

Kroniek

The Annual Meeting of the American Educational Research Association, April, 20-24, 2000, New Orleans.

Creating Knowledge in the 21st century

Inleiding (H. Dekkers, Katholieke Universiteit
Nijmegen)

Het was te verwachten dat de ligging en de aard van de stad waar de 81ste editie van het jaarlijkse congres van de AERA werd gehouden borg stond voor een groot aantal bezoekers. Met een recordaantal van 13.500 bezoekers kwam die verwachting ruim uit. Vanuit de twaalf divisies en ongeveer honderd grotere of kleinere Special Interest Groups (SIGs) werden ongeveer duizend symposia met ruim drieduizend presentaties aangeboden.

De Nederlandse onderwijsresearch was goed vertegenwoordigd. Een aantal Nederlandse bezoekers van het congres doet zoals gebruikelijk in deze rubriek in eerste instantie verslag van de divisies die aansluiten bij hun eigen onderzoeksthematieken. Dat laat onverlet dat interessante bijdragen vaak gevonden werden onder de vlag van andere divisies.

Divisie A: Administration (Meta L. Krüger, Universiteit van Amsterdam en Bob Witziers, Universiteit Twente)

Divisie A was dit jaar verdeeld in vijf secties waarbinnen de papers waren gegroepeerd: leiderschap, schoolverbetering, de school en haar omgeving, leiderschapstraining en -ontwikkeling, en schoolorganisatie. Er waren 83 symposia en papersessies in de aanbieding binnen deze divisie, elk met meestal vier en soms drie papers. Zoals elk jaar dus weer een ruim aanbod, waarin vooral vertrouwde onderzoeksthema's aan bod kwamen, zoals onderwijsstandaarden en accountability, de school als

lerende organisatie, docenten als leiders, samenwerking tussen schoolomgeving, ouders en school, en onderzoek naar onderwijshervormingen. Daarnaast kwamen ook recentere thema's aan bod, die we kunnen samenvatten in twee nieuwe trends, namelijk de hernieuwde belangstelling voor de rol van de schoolleider en de aandacht voor het ontstane tekort aan gekwalificeerde kandidaten voor het schoolleiderschap.

Het benadrukken van het belang van de schoolleidersrol blijkt te zijn teruggekeerd in het onderzoek. In de jaren 80 en 90 was er vanuit de Verenigde Staten een geleidelijke opkomst van de gedachte dat de schoolleider niet zo belangrijk was voor het functioneren van de school. Enerzijds speelden daarbij de teleurstellende resultaten uit het onderzoek naar de relatie tussen onderwijskundig leiderschap en leerprestaties een rol, anderzijds werd in het kader van 'empowering teachers' en 'school restructuring' steeds meer aandacht besteed aan de invloed in de school van het leiderschap door docenten. De hernieuwde aandacht voor de rol van de schoolleider lijkt voort te vloeien uit het feit dat de maatschappelijke context waarbinnen scholen opereren essentieel is veranderd, gecombineerd met het sterker worden van de invloed van de externe omgeving op scholen. Zo kwam in verschillende sessies naar voren dat scholen steeds meer verweven raken met hun maatschappelijke omgeving (buurten, ouders, bedrijven). Daarnaast werd veelvuldig gewezen op de eisen die vanuit overheidswege op scholen afkomen. Schoolleiders zijn degenen die vaardigheden moeten ontwikkelen om met die vergrote invloed van de omgeving om te gaan, er een visie over te ontwikkelen en deze te vertalen in doelen voor de school. De schoolleider krijgt een sleutelrol te vervullen in het helder maken naar docenten toe van het belang van onderwijsinnovaties op het niveau van de klas en van de gewenste richting van deze onderwijsinnovaties. Al deze ontwikkelingen vereisen sterk leiderschap. Overigens beperkt de hernieuwde aandacht voor de rol van de schoolleider zich niet alleen tot de Verenigde Staten, zo bleek bijvoorbeeld tijdens een sessie waarin vergelijkend onderzoek werd gepresenteerd naar leiderschap uit Amerika, Engeland en Denemarken. Paul Bredeson, discussiant in deze sessie,

zag de trend van terugkeer tot 'one strong leader' als een gevaarlijke terugkeer naar eerdere ideeën.

Met deze herleefde aandacht voor de rol van de schoolleider is ook een hernieuwde zoektocht naar kenmerken van de effectieve schoolleider ontstaan. Die kenmerken worden nu echter niet meer gezocht in rijtjes do's and don'ts, zoals we zagen in het begin van het onderzoek in de jaren 80 vanuit het schooleffectiviteitsonderzoek, maar eerder in de visie en de cognities van de schoolleider. Ook wordt in verschillende presentaties ingegaan op de ethiek van het leiderschap en de invloed van de waarden van de schoolleider. Zo vat een discussiant de gepresenteerde papers in een bepaalde sessie samen met de uitspraak: 'successful principals are value led'. De visie, cognities en waarden van de schoolleider worden dus gezien als van belang voor het effectief kunnen aansturen van de schoolorganisatie. Dit idee wordt gekoppeld aan eerdere ideeën over teacher leadership. Teacher leadership is alleen goed mogelijk als er een schoolleider is die daar sturing aan geeft via het formuleren van een visie voor de school. Zoals Leithwood het verwoordde: 'de visie van de schoolleider bevindt zich niet op het niveau van onderwijs en klas. De gap die daardoor ontstaat moet worden opgevuld door teacher leadership'. Ook vindt een verbinding plaats met eerdere ideeën over de school als lerende organisatie. Een discussiant in een sessie over environmental leadership concludeerde: 'gedeeld leiderschap in lerende organisaties maakt van het leidinggeven een hoogst morele en menselijke activiteit, waarmee de cognitieve processen belangrijk worden waardoor het leidinggeven wordt gestuurd'. De waarden en de cognitieve processen van docenten zoals verankerd in de schoolcultuur, het samenwerken en samen leren van docenten ('the school as professional community') worden gezien als essentieel voor een effectieve schoolorganisatie. Er werden uitdrukkingen gebruikt zoals 'culture as cognitive process', 'culture as shared learning' en 'organizations as cognitive accomplishments, constructed by meanings that people hold to them'. De vraag die daarbij opkomt is natuurlijk hoe zo'n 'school als professionele gemeenschap' tot stand kan komen. Inmiddels is duidelijk geworden dat daarvoor de aanwezigheid van een schoolleider nood-

zakelijk is die daar sturing aan geeft; een schoolleider die op basis van waarden en visie de cognitieve processen van docenten weet te veranderen. Daarmee lijkt de kring rond.

De tweede trend die we willen bespreken is de crisis in het schoolleiderschap die in de Verenigde Staten recent wordt gesignaleerd. Een crisis in die zin dat de baan zo zwaar blijkt te zijn geworden, dat er steeds minder belangstelling voor bestaat en dat steeds meer schoolleiders en superintendents vertrekken uit hun functie. Daardoor is niet alleen een tekort aan kandidaten ontstaan voor de opvallende vacatures, maar is ook de kwaliteit van potentiële kandidaten lager geworden. Er waren twee sessies speciaal gewijd aan deze crisis, onder de naam: 'Superintendents and principals: who will fill their shoes?' en 'The challenges of educational leadership: is the job possible and does anyone want it?' Galvin vond in zijn onderzoek dat slechts 15% van de schooldirecteuren 10 of meer jaren volmaakte, terwijl 13% de functie niet langer dan één jaar vervulde. In een periode van twee jaar moet maar liefst 31% van alle schooldirecteuren worden vervangen. Naar aanleiding van de crisis is in de Verenigde Staten onderzoek op gang gekomen naar de mobiliteit van schoolleiders, maar nog niet naar hoeveel van die mobiliteit bestaat uit onvrijwillig en voortijdig vertrek van schoolleiders, alhoewel wordt vermoed dat dit toeneemt. Er werd wel Nederlands onderzoek naar voortijdig vertrek gepresenteerd (Krüger, Van Eck en Vermeulen), waaruit naar voren kwam dat meer dan een kwart van de Nederlandse schoolleiders die mobiliteit vertonen, voortijdig vertrekken. Vaak gaat dit voortijdige vertrek gepaard met een arbeidsconflict. Het voortijdig vertrek blijkt significant hoger onder vrouwelijke directeuren dan onder mannelijke. Opvallend is ook dat door het ontstaan van de schoolleiderscrisis in de Verenigde Staten de variabele sekse een interessantere variabele lijkt te worden in het onderzoek. Volgens Diana Pounder, discussiant in één van de sessies, staat 'gender' al jaren op de onderzoeksagenda, maar wordt het nu pas serieus genomen. Ook mannen betrekken de variabele in hun onderzoek, terwijl onderzoek naar sekseverschillen van zowel mannelijke als vrouwelijke onderzoekers nu ook vaker vergelijk-

kend van aard wordt. Veel van het eerdere onderzoek op dit terrein werd namelijk alleen onder vrouwen verricht, waardoor er nogal wat stereotypen en vooroordelen over vrouwelijk leiderschap naar boven zijn gekomen.

In menige sessie werd het feit gememoreerd dat veel van het onderwijsonderzoek naar schoolorganisatie en leiderschap zich nu al vele jaren concentreert op scholen in verandering en de rol van de schoolleider in veranderende contexten. Verschillende malen werd daarbij Larry Cuban geciteerd: 'reform again and again and again...'. Velen spraken dan ook hun twijfel uit over het nut van al het onderzoek en vroegen zich af of het onderzoek de realiteit wel dekt. Men probeert nieuwe metaforen en onderzoeksthema's te vinden 'that are better grasping reality'. Een aardige titel van een sessie in dit verband is: 'Educational administration: the field that mistook its research for a hat'. Misschien ook een thema om over na te denken met betrekking tot het Nederlandse onderzoek naar onderwijsbeleid, schoolorganisatie en schoolleiderschap.

Divisie B: Curriculum (Jan van den Akker, Universiteit Twente)

Laat mij starten met een bevinding die ik al vaker heb verkondigd (en die ik door veel collega's bevestigd hoor): curriculumonderzoekers kunnen zeer veel van hun gading vinden op de AERA, maar doorgaans meer buiten de Divisie Curriculum dan daarbinnen. Het hangt vooral af van de verdere inhoudelijke specialisatie of onderzoeksoriëntatie waar men de meest vruchtbare informatie aantreft.

Een hoogtepunt vanuit Nederlands perspectief was het symposium '*Curriculum evolution between policies and practices: Examples from Dutch curriculum research*'. De VOR-divisie Curriculum had dit jaar het genoegen het jaarlijkse trekkingsrecht van de VOR op zo'n invited AERA-sessie te mogen verzilveren. Onder voorzitterschap van schrijver dezes was een symposium belegd waarin het Nederlandse curriculumonderzoek vanuit een drietal geijkte curriculumperspectieven werd belicht. Het eerste paper (van Jan Terwel, Monique Volman en Wim Wardekker, VU) richtte zich op een inhoudelijke analyse van de curriculumvernieuwing in het voortgezet onderwijs in de

laatste decennia. Voorbeelden werden vooral ontleend aan het wiskunde- maar ook het talenonderwijs. En natuurlijk passeerden de basisvorming en Het Studiehuis de revue. Het tweede paper (van Wilmad Kuiper en Nienke Nieveen, UT) belichtte met name het technisch-professionele perspectief van curriculumontwikkeling. Daarin was veel aandacht voor hoe ontwikkelbenaderingen evolueren onder invloed van zowel beleidsveranderingen, lessen uit reconstructie-onderzoek naar ontwikkelpraktijken, als nieuwe ICT-tools voor ontwerpers. In het derde paper (van Jos Letschert, SLO en Joseph Kessels UT) domineerde het sociaal-politieke perspectief. De auteurs legden een verbinding tussen enerzijds onderzoek naar de totstandkoming van de kerndoelen in het basisonderwijs en anderzijds de relationele benadering van opleidingsontwikkeling in het bedrijfsleven. Zij vonden elkaar in een accent op het onontkoombare onderhandelingskarakter van veel curriculumontwikkelactiviteiten. Discussiant John Olson (Canada) plaatste een aantal boeiende opmerkingen. Hij onderstreepte het internationale karakter van veel gesignaleerde problemen en trends in de papers en illustreerde dat met verbindingen naar verwant onderzoek. Een greep uit zulke internationale thema's: discussies over de kwaliteit van het onderwijs; spanning tussen sociale en individuele accenten in het curriculum; de noodzaak tot meer flexibele en interactieve benaderingen van curriculumontwikkeling; de rol van leraren in curriculumontwerp en -implementatie; de implicaties en uitdagingen van ICT voor curriculumontwikkeling.

Een tweede vermeldenswaardige sessie had als titel: 'Does globalization imply fundamental curriculum change?' Verschillende sprekers vertolkten daarin hun visie dat globalisering het curriculum wel/niet/misschien zal veranderen. Larry Cuban (Stanford) beled daarin het standpunt dat er geen fundamentele veranderingen te verwachten zijn. Naar zijn mening zal, net als met andere externe invloeden op het onderwijs (bijv. ICT), ook globalisering wel wijzigingen teweeg brengen in het curriculumdenken op retorisch en formeel beleidsniveau, maar nauwelijks serieuze consequenties hebben voor wat er zich daadwerkelijk op micro-niveau (in leerboeken en klaspraktijken) afspeelt. Daartoe blijken bestaande gewoonten

en opvattingen in de schoolpraktijk over wat 'goed' onderwijs is te krachtig. Hoewel geen der overige sprekers overtuigd was van de komst van ingrijpende curriculaire veranderingen, werden toch enige trends gesignaleerd die dat enigszins plausibel maken. Meest in het oog springend daarbij: de verminderende exclusieve functie en efficiency van de school als leerinstituut, en de doorwerking van globalisering in examinering en certificering van leren. Beide verschuivingen kunnen op den duur ook implicaties hebben voor de doelen, inhouden en werkvormen in het curriculum.

Divisie C: Learning & Instruction (P. Robert-Jan Simons, Katholieke Universiteit Nijmegen)

Dit jaar mocht ik het 'slot' dat EARLI (the European Association for Research on Learning and Instruction) jaarlijks heeft in de AERA conferentie organiseren. Helaas was er het een en ander misgegaan: ons symposium over cscl (computer supported collaborative learning (met De Corte, Hakkarainen, Lehtinen, Scardamalia, Vosniadou)) was door een miscommunicatie tussen EARLI en AERA niet in het programmaboek opgenomen. Gelukkig was het wel prominent in het programmawijzigingenboekje opgenomen. Maar wij waren er niet gerust op dat er nog mensen zouden komen. Ik kwam pas maandagavond aan en dinsdagmorgen 8 uur zou het symposium al plaatsvinden. Twee van de presentatoren waren niet gekomen vanwege de ontstane onzekerheid. 's Morgens bleek de discussiant (Marlene Scardamalia) het vliegtuig gemist te hebben. Toch was het nog wel een geslaagd symposium waarin wij terugblikten op ons Europees project over samenwerkend computer gestuurd leren. Er waren toch nog ongeveer twintig mensen en we kregen Karl Bereiter (Marlene's man) als stand-in discussiant. De conclusie van het symposium was dat cscl, ondanks de positieve resultaten bij wiskunde en natuurkunde, vooral een betere inbedding in de lopende curricula nodig heeft en dat de rol van de docent veel aandacht behoeft, omdat docenten veel moeite blijken te hebben met de sturing van het leren: velen lieten alles gebeuren zonder enige sturing (een vorm van doorschieten).

Het meest bijgebleven van de AERA conferentie van dit jaar is voor mij de presidential

address van Lorrie Shepard. Komend van een achtergrond in de toetstheorie, keek zij (terug) naar de relatie tussen leertheorie (filosofie) en toetstheorie (meettheorie). Haar centrale stelling was dat er zich in de leertheorie in de afgelopen jaren een verschuiving heeft voorgedaan van behaviorisme naar constructivisme, cognitivisme en sociale leertheorie (deze drie schoof zij voor het gemak even op een hoop). Deze verschuiving in denken over leren is niet gepaard gegaan met een verschuiving in de toetstheorie. In de zestiger jaren was er een perfecte afstemming tussen de leertheorie en de toetstheorie. In beide lag de nadruk op een soort Taylorisme: nadruk op opdelen in kleine eenheden, objectiviteit en metingen per eenheid. De nieuwe leertheorieën benadrukken echter meer actief en zelfstandig leren, probleem oplossen, authenticiteit, holisme, leercultuur en leermotivatie. De vigerende toetstheorie is echter nog steeds hetzelfde gebleven. Onderzoek laat zien dat leerkrachten noch wetenschappers de traditionele behavioristische toetstheorie kunnen loslaten. Leerkrachten laten leerlingen bijvoorbeeld bewust twee soorten aantekeningen maken: die ze voor de toets nodig hebben en die echt belangrijk zijn. Onderzoek laat zien dat (traditionele) toetsen een belemmerend en soms zelfs desastreuze invloed hebben op het leerklimaat. Wetenschappers hebben eigenlijk nog geen alternatieve toetstheorie ontwikkeld en blijven bij de traditionele, behavioristische toetstheorie. Shepard's pleidooi voor een nieuwe toetstheorie, die beter aansluit bij de constructivistische (en cognitivistische, sociale) leertheorie, verdient mijn inziens alle steun. Zij stelt de volgende kenmerken van zo'n toetstheorie voor: nadruk op holistisch meten, authentieke situaties centraal, met inbegrip van vormen van self-assessment, gericht op het vaststellen van het leren en minder van de prestaties, met een accent op het diagnosticeren van mogelijkheden en noodzaak voor verder leren en gericht op het bijdragen aan een leerklimaat in plaats van de voorname belemmering ervan te vormen.

Ik had me veel voorgesteld van de lezing van Howard Gardner over zijn multipale intelligentietheorie. Helaas werd zijn presentatie gestoord door een soort brandalarm dat na een minuut of vijf afging en twintig minuten aanhield. Op bewonderenswaardige wijze heeft hij

geprobeerd er toch nog wat van te maken, maar het was, althans voor mij, onmogelijk om mij op zijn theorie te concentreren. Veel mensen liepen trouwens ook de zaal uit. Wat mij is bijgebleven is dat hij zijn theorie over zeven vormen van intelligentie heeft uitgebreid met een nieuwe vorm van intelligentie, die hij de existentiële intelligentie noemt. Het betreft de vorm van intelligentie die te maken heeft met het centraal stellen van ethische kwesties en levensvragen. Ook viel mij op dat Gardner vooral evidentie heeft gezocht voor zijn theorie in de evolutietheorie en in multidisciplinair onderzoek. In zijn antwoorden op de vele vragen over de onderwijskundige consequenties van zijn theorie was hij wat vaag. Hij zou binnenkort zijn eigen theorie gaan toepassen in zijn eigen onderwijspraktijk (!). Verder was zijn boodschap vooral dat goed onderwijs een beroep doet op alle acht de typen intelligentie en dus combinaties heeft van 'narrative', kwantitatief – logische benaderingen, existentiële benaderingen, 'hands-on experience' en kunstzinnige invalshoeken.

Tot slot noem ik nog enkele opvallende resultaten van onderzoek: reflectieve portfolio's hebben een groot effect op het zelfvertrouwen en de leerbereidheid van volwassenen en angst en boosheid zijn gerelateerd aan nieuwsgierigheid die op zijn beurt weer correlaties met job performance laat zien.

De meest interessante ervaring was overigens een informeel optreden van Gay Salomon tijdens een symposium over de vergelijking tussen verschillende typen media. Hij zat al de hele tijd op de punt van zijn stoel en kon zich bijna niet inhouden. Toen de plenaire discussietijd aanbrak, stak hij dan ook onmiddellijk van wal met een scherp betoog: het heeft geen zin om de effecten van verschillende media tegen de zelfde standaard af te zetten. Het gaat er juist om dat verschillende typen media (internet, real life, video, e.d.) verschillende typen uitkomsten realiseren. Zoek dus de sterke en zwakke kanten op van de media in plaats van het vergelijkend experimenteel onderzoek.

Divisie D: Measurement and Research

Methodology (Cees Glas, Faculteit Toegepaste Onderwijskunde, Universiteit Twente)

Dit jaar was 'Creating Knowledge in the 21st Century: Insights from Multiple Perspectives'

het thema van de AERA. Als men het aanbod van onderwerpen in Divisie D bekijkt, kan men constateren dat er zeker sprake was van verschillende perspectieven in methodologische benaderingen. Dat wil echter niet zeggen dat er veel communicatie tussen de benaderingen plaats vond. Het menu van Divisie D bestond uit een breed aanbod van kwalitatieve naar kwantitatieve methodologie. Om een indruk van het spectrum te geven, gaan we eerst naar de sessie 'Qualitative Research Methods: Case Studies and Others'. Kathleen Everling geeft een overzicht van onderzoeksparadigma's: positivisme, post-positivisme, constructivisme, post-modernisme, feminisme, marxisme en post-kolonialisme. Everling verklaart 'Post-colonial feminist' te zijn. Als mogelijke onderzoeksmethoden beschrijft ze, naast kwantitatieve en kwalitatieve methoden, methoden die 'arts-based', 'visual', 'historical' en 'philosophical' zijn. Uit haar onderzoek blijkt dat haar collega-docenten zich over het algemeen niet veel aan genoemde paradigma's gelegen laten liggen, studenten blijken er meer in geïnteresseerd te zijn. Daarna constateert Everling, met een beroep op Paulo Freire, dat het huidige onderzoek vooral leidt tot onderdrukking en ze vraagt zich af waarom we überhaupt onderzoek doen. Tenslotte breekt ze dan een lans voor de 'arts-based-approach': 'Perhaps the arts can begin to provide us a way to begin to dialogue about research'. Het vrolijke publiek in het overvolle zaaltje stelt geen kritische vragen, het is duidelijk dat de 'measurement community' al lang niet meer naar deze sessies komt kijken. Dat is jammer, compartimentering is een belangrijk principe voor het opslaan van vuurwerk, voor de wetenschappelijke discussie is het minder bevorderlijk.

Op plaatsen waar onderzoekers elkaar wel ontmoeten, kan die ontmoeting echter soms leiden tot een wat ongemakkelijk samenzijn. Sessies waar het label 'constructivisme' in de titel voorkomt, bestaan vaak uit een combinatie van presentaties van onderzoekers die inhoudelijk geïnteresseerd zijn in het epistemologische werk van Jean Piaget, onderzoekers die inhoudelijk geïnteresseerd zijn in constructivisme als pedagogisch principe (bijvoorbeeld aanhangers van de ideeën van John Dewey of onderzoekers op het gebied van onderwijs in exacte vakken) en onderzoekers die het constructivisme als

methodologisch principe omarmen. Een combinatie van een presentatie van Maria Trifonova over het getalsbegrip bij Piaget met een presentatie van Julie Rainer over 'constructivist pedagogy in primary classrooms' leidt niet direct tot een vruchtbare gedachtewisseling.

Nauw verwant met de opkomst van het constructivisme als pedagogisch principe is de tendens naar authentiek meten. In hun overzicht van de bijdragen aan Divisie D van de AERA in 1996 stellen Knoors en Dochy 'De laatste jaren is er sprake van een verschuiving van een test-cultuur naar een 'assessment'-cultuur; van een (puur) psychometrische aanpak naar een aanpak waarin de aandacht vooral uitgaat naar de authenticiteit van het testmateriaal'. Verder signaleren deze auteurs een '...stijgende bewustwording van het niet-authentieke karakter van examens en het verlies aan geloof hierin als valide toetsen van leerresultaten'. Even los van de vraag of deze observatie strookt met de werkelijkheid, is het interessant na te gaan of de verschuiving van psychometrische onderwerpen naar onderwerpen over authentiek meten in Divisie D zich heeft voortgezet. Divisie D omvatte dit jaar 73 sessies. Uit de titels van de sessies was op te maken dat 31 over psychometrische onderwerpen gingen, terwijl er slechts 2 sessies te vinden waren die over authentiek meten handelden. Daarnaast waren er 10 sessies over kwalitatieve onderzoeksmethoden, misschien kunnen we ook die nog rangschikken onder een kopje 'verschuiving in de meetcultuur'. Uit mijn uiteraard nogal informeel onderzoekje blijkt dat de verschuiving naar authentiek meten, althans in Divisie D, niet verder doorzet.

Verder wandelend langs het spectrum komen we bij de 'measurement community' die behalve in Divisie D ook op de parallel aan de AERA gehouden Annual Meeting of the NCME (National Committee for Measurement in Education) bijeenkomt. Weinig theoretische presentaties dit jaar. Vorig jaar was de meest opmerkelijke theoretische sessie die waar Patz en Junker enerzijds, en Wilson en Hoskens anderzijds, concurrerende IRT modellen presenteerden voor het analyseren van data van multiple beoordelingen. Dit jaar presenteerden ze een gezamenlijk paper waarin hun benaderingen vergeleken werden. Het model van Wilson en Hoskens blijkt zeer robuust en een-

voudig te schatten, het model van Patz en Junker lijkt theoretisch beter onderbouwd, maar hun Bayesiaanse MCMC schattingstechniek is op dit ogenblik nog behoorlijk tijdrovend.

De nadruk lag dit jaar echter vooral op praktische aangelegenheden. Naast 'computerized adaptive testing', 'differential item functioning' en 'test equating' was het belangrijkste topic 'accountability and standard setting'. Een mooi overzicht van de problematiek werd gegeven in de sessie 'Beyond Magical Thinking: Designing Ideal Assessment and Accountability Systems', onder andere met Robert Linn en Edward Haertel. Om tot een verbetering van het onderwijs te komen, zijn er in de Verenigde Staten plannen ontwikkeld om de prestaties van scholen te evalueren met behulp van speciaal voor dat doel ontwikkelde toetsen. In Nederland zijn vergelijkbare voorstellen gedaan door de Onderwijsraad. De ontwikkeling van het algemene niveau van het onderwijs wordt in Amerika reeds geëvalueerd in NEAP (National Assessment of Educational Progress), in Nederland gebeurt dat in het PPON (Periodieke Peiling van het Onderwijsniveau). Deze projecten zijn echter niet op de prestaties van specifieke scholen gericht, het gaat hier meer om de globale trend in het onderwijsniveau. Een van de sprekers merkte op dat NEAP bij het grote publiek nauwelijks bekend is, de situatie van het PPON zal in Nederland nauwelijks anders zijn. Het is te verwachten dat een project dat specifiek op scholen gericht is veel meer belangstelling zal genereren. Een probleem in de Verenigde Staten is dat het onderwijs per staat geregeld is, zodat vergelijkingen tussen staten alleen via vrijwillige toetsen gemaakt kunnen worden. Vandaar het initiatief een aantal Voluntary National Tests (VNT) te ontwikkelen. Opvallend is dat ook in de 'Measurement Community' kritische kanttekeningen bij de voorstellen geplaatst worden. Linn merkte op dat vooral politici het voorstel omarmen omdat het een relatief goedkope manier is om iets aan het niveau van het onderwijs te doen. De effectiviteit van die aanpak staat uiteraard allerminst niet vast. Linn gaf een aantal aanbevelingen om de kwaliteit van de meeting te waarborgen. Zorg ervoor dat de scholen de steekproefleerlingen niet kunnen manipuleren, gebruik meerdere toetsen en indicatoren, gebruik multilevelmodellen voor de

analyses en rapporteer nadrukkelijk de (on)nauwkeurigheid van de resultaten. Een andere aanbeveling van Linn was om zowel resultaten die gecorrigeerd zijn voor achtergrondvariabelen (zoals SES) als ongecorrigeerde resultaten te rapporteren. Het eerste om scholen niet onrechtvaardig te beoordelen, het tweede als een 'safeguard against institutionalizing low expectations'. Haertel bepleitte om meer nadruk te leggen op de longitudinale ontwikkeling binnen scholen dan op vergelijking van scholen onderling. Overwegingen die ook in de Nederlandse situatie aandacht verdienen.

Het NVT-initiatief genereert ook nieuw psychometrisch onderzoek. In verschillende sessies werd aandacht besteed aan het equiveren van de jaarlijkse afnamen van de NVT en het linken van NVT, NEAP en TIMMS gegevens. De aanbevelingen kwamen neer op een inhoudelijke afstemming van de te meten constructen en het via IRT calibreren van de verschillende toetsen. Over het algemeen werd door de presentatoren (Holland, Muraki, Kolen, Van der Linden) veel nadruk gelegd op unidimensionaliteit van vaardigheidsstructuren. Eugene Johnson onderzocht mogelijkheden om toetsen met gecorreleerde vaardigheidsstructuren te equiveren. Zijn conclusie was negatief omdat uit simulatiestudies bleek dat het verlies aan nauwkeurigheid vrij groot was.

Alles overziend mogen we dus zeggen dat er voldoende leven is in de brouwerij van 'Measurement and Research Methodology'.

Divisie G: Social Context of Education

(G. Driessen, ITS)

Gerelateerd aan Divisie G werd dit jaar, zoals gebruikelijk, een groot aantal sessies georganiseerd door de Special Interest Group 'Research Focus on Black Education'. Verreweg de meeste deelnemers aan deze overigens goed bezochte sessies waren African-Americans; Caucasians waren slechts matig vertegenwoordigd. De papers vertegenwoordigden een brede range aan topics en methoden, alhoewel exploratief onderzoek en kwantitatief onderzoek met minder dan 150 respondenten de boventoon voerden.

Interessant was het paper van McMillian en Meece over de ontwikkeling van academische

identificatie bij Afrikaans-Amerikaanse kinderen. Zij gaan uit van de 'disidentification hypothesis' die stelt dat disidentification een zelfbeschermende strategie is, waardoor leerlingen hun zelfbeeld niet meer relateren aan academische gebieden, maar zich richten op andere gebieden. In een studie onder 164 jongens en meisjes uit laag-inkomen milieus werden de relaties tussen gender, wiskunde en leesprestaties, en zelfbeeld geanalyseerd. De resultaten geven aan dat, hoewel jongens een positiever zelfbeeld hebben en meisjes betere leesprestaties leveren, de correlatie tussen leesprestaties en zelfbeeld hoger is bij meisjes.

Ook interessant was het paper van Williams-Smith en Causey die twee case-studies uitvoerden bij Charter Schools. De onderzoeksvraag van de eerste studie was of Causey's Transformational Mentoring Model (meer dan één mentor) tot betere resultaten leidt dan de traditionele aanpak met slechts één mentor. Pre- en post-test vergelijkende analyses toonden aan dat geen van de bevindingen significant was. Toch zijn volgens de onderzoekers de onderzoeksresultaten van onderwijskundig belang, vanwege de speciale behoeften en achtergronden van deze zwarte achterstandsleerlingen. De tweede studie vergeleek leerlingen in Charter Schools met leerlingen in traditionele 'non magnet public schools'. De resultaten op de Charter Schools bleken niet alleen significant beter dan de gemiddelde scores van beide vergelijkingsgroepen, maar zelfs significant hoger dan de hoogste scores in de range van beide vergelijkingsgroepen.

In een studie van Gladney en Greene werd de relatie tussen fictieve verwantschap, raciale identiteit, peer invloed, houding tegenover school en toekomstige doelen enerzijds, en prestaties anderzijds van 270 Afrikaans-Amerikaanse high school leerlingen onderzocht. De resultaten van een serie regressie-analyses lieten zien dat houding tegenover school en bepaalde indicatoren van raciale identiteit de sterkste voorspellers zijn van prestaties, hetgeen betekent dat voor Afrikaans-Amerikaanse leerlingen een positieve houding tegenover school en een sterk onbehagen met betrekking tot hun identiteit samengaat met hoge prestaties. Aan de andere kant rapporteerden leerlingen die geen problemen hadden met hun iden-

titeit hoge toekomstige doelen, terwijl degenen die negatieve invloeden van peers ervoeren laag scoorden op toekomstige doelen.

Miller en Dedrick voerden een longitudinale studie uit naar het sociaal en academisch zelfbeeld van begaafde en niet-begaafde Afrikaans-Amerikaanse kinderen. Ondanks onderwijshervormingen zijn Afrikaans-Amerikaanse leerlingen nog steeds ondervertegenwoordigd in programma's voor begaafde kinderen. Mogelijk is dit te wijten aan bias in intelligentie-tests en gebrek aan aandacht voor culturele factoren. Om aan dit probleem tegemoet te komen werden alternatieve selectieprocedures voor kinderen uit minderheden ontwikkeld, waarin gebruik werd gemaakt van verschillende informatiebronnen. De studie betrof 148 leerlingen. Drie groepen werden onderscheiden: leerlingen die toegelaten waren tot de programma's voor begaafden via de traditionele benadering, leerlingen die via de alternatieve benadering werden toegelaten, en leerlingen die genomineerd waren, maar niet geschikt werden bevonden. Groei-curve analyse (multi-level) werd gehanteerd om de veranderingen in zelfbeeld te meten. Afrikaans-Amerikaanse leerlingen in het programma voor begaafden (via beide methoden) bleken relatief positieve gevoelens over zichzelf te hebben. Daartegenover toonden niet gekozen leerlingen negatieve veranderingen in hun zelfbeeld.

Een symposium dat een interessante discussie had kunnen opleveren was 'Multicultural Education in the 21st Century', voorgezeten door James Banks en met kopstukken uit de Verenigde Staten en United Kingdom. Na inleidende statements door elk van de participanten over onder andere de zwakke link tussen theorie en praktijk, multi-cultureel onderwijs als feit en als ideaal, de relatie met sociaal milieu en gender, objectiviteit, kwalitatief versus kwantitatief onderzoek, proletarisering van de middenklasse en anti-racistisch onderwijs, was de sessietijd helaas verstreken voordat een discussie van de grond kon komen – een gemis te kans.

Divisie H: School Evaluation and Program Development (R.J. Bosker, Universiteit Twente)

Uit de veelheid van sessies met name bericht over twee symposia die ook voor de Neder-

landse situatie niet van belang ontbloot zijn. Het onderwerp van het eerste symposium sloeg ook bij Amerikanen aan, want maar liefst 300 aanwezigen werden aangetrokken door de uitdagende titel 'school scandals waiting to happen: consequences of current pressures to raise test scores', of door de namen van de inleiders (o.a. James Popham). Wie echter verwacht had dat de inleiders de vloer zouden aanvegen met de test-praktijk kwam bedrogen uit. Lygon begon met te betogen dat er altijd bedriegers zijn, of het nu gaat om belastingformulieren, sporten, verkeersregels, of om testen en dat er goede manieren zijn om dit probleem op te vangen. Het feit dat er bedriegers zijn is in elk geval geen reden om er mee op te houden. Perlman ging vervolgens in op de bijzonderheid dat de tests nu niet alleen meer 'high stakes' zijn voor leerlingen, maar ook voor scholen en leerkrachten. Er is nu een nieuwe industrie op komst, die mede vanwege dat laatste, scholen en leerkrachten hulp biedt bij het voorbereiden van leerlingen op de toetsafname, zulks natuurlijk met in acht neming van de wettelijke voorschriften (d.w.z. de leerlingen leren in het algemeen met testen om te gaan, niet specifiek met de items die in de diverse tests aan de orde zijn). Stiggins ging in haar bijdrage in op de taxonomie van denkvaardigheden (reproduceren – analyseren – vergelijken – concluderen – evalueren), en maakte aannemelijk dat als in het onderwijs aan de hogere orde denkvaardigheden aandacht wordt besteed, leerlingen ook betere testcores zullen behalen, want deze vormen de eigenlijke 'latente trek'. Aan Popham tenslotte de eer toch nog een kritische noot te kraken: veel van de testen zijn norm- (en niet criterium-) georiënteerd. Dat betekent dat leerkrachten met hun klassen vaak geen vooruitgang kunnen boeken, omdat ze in de normpopulatie afhankelijk zijn van de prestaties van anderen. Een pleidooi dus voor criteriumgerelateerde toetsen met daarin een absolute standaard gedefinieerd. Al met al vormt het-geen gepresenteerd werd in het symposium geen aanleiding met de CITO-toetsen te stoppen, of om de praktijk van het openbaar maken van schoolgemiddelde CITO-scores te staken, wel om nog eens na te denken over absolute standaarden voor de evaluatie van leerlingvorderingen en voor de evaluatie van prestaties van scholen.

In het symposium over klassengrootte werden de laatste resultaten van het STAR-experiment bekend gemaakt. De onderzoekers heranalyseerden oude gegevens uit dit spraakmakende experiment naar de effecten van klassengrootte in Tennessee en presenteerden nieuwe gegevens over het beklijven van effecten. Gerber en Finn richtten zich in hun bijdrage op het effect van de klassenassistent. In het oorspronkelijke experiment waren er immers drie condities: kleine klassen (13-17 leerlingen), grote klassen (22-25 leerlingen) en grote klassen met een klassenassistent. Dat een klassenassistent geen nut heeft (in termen van verbeterde leerlingprestaties) is een bevinding die is ondergesneeuwd in alle heisa rond dit experiment. De heranalyse bevestigde dit onderzoeksresultaat nog eens, en maakt eveneens aannemelijk dat deze bevinding niet gedifferentieerd hoeft te worden naar het soort activiteit dat de klassenassistent uitvoert (administratie, instructie, anders). De auteurs concluderen dan ook onomwonden dat 'if the primary function of paraprofessionals is to enhance pupils' academic achievement, this expenditure represents a significant misuse of resources.' De andere bijdragen richtten zich op de lange termijn effecten van de kleine klas (vergeleken met de grote). Hier zijn er een paar:

1. verhoogde kansen op tijdige afronding van de middelbare school, op deelname aan de moeilijkere cursussen in de middelbare school, op deelname aan een toelatingsexamen voor de universiteit (met name voor allochtone leerlingen) en op universiteitsbezoek,
2. verlaagde kansen om voortijdig de school te verlaten, en
3. voor degene die uiteindelijk in de gevangenis belanden (1,1 procent van de onderzoeksgroep): lagere straffen voor de leerlingen uit de kleinere klassen, maar dit laatste effect kon de significantietoets niet doorstaan.

De opbrengsten zijn volgens de auteurs derhalve groot. Men vraagt zich echter af waarom die effecten dan optreden. Waarom niet de causale keten ontrafeld? Het moet toch aannemelijk gemaakt kunnen worden dat beginnende geletterd- en gecijferdheid (de primaire opbrengsten van de kleine klassen) deze effecten interne-

diëren? Zonder een dergelijke analyse blijft het 'fact finding'. Naast de symposia was er wederom een aantal interessante papers over de 'charter school' beweging. Maar aangezien al onze Nederlandse scholen eigenlijk 'charter schools' zijn (al is het idee dat scholen op contractbasis werken misschien zo gek nog niet – vergelijk het onderwijskansenplan) waren de lezingen iets minder wereldschokkend. Ook dit jaar dus binnen de algemene container van het evaluatie-onderzoek genoeg wetenschappelijk en maatschappelijk interessante vraagstukken en interessante nieuwe inzichten, en wellicht zouden we de Amerikanen eens moeten trakteren op een symposium over hoe onze scholen opereren onder marktcondities.

Divisie K: Teaching and Teacher Education

(J. van Driel, ICLON, Universiteit Leiden)

Bij deze Divisie ging dit jaar veel aandacht uit naar vragen omtrent de effecten van lerarenopleidingen en nascholingsprogramma's. Zo had M. Cochran-Smith haar lezing als vice-president van Divisie K als titel 'The outcomes question in teacher education' gegeven. Hierin plaatste zij de huidige belangstelling voor *outcomes* in een historisch perspectief waarin het onderzoek naar leraren en lerarenopleiding achtereenvolgens is gedomineerd door attributies, effectiviteit van leraren en cognities van leraren. Zij wees op resultaten van studies waaruit is gebleken hoe belangrijk de invloed van de kwalificaties en de opleiding van leraren is op leerresultaten van leerlingen. Tegelijkertijd vermeldde zij het bestaan van studies met hele andere uitkomsten. Uiteindelijk kreeg haar betoog een politieke teneur waarin zij stelling nam tegen marktwerking in het onderwijs en tegen een te zwaar accent op standaarden, omdat dergelijke ontwikkelingen in het nadeel zouden zijn van scholen met een grote culturele diversiteit.

In een andere sessie reageerden enkele bekende onderzoekers op de kritiek waaraan lerarenopleidingen in de Verenigde Staten in toenemende mate worden blootgesteld. Die kritiek komt er, kort gezegd, op neer dat lerarenopleidingen weinig niveau hebben ('Micky Mouse-cursussen') en teveel gericht zijn op 'softe' zaken, zoals reflectie. In plaats daarvan zouden lerarenopleidingen, volgens de

critici, leraren moeten toerusten met kennis en vaardigheden om feiten over te dragen, orde te houden, enzovoort. Het meest overtuigend waren de reacties van L. Darling-Hammond en D. Berliner op deze kritiek. Beiden baseerden zich hierbij voornamelijk op 'harde', empirische studies waaruit onder meer naar voren is gekomen dat er een behoorlijk sterk verband bestaat tussen de didactische voorbereiding en de pedagogical knowledge van leraren en hun effectiviteit. Ook blijkt hieruit dat leerlingen die worden onderwezen door gekwalificeerde leraren betere leerresultaten boeken en minder uitvallen. Het idee dat leraren eigenlijk genoeg hebben aan een bepaalde dosis vakkennis wordt door dergelijk onderzoek duidelijk onderuit gehaald. Het verkorten van lerarenopleidingen is dan ook geen goed middel om lerarentekorten tegen te gaan, te meer omdat veel leraren na een dergelijke opleiding al vlot het beroep vaarwel zeggen.

De betekenis van het bovenstaande voor het ontwerp van lerarenopleidingen en nascholingsprogramma's voor specifieke doelgroepen kwam naar voren in diverse andere sessies. Hierin werden kenmerken beschreven van opleidingen die daadwerkelijk effect hebben ('make a difference'). Hierbij werd vaak gewezen op de gunstige effecten van verschillende vormen van samenwerking en uitwisseling tussen leraren. Bij initiële opleidingen ging het bijvoorbeeld om samenwerking tussen beginners en expert leraren, waarbij de beginners konden experimenteren met nieuwe werkvormen. Bij nascholing, ofwel *professional development*, was de belangrijkste trend het opzetten van *communities* van leraren, waarin ervaringen uitgewisseld worden, artikelen gezamenlijk bestudeerd worden, enzovoort. Voor dergelijke communities is een duidelijke structuur nodig, alsmede voldoende tijd en geld. Uit onderzoek bleek dat de 'intensiteit' van zulke communities positief gecorreleerd is met studieresultaten van leerlingen.

In het onderzoek naar de praktijk van het onderwijzen wordt in toenemende mate gebruik gemaakt van de mogelijkheden die multimedia bieden om een zo compleet en authentiek mogelijk beeld van de lespraktijk vast te leggen. Zo schetste M. Lampert een aanpak waarin videobeelden van lessen in een database werden gecombineerd met volledige trans-

cripten van deze lessen, persoonlijke gegevens van de leerlingen en hun werk, logboeken van de docent, enzovoort. Het bleek echter lastig om in een dergelijke 'zee' van gegevens het hoofd boven water te houden. Een theoretisch kader is daarbij onontbeerlijk, zoals bleek uit een presentatie van het project *Learning to Teach Secondary Mathematics* (LTSM; H. Borko et al.). In dit project, dat nog enige tijd doorloopt, wordt de ontwikkeling van cognities van leraren bestudeerd in relatie tot hun gedragingen in de praktijk vanuit een situationeel perspectief. Hierbij wordt een groep aanstaande wiskundeleraren gevolgd tijdens hun opleiding en hun eerste jaar als leraar. Registraties van lessen worden geanalyseerd in samenhang met gegevens die betrekking hebben op de cognities van de leraren over domeinen als 'vakdidactiek' en 'professionele identiteit'. Het situationeel perspectief blijkt hierbij vooralsnog zeer bruikbaar te zijn. Onderzoek van dit type markeert een duidelijke progressie ten opzichte van het 'traditionele' onderzoek naar *teachers' knowledge and beliefs* dat enigszins is blijven hangen in eindeloze reeksen van gevalbeschrijvingen.

Referenties

De lezing van M. Cochran-Smith is te vinden op de website van haar instituut:

<http://www.bc.edu/teachered>

L. Darling-Hammond verwees naar een artikel van haar hand dat is te vinden bij de Education Policy Analysis Archives:

<http://epaa.asu.edu/epaa/v8n1/>

Informatie over het LTSM project is te vinden op <http://ltsm.colorado.edu/ltsm/index.html>

Aan deze kroniek werkten mee: J. van den Akker, R. Bosker, J. van Driel, G. Driessen, C. Glas, M. Krüger, R.J. Simons, B. Witziers.

De eindredactie was in handen van H. Dekkers.