde Staten kennelijk weinig aandacht en interesse blijkt te bestaan voor het methodologisch- en statistisch verantwoord verzamelen c.q. analyseren van onderzoeksgegevens op dit terrein. Wij zijn geneigd vaak de blik naar het westen te richten, daar waar men geïnteresseerd is in nieuwe ontwikkelingen op het gebied van de onderwijsresearch. Ik moet tot de conclusie komen dat in Nederland de onderwijsresearch, mondiaal gezien, op een zeer aanvaardbaar peil staat, zeker wanneer men dat vergelijkt met hetgeen men er in het 'mekka' van de onderwijsresearch, de Verenigde Staten, van terecht brengt.

Tot slot nog enige opmerkingen naar aanleiding van de roep om implementatie van maatregelen waarvan onderzoek zou hebben uitgezonden dat deze het rendement van het onderwijs zouden verhogen. Hiervoor is reeds gezegd dat inzicht in de causaliteit van de aangetoonde relaties ontbreekt, dat een theoretisch kader waarmee de relaties verklaard c.q. voorspeld kunnen worden eveneens zeer gebrekkig aanwezig is en dat de empirische

basis onder veel onderzoek ronduit zwak is. Het lijkt derhalve hoogst onverstandig om reeds in dit stadium van kennisonwikkeling rond schooleffecten een begin te maken met de implementatie van de spaarzame bevindingen, ook al is de roep om rendementsverbetering en kwaliteitsverhoging nog zo groot. In de Verenigde Staten worden vele effectiviteitsprojecten uitgevoerd; op de AERA worden trots de 'resultaten' bekend gemaakt. Bij de presentaties van dergelijke projecten vertelt men er veelal niet bij dat er, afhankelijk van het project, honderden miljoenen dollars in worden gestoken. Hoewel extra geld in het onderwijs stoppen niet automatisch tot kwaliteitsverbetering leidt, geldt eveneens dat het heel goed mogelijk is dat een eventueel succes van dergelijke projecten vooral te danken is aan extra leermiddelen, mankracht e.d. en veel minder aan hetgeen men aan maatregelen op schoolniveau heeft geïmplementeerd. Een andere set aan maatregelen had met dezelfde extra inzet in middelen en mensen wellicht tot hetzelfde (of een beter) resultaat geleid. Met wetenschap of onderzoek hebben dit soort presentaties weinig te maken.

#### *Conclusie*

Op grond van, onder meer het bijwonen van dit AERA-congres, kom ik tot de conclusie dat er nog heel wat water door de Mississippi moet stromen, c.q. dat er nog zeer veel grondige evaluatiestudies zouden moeten worden verricht, voordat men verantwoord kan stellen dat de implementatie van bepaalde maatregelen op het niveau van de school, met een acceptabele mate van zekerheid, zal leiden tot een verbetering van de leerprestaties van de leerlingen.

#### *Peilingsonderzoek (Th. Eggen en P. Sanders, CITO, Arnhem)*

De resultaten en de implicaties van grootschalig periodiek peilingsonderzoek kregen veel aandacht. Dit geldt met name voor het National Assessment of Educational Progress (NAEP) programma, maar ook voor peilingen in verschillende Amerikaanse staten. Opvallend veel aandacht was er ook voor de methodologie van peilingsonderzoek. Verantwoordelijk hiervoor is de grote reorganisatie in de opzet van NAEP sinds de uitvoering in

1983 in handen is gekomen van ETS. De belangrijkste veranderingen in de methodologie bestaan uit het toepassen van itemrespons-theorie gecombineerd met geavanceerde vormen van matrix proefopzetten. Omdat bij peilingsonderzoek niet de individuele meting van vaardigheden vooropstaat, maar metingen van (veranderingen in) vaardigheden van subgroepen in de populatie, werd reeds langer gebruik gemaakt van mogelijkheden die multiple matrix sampling technieken bieden voor efficiënte afnamedesigns. Inmiddels lijkt ook algemene overeenstemming te bestaan over het feit dat itemrespons-theorie het beste gereedschap biedt voor het meten van veranderingen in de tijd. Dit laatste komt met name doordat bij modelgeldigheid met itemrespons-theorie vaardigheden onafhankelijk van de specifiek gebruikte items kunnen worden geschat.

Echter de combinatie van de efficiëntere designs met het ander meetperspectief, waarvoor de itemrespons-theorie in eerste instantie niet ontwikkeld is, heeft om nogal wat uitbreidingen van de bestaande methodologie gevraagd. In de papers van Mislevy en Johnson werd op een aantal oplossingen voor hiermee samenhangende problemen ingegaan, zoals die in NAEP worden toegepast. Een van de in NAEP ontwikkelde toepassingen verdient speciale vermelding. Doordat in de designs individuen op specifieke leerstofgebieden slechts een gering aantal opgaven maken is het niet mogelijk goede individuele schattingen van vaardigheden te krijgen. Voor het hoofddoel van peilingen is dit geen probleem, maar voor het uitvoeren van nadere analyses naar relaties tussen kenmerken van leerlingen en vaardigheden zijn individuele schattingen van deze vaardigheden zeer gewenst. Door het genereren van zogenaamde 'plausible values' wordt dit binnen NAEP mogelijk. Daarbij wordt op originele en succesvolle wijze gebruik gemaakt van Rubins multiple imputatie technieken. Door de NCME werd de bijdrage van NAEP met name hiervoor met een prijs gehonoreerd.

Twee andere ontwikkelingen op het gebied van het peilingsonderzoek zijn nog vermeldenswaard. Allereerst was in de NAEP-bijdragen duidelijk dat wordt overgegaan op het ontwikkelen van schalen bij inhoudelijk niet te ruim gedefinieerde onderwerpen. Alhoewel nog niet zover wordt gegaan om de

zeer fijnmazige opdeling zoals die al een aantal jaren in de Californian Assessment Program (CAP) wordt toegepast te benutten worden de voordelen ervan ingezien. Door de beperking van de reikwijdte van de schalen kan beter vat worden gekregen op het mogelijk differentieel veranderen van itemkenmerken in de tijd ('parameter drift') en nemen de interpretatiemogelijkheden van de resultaten voor inhoudsdeskundigen toe.

Verder werd in bijdragen van Bock en Zimowski melding gemaakt van de toepassing en evaluatie van het zogenaamde duplex design voor peilingsonderzoek in California, nadat het eerder was uitgetoetst in Illinois. Een van de nieuwe elementen in dit design is dat een twee stappen toetsprocedure wordt toegepast. Op grond van een voortoets worden leerlingen ingedeeld in vaardigheidsgroepen en wordt op basis daarvan beslist wat de moeilijkheid is van de vervolgtolsten die de leerlingen maken. De efficiëntie van dit design bleek groot. Tenslotte blijft het opmerkelijk dat in de Verenigde Staten in de ontwikkeling van de methodologie van peilingsonderzoek, maar daar niet alleen, de produktie van werkbare procedures over het algemeen een hogere prioriteit heeft dan de theoretische fundering van de procedures. Zo wordt bijvoorbeeld in NAEP het drie parameter logistisch testmodel, ondanks de onmogelijkheid de geldigheid van dit model op verantwoorde wijze te toetsen, om duidelijk pragmatische redenen gebruikt.

#### *Adaptief toetsen*

Een in Nederland nooit echt populair geworden onderwerp dat in de Verenigde Staten zeer veel aandacht heeft gehad en nog heeft is adaptief toetsen. Met adaptief toetsen wordt bedoeld: een toetsprocedure die een vaardigheid van een persoon meet middels selectie van items die afgestemd zijn op het vaardigheidsniveau van de persoon tijdens de toetsafname. Dit toetsen op maat heeft in de Verenigde Staten een flink aantal operationele toepassingen. Het combineren van de toegenomen mogelijkheden van de microcomputer en het gebruiken van itemresponstheorie hebben dit mogelijk gemaakt. Reckase presenteerde een overzicht van de stand van zaken. Hij onderscheidt de volgende vier elementen in een goed adaptief toetsstelsel: een itemverzameling, een item selectieprocedure, een scoringsprocedure en

een stopregel. Voor alle vier elementen zijn technieken ontwikkeld en reeds geïmplementeerd. In de toekomst zal met name nog aandacht worden besteed aan het verfijnen van de bestaande technieken. Het scoren van de personen, in een itemresponscontext het schatten van vaardigheden, heeft nog steeds te kampen met onopgeloste statistische problemen. Reckase verwacht vooral dat de uitbreiding van de modellering van de itempersooninteractie nog veel onderzoek zal vragen. Hij noemt hierbij de ontwikkeling van niet-monotome, polychotome en meerdimensionale itemresponsmodellen.

#### *Itemresponstheorie*

Onder deze titel worden al jaren lang veel papers gepresenteerd. Ook dit jaar was het niet anders. Opvallend waren de bijdragen die zich met name richten op de uitbreidingen van bestaande itemresponsmodellen naar modellen die rekening houden met leren en verschillen in instructie en naar meerdimensionale modellen en modellen voor polychotome items. Ook opmerkelijk was dat er in dit verband, in tegenstelling tot de grote aandacht die hiervoor in Nederland bestaat, geen papers waren die de itemresponstheorie direct relateerden aan de ontwikkeling van itembanken. Vermeldenswaard is tenslotte een sessie met voornamelijk Nederlandse bijdragen (Verhelst, Van der Linden, De Gruyter en Glas) waarin aandacht werd besteed aan het schatten van itemparameters in onvolledige en dus in principe meer efficiënte designs.

#### *Equivalenten*

Evenals in voorgaande jaren werd ook dit jaar weer een aantal papers over methoden voor het equivalenten van toetsen gepresenteerd. Aangezien het een probleem betreft waarmee vooral instituten die toetsen ontwikkelen te kampen hebben, was het merendeel van de paperlezers dan-ook afkomstig van Educational Testing Service (ETS) uit Princeton. De papers betroffen voornamelijk vergelijkingen tussen equivalentdesigntypes (wel of geen overlap tussen proefpersonen en/of items) en/of equivalentfuncties (lineair, niet-lineair en equipercentieel) in het kader van zowel de klassieke testtheorie als de itemresponstheorie. Ook deze papers bevestigden opnieuw dat er (nog) niet één equivalentmethode aangemerkt kan worden als de methode waarmee de meeste in



de praktijk voorkomende problemen en situaties opgelost kunnen worden.

#### *Item bias*

Opvallend veel papers hadden betrekking op procedures voor het detecteren van items met bias. In plaats van 'item bias' wordt tegenwoordig ook wel de term 'differential item functioning' (DIF) gebruikt. Een item wordt biased genoemd als personen met dezelfde vaardigheid, maar afkomstig uit verschillende groepen (mannen/vrouwen, blanken/zwarten), niet dezelfde kans hebben het item goed op te lossen. Van de vele procedures die zijn voorgesteld zijn vooral die gebaseerd op de itemresponstheorie en de Mantel-Haenszel (MH) methode populair. Hoewel de MH-procedure het voordeel van de eenvoud heeft, is deze minder gevoelig voor het detecteren 'non-uniform' biased items, d.w.z. een item waarbij het verschil in de kans op het goed oplossen van het item door twee verschillende groepen niet gelijk is op alle vaardigheidsniveaus. Als alternatief voor de MH-procedure werd door Swaminathan en Rogers een logistische regressie procedure voorgesteld, waarmee het mogelijk bleek te zijn zowel uniform biased als non-uniform biased items te kunnen detecteren.

Kritische kanttekeningen bij de tot nu toe ontwikkelde indices voor item bias werden door Skaggs geplaatst. Uit zijn onderzoek, maar ook uit vele verslagen van empirisch onderzoek, bleek dat de overeenstemming tussen de verschillende indices nogal gering was.

#### *Polychotoom scoren van items*

Meerkeuze items worden meestal dichotoom gescoord, dat wil zeggen dat items die goed beantwoord worden het gewicht één krijgen en items die fout beantwoord worden het gewicht nul krijgen. De som van de enen en nullen uit de itemresponsvector van een persoon is zijn toetscore. Aangezien er de laatste jaren steeds meer belangstelling bestaat voor het poly(cho)toom scoren van items, was er een papersessie gewijd aan dit onderwerp.

In hun overzicht van de literatuur onderscheiden Haladyna en Sympson twee benaderingen. De eerste betreft het toekennen van verschillende gewichten aan de itemresponscategorieën of alternatieven. De toetscore is een lineaire functie van de itemresponsvector van een persoon. Een veelvuldig toegepaste

methode uit deze benadering is die waarbij coëfficiënt alpha gemaximaliseerd wordt. Voordelen van deze methode zijn dat de toets niet unidimensioneel hoeft te zijn en dat kleine steekproeven, kleiner dan duizend personen, voldoende zijn. Nadelen zijn dat de gewichten steekproefafhankelijk zijn en dat het antwoord dat fout is soms een hoger gewicht krijgt dan het antwoord dat goed is. Uit vergelijkingen die gemaakt zijn tussen toetsen waarvan de items zowel dichotoom als polychotoom gescoord worden, blijkt dat voor het bereiken van een gelijke betrouwbaarheid de dichotoom gescoorde toets (veel) langer moet zijn dan de polychotoom gescoorde toets.

De tweede benadering voor het polychotoom scoren van items is afkomstig uit de itemresponstheorie. Deze benadering heeft als voordeel dat als aan de aanname van het polychotome itemresponsmodel voldaan is (unidimensionaliteit, items passen bij het model) de parameters van een item onafhankelijk zijn van de parameters van andere items uit de toets. De resultaten van zowel simulatiestudies als empirische studies laten zien dat – vergeleken met andere modellen – de toepassing van polychotome IRT modellen een aantal voordelen heeft. Zo worden vaardigheden van personen nauwkeuriger geschat, met name van personen waarvoor de items niet te gemakkelijk zijn, kan de toetslengte gereduceerd worden en kunnen personen met 'afwijkende' itemresponsvectoren beter gedetecteerd worden.

Aan deze kroniek werkten mee: *G. J. J. Biesta, Th. Eggen, L. F. W. de Klerk, J. Lowyck, W. J. Nijhof, J. L. M. Remmers, P. Sanders, P. R. J. Simons, W. G. R. Stoel, J. N. Streumer.*

De eindredactie werd verzorgd door *W. J. Nijhof*