

**Annual Meeting American  
Educational Research  
Association  
18-22 april 1995  
San Francisco**

---

Inleiding (J. N. Streumer, Universiteit Twente)

De jaarlijkse conferentie van de American Educational Research Association heeft in 1995 minstens in één opzicht een record gebroken: meer dan 11.000 bezoekers werden geregistreerd. Dat impliceert dat ongeveer de helft van alle AERA-leden (in totaal 22.000) was afgereisd naar San Francisco. In totaal werden 1461 sessies georganiseerd. Iets meer dan de helft van de ingezonden papervoorstellen werd goedgekeurd en in het programma opgenomen. Dit wijkt niet af van voorafgaande jaren. Het programma bestond uit paper- en symposiumsessies, experimentele sessies, trainingsactiviteiten, poster sessies, presentaties voor nieuwe leden en graduate studenten, invited presentations, tentoonstellingen, off-site sessions en talrijke bijeenkomsten met een zakelijk en/of gezelligheidskarakter.

Het thema voor 1995 was: Partnership for a New America in a Global Community.

Ook in 1995 waren de Nederlanders goed vertegenwoordigd, zowel in de rol van conferentiebezoeker als paperpresentator. Nauwkeurige statistieken over 1995 wat betreft de nationaliteit van deelnemers en paperpresentatoren zijn nog niet bekend. De verwachting is echter dat het aantal presentatoren dit jaar hoger is uitgevallen dan vorig jaar. Om een idee te geven van de Nederlandse aanwezigheid in 1994 volgen hieronder enige gegevens, ontleend aan een artikel van Richard White in het augustus-september nummer van de Educational Researcher. Van een totaal van 4674 presentatoren in 1994 waren er 37 met een Nederlandse nationaliteit en 5 met een Belgische nationaliteit. Nederland eindigde wat betreft het aantal deelnemende presentatoren op de

vijfde plaats, voorafgegaan door de Verenigde Staten (4100 presentatoren), Canada (235), Australië (81), Engeland (77). België eindigde op een gedeelde 11e plaats.

Dit jaar zal de AERA-conferentie plaats vinden in New York van 8-12 april. Het thema zal zijn: Research for Education in a Democratic Society. De programmavoorzitter voor 1996, Ann Lieberman, zegt daarover: "In general, the theme suggests the importance of educational research to a variety of issues that deal with learning, teaching, and access to knowledge for increasing numbers of students in our own country and around the world. In particular it asks that we work to achieve a better understanding of the many new small learning communities being created today that are successfully engaging a diverse population of students. And further, that we look carefully at small- and large-scale efforts that are attempting to transform relationships between district and school, school and family, teacher and student, and researcher and practitioner (Lieberman, 1995 in: Educational Researcher, June-July 1995, p. 43).

De papervoorstellen voor dit jaar zijn inmiddels al weer becommentarieerd en de inzenders zijn op de hoogte van het feit of hun paper is geaccepteerd en geprogrammeerd. Zal de trend van de afgelopen jaren wat betreft een groeiend aantal Nederlandstalige presentaties en deelnemers zich ook dit jaar weer voortzetten?

**Beroeps- en bedrijfsopleidingen (M. Mulder, Universiteit Twente)**

Het gebied van de beroeps- en bedrijfsopleidingen wordt binnen de AERA niet afgedekt door een volwaardige divisie. Er zijn diverse Special Interest Groups (SIGs) op het gebied van de bedrijfsopleidingen, het beroepsonderwijs, opleidingen binnen defensie, opleidingen binnen de bedrijfskunde en de volwasseneneducatie. Aangezien ik als programmavoorzitter verantwoordelijk was voor het programma van de SIG Training in Business and Industry zal ik vooral verslag doen over de bijdragen van deze SIG. Het programma bestond uit vier sessies: twee gewone papersessies, een sessie voor zogenaamde roundtables waar geïnteresseerden intensief van gedachten kunnen wisselen over

een paper van de onderzoeker, en een business meeting waarin A. Tuijnman een inleiding zou houden over de relatie tussen opleiding en economische ontwikkeling, maar helaas moest hij enige tijd voor de conferentie afzeggen.

De belangstelling voor het programma van de SIG is al jaren vrij constant. Meestal varieert het bezoekersaantal bij de sessies van twintig tot dertig personen, met uitschieters naar veertig tot vijftig personen, wat voor een SIG redelijk goed genoemd kan worden. Ervaren bezoekers van de AERA weten dat er soms bijeenkomsten van divisies plaatsvinden in zalen met een capaciteit van honderden personen, terwijl er dan vier of vijf inleiders verslag doen van hun onderzoek voor nog geen tien bezoekers in de zaal. Ook in 1995 was dat weer het geval.

De onderwerpen van de sessies in de SIG waren 'de positie van evaluatie, team oriëntatie en klanttevredenheid', 'evaluatie van Electronic performance systems, training courseware en management training', en 'arbeidsbetrokkenheid, evaluatie en kwaliteit van opleidingsprogramma's'. Zoals zo vaak tijdens internationale wetenschappelijke conferenties, was ook binnen dit programma sprake van weinig coherentie. De onderwerpen die aan de orde kwamen, waren zeer uiteenlopend en zeer verschillend van kwaliteit.

Ik beperk me bij het bespreken van de papers tot enkele naar mijn oordeel redelijk goede. Daarbij laat ik mijn eigen bijdrage aan het programma buiten beschouwing. Ik laat het liever aan anderen over om daar een kwaliteitsoordeel over uit te spreken.

Allereerst het onderzoek van Gauthier van de Universiteit van Montreal. Zij heeft een onderzoek gedaan naar de transfer van de resultaten van een opleiding naar de werksituatie van managers en medewerkers van een instelling in de dienstensector. Dit is een belangrijk onderzoeksthema omdat vaak wordt gesteld dat maar tien tot twintig procent van de inhoud van opleidingen leidt tot werkelijke transfer op de werkplek. Het onderzoek borduurt voort op een conceptueel raamwerk van Baldwin en Ford, waarbij naast opleidingscondities, persoonskarakteristieken en kenmerken van de werkomgeving worden gezien als factoren die leren en retentie alsmede korte en lange termijn transfer beïnvloeden. Het onderzoek van Gauthier richt zich op de vraag of het organisatieklimaat en de

betrokkenheid bij het werk van invloed zijn op de transfer. In het onderzoek is gebruik gemaakt van vier experimentele groepen en een controle-groep ( $n = 75$ ). Bij de verschillende experimentele groepen zijn op verschillende momenten gegevens verzameld. Als methoden zijn gebruikt: een observatie door externe observatoren van het werk twee maanden voor en tien maanden na de opleiding en vragenlijsten die zijn ingevuld door de deelnemers aan de opleiding. Het gaat om vier vragenlijsten voor het meten van respectievelijk de beheersing van de opleidingsdoelen, het werkgedrag in de werksituatie, de betrokkenheid bij het werk en het organisatieklimaat. Uit de resultaten van het onderzoek, die overigens niet zo helder worden beschreven, blijkt dat er sprake is van leren, retentie en opleidingseffecten, en dat betrokkenheid van het werk en organisatieklimaat inderdaad van invloed zijn op de transfer.

Vervolgens het onderzoek van Russ-Eft en Ravishankar van Zenger Miller in San Jose. Zij onderzochten de houding van managers en medewerkers ten opzichte van hun werk en vergeleken daarbij organisaties naar de mate waarin zij verschillen in hun oriëntatie op teams. In veel organisaties wordt gewerkt aan teamvorming, omdat veelal gestipuleerd wordt dat het werken in teams als resultaatverantwoordelijke eenheden leidt tot een grotere efficiëntie en effectiviteit van het werk. De onderzoekers hebben 81 managers en 364 medewerkers uit tien organisaties een oordeel laten geven van hun functie en organisatie. Daarnaast hebben onafhankelijke experts op het gebied van zelfgestuurde werkteams een oordeel gegeven over de mate waarin de organisatie is gericht op het werken in teams. Deze gerichtheid van organisaties duiden de onderzoekers aan met de term 'team oriëntatie'. Uit de resultaten van het onderzoek blijkt dat de team oriëntatie van organisaties inderdaad samenhangt met de houding van managers en medewerkers ten opzichte van de eigen functie en organisatie. Bovendien blijkt dat in teamgerichte organisaties de oordelen van managers en leidinggevendenden over hun functie en de organisatie meer met elkaar overeenkomen dan in organisaties met een lagere mate van team oriëntatie. De onderzoekers concluderen dat team oriëntatie een belangrijk middel is om het klimaat in organisaties te verbeteren. De vraag is echter of op

grond van het onderzoek wel gesproken kan worden over een causale relatie tussen team oriëntatie en organisatieklimaat. Er zijn mijns inziens evenzovele gronden om aan te nemen dat het organisatieklimaat van invloed is op de mate van team oriëntatie. Als antwoord op die vraag zou het interessant zijn om na te gaan in welke mate opleidingen in team oriëntatie van invloed zijn op het verbeteren van het organisatieklimaat.

Tenslotte het onderzoek van MacDonald, Bradley Cousins, Bailetti en Rahman van de Universiteit van Ottawa en de Carleton Universiteit. Zij hebben een model geconstrueerd voor betekenisvolle evaluatie van managementopleidingen in ondernemingen met een sterk technologisch karakter. Zij zien dit model als een alternatief voor de vele evaluaties waarbij alleen wordt ingegaan op de reacties van de deelnemers of waarbij de verhouding tussen kosten en baten wordt geïventariseerd. Het uitgangspunt van de onderzoekers is dat eerst moet worden nagegaan wat de factoren zijn die het succes bepalen van produktontwikkelingsprojecten. Zij hebben daarvoor tien managers uit negen verschillende ondernemingen een lijst van 47 succesfactoren voorgelegd. Deze personen hebben die lijst beoordeeld met behulp van een kaartsorteertechniek en importantiëweging. De gegevens zijn verwerkt met behulp van multi-dimensionale schaaltechnieken en een clusteranalyse. Het resultaat is een model met karakteristieken van individuen, groepen en werkobjecten die succesvol zijn bij produktontwikkelingsprojecten. Dit model kan worden gebruikt bij evaluaties van managementopleidingen in technologisch geïntereerde ondernemingen.

Er zullen dit jaar voor het eerst proceedings verschijnen van de papers die zijn gepresenteerd in het programma van de SIG Training in Business and Industry. Deze proceedings zullen worden uitgegeven door de Faculteit der Toegepaste Onderwijskunde onder de titel 'AERA Research Papers on Training in Business and Industry 1995'.

De SIG op het gebied van het beroepsonderwijs is voornemens een divisie op te richten getiteld 'Education and Work'. Initiatiefnemer is prof. dr. Curtis Finch. Het is de bedoeling om niet te

concurreren met de bestaande SIGs, ofschoon de oprichting van deze divisie wel de nodige effecten zal hebben op de SIGs. Veel leden van verschillende SIGs zullen namelijk belangstelling hebben voor deze nieuwe divisie. Hoe dan ook, de intentie is de divisie op te richten naast de bestaande SIGs, zodat deze hun eigen activiteiten gewoon kunnen voortzetten.

Vergeleken met de Nederlandse situatie binnen de Vereniging voor Onderwijsresearch zou deze AERA-divisie verwant zijn aan de VOR-divisie Beroeps- en Bedrijfsopleidingen. Ik ondersteun het initiatief om te komen tot deze divisie omdat ik denk dat het veel voordelen biedt voor onderzoekers op het terrein van de beroeps- en bedrijfsopleidingen. Nu is hun onderzoek sterk versplinterd in allerlei SIGs binnen de AERA. De beoogde divisie kan waarschijnlijk meer synergie in de presentatie van onderzoeksresultaten bewerkstelligen. Voorwaarde voor het oprichten van een divisie binnen de AERA is echter dat een substantieel deel van de AERA-leden het initiatief ondersteunt. Geïnteresseerde VOR-leden van de divisie Beroeps- en Bedrijfsopleidingen kunnen zich voor verdere informatie melden bij de auteur van dit deel van de AERA-Kroniek. Hij treedt op als contactpersoon voor de Amerikaanse initiatiefnemer.

### **Instructie- en Communicatietechnologieën**

(J.M. Pieters, Universiteit Twente)

De tijd dat er op de AERA presentaties werden verzorgd waarbij de onderzoeker zijn of haar resultaten presenteerde met handgeschreven en/of moeilijk leesbare transparanten lijkt voorgoed voorbij. De presentaties zijn goed verzorgd en de moderne media hebben hun intrede gedaan. Het gebruik van Power Point of vergelijkbare presentatiesystemen, off-line dan wel on-line, is vanzelfsprekend. Bij een aantal sessies zag ik de discussiant tijdens de presentaties zijn of haar discussie voorbereiden op de computer, vanzelfsprekend van notebook formaat. De discussie zelf was daardoor zeer to-the-point en visueel van hoog niveau. Ook de dynamica heeft zijn intrede gedaan. Tijdens de presentatie dynamische voorstellingen het betoog laten ondersteunen, soms zelfs met aansluiting op een externe computer, komt ook voor.

Is er dan ook aan de inhoud van de presentaties ook een en ander veranderd? Jazeker, de nieuwe technologieën bieden mogelijkheden op ruimere schaal ingenieuzer te experimenteren, gegevens te verzamelen en te verwerken, maar vooral de leeromgeving van een nieuwe dimensie te voorzien, namelijk interactiviteit. Interactiviteit kan zich op verschillende manieren openbaren: in kennisnetwerken, in kennisconstructie met een scala aan participanten, in collaboratie, en in situationisme en realisme.

Wanneer we de presentaties in Divisie C overzien dan constateren we een overwegend gebruik van termen als: constructivism, interacting, (learning) community, knowledge organization, communication, computer-mediated instruction, learning environments, discourse communities, en virtual communities. Een goed voorbeeld van het gebruik van nieuwe technologieën in het onderwijs (in constructivistische zin) gaf het symposium met de titel 'Applications in networked multimedia for knowledge construction, collaboration, and curriculum reform'. Een sterke nadruk lag in dat symposium op de technische toepassingsmogelijkheden waarbij de theoretische en in het verlengde daarvan empirische onderbouwing het moest afleggen tegen de aantrekkingskracht van het nieuwe medium, ondanks het gebruik van de termen knowledge construction en collaboration. Ook in het symposium 'The Learning Community: A Star schools project building a network of school-community partnership through multiple technologies' waarin verslag gedaan werd van de eerste ervaringen in dit grootschalige project, overheerste het imponerende karakter van de nieuwe technologieën en optimistische visies op een succesvolle implementatie. Deze door de technologie gedomineerde aanpak, eerst de technische implementatie en pas daarna je druk maken over de theoretische verantwoording, stak schril af tegen de wat meer pessimistische, misschien ook wel meer realistische kijk op de zaak van door de onderwijskundig-theoretische en -praktische wol geverfde onderzoekers als Fullan, Rothkopf en Berliner die in het symposium 'Getting instructional research results in our schools: A loveless marriage, temporary misunderstanding, or promising partnership' de traditionele R&D benadering (met eigentijdse variant) verdedigden. Een tussenpositie

nemen de onderzoekers in die in het symposium 'Schools for thought: Transforming classrooms into learning communities' vanuit zeer geprononceerde theoretische opvattingen fraaie praktische telematische innovaties presenteerden. Innovaties die naar de mening van de onderzoekers, gesteund door weliswaar beperkte empirische evaluaties, er wel degelijk toe deden. Dat konden deze onderzoekers elkaar en het publiek in een aantal symposia vol overtuiging duidelijk maken. Het gaf ten langen leste de indruk van een rondreizend circus. De onderzoekers, in paren: Brown en Campione met hun Communities of practice, Scardamalia en Bereiter met CSILE, en Goldman en Bransford met Jasper. Het moet gezegd: de innovaties zijn theoretisch en praktisch indrukwekkend, maar waarom moet dat in een flink aantal symposia telkenmale herhaald worden zonder dat er veel nieuwe feiten aan worden toegevoegd?

Een interessante ontwikkeling die ook op de AERA zichtbaar werd in de paper-presentaties, is de verbinding van communicatie en interactie. De nieuwe technologieën, in het bijzonder een nieuwe communicatietechnologie als Internet, maakt het mogelijk geheel nieuwe en tot voor kort nauwelijks denkbare innovaties op het terrein van instructie en leren door te voeren. De lerende wordt steeds minder afhankelijk van de produkten van de instructieontwerper maar wordt meer in staat gesteld in rijke omgevingen, met behulp van de communicatietechnologie, rond te dolen en daardoor zijn eigen kennispakket samen te stellen. Symposia en presentaties met in de titel knowledge building of knowledge construction, met of zonder collaborative, maakten hier steeds meer melding van. Zoals bijvoorbeeld het door OISE gedomineerde symposium 'Collaborative knowledge building on a computer network', en de interessante paper sessie 'Technology enhanced learning environments' met bijdragen van Woodruff 'The effects of computer-mediated communication on collaborative discourse in knowledge building communities' en van Jonassen over case-based learning environments. Ook de bijdragen van Levin (Illinois) in een aantal symposia, waarin hij soms optrad als discussiant, wezen duidelijk op mogelijkheden in die richting.

Maar ondanks de indruk die presentaties en

demonstraties maken, er blijft de vraag naar de effectstudies (wat is nu het effect van die mooie omgevingen?) en naar de ontwerprichtlijnen voor die fraai vormgegeven omgevingen. Sommige onderzoekers trachten daarop antwoorden te vinden door communicatienetwerken te observeren en de interacties tussen docenten, domein-experts en studenten systematisch te inventariseren en zodoende een beeld te schetsen van de werking van niet alleen de technologische omgeving, maar met name van de inhoudelijke impact ervan. Een voorbeeld van het laatste werd geleverd in het symposium 'Analyses of computer-mediated communications within on-line teaching/learning' door Harris en Jones in hun bijdrage 'A study of on-line communication among subject matter experts, teachers, and students: Message flow and functions'.

### Professionele ontwikkeling van leraren

(M. G. Roes, Universiteit Twente)

Het realiseren van onderwijsvernieuwingen is en blijft een complexe aangelegenheid. Dit bleek weer eens uit de papersessies waarin verslag werd gedaan van Amerikaans onderzoek naar de effectiviteit van programma's voor de professionele ontwikkeling van leraren.

Leraren in de VS zien zich voor ingrijpende onderwijsvernieuwingen geplaatst: nieuwe benaderingen van de exacte vakken, nieuwe instructiemodellen als 'complex instruction' en coöperatief leren. Deze innovaties hebben met elkaar gemeen dat ze uitgaan van een constructivistische opvatting van het leerproces.

De ondersteuning van docenten bij de invoering van deze vernieuwingen wordt intensief aangepakt. Langer lopende trajecten bestaande uit cursussen, follow-up workshops en begeleiding op de werkplek in de vorm van coaching, zijn geen uitzondering.

In veel van de gepresenteerde studies waren de 'knowledge and belief systems' van leraren als uitgangspunt genomen voor het inrichten van de professionele ondersteuningsarrangementen. Via veranderingen in kennis en opvattingen van docenten worden veranderingen in de lespraktijk nagestreefd. In de uitwerking van deze programma's betekent dit dat veel belang wordt gehecht aan mogelijkheden voor continue dialoog met collega's en programma-

staf én voor reflectie op de eigen lespraktijk.

Een voorbeeld hiervan was het paper van Swafford, Jones en Thornton met als titel 'Changing middle grades teachers' classroom practices'. Dit paper werd gepresenteerd tijdens een sessie waarin de vernieuwing van het wiskunde-onderwijs centraal stond. De onderzoekers hadden de effecten bestudeerd van een driejarig ondersteuningsprogramma dat gericht was op de implementatie van nieuwe standaarden voor het wiskunde-onderwijs. Het nascholingsprogramma bestond jaarlijks uit de volgende componenten: een zomercursus wiskunde van vier weken; een ééndaags research-seminar over het leren van wiskunde door leerlingen én zes middagseminars, waarbij gewenste lespraktijken werden geanalyseerd, onder meer op basis van inbreng van de deelnemers. Ook was tijdens deze middagen gelegenheid voor formele en informele uitwisseling tussen de deelnemers. Als huiswerk moest iedere deelnemer jaarlijks twee lessen op video opnemen en analyseren en een reflectie-logboek bijhouden. Aan het eind van elk 'cursus'jaar moesten ze een samenvatting van deze logboeken maken en bespreken met de programmaleiding. Tenslotte was er op de scholen van de deelnemers tijd voor onderlinge samenwerking uitgetrokken. Om de effecten van dit programma te onderzoeken, hadden de onderzoekers de 48 deelnemers aan het begin en aan het eind van het traject een vragenlijst voorgelegd over hun opvattingen, lespraktijk en leerstofaanbod en een wiskundetoets afgenomen. Zes deelnemers waren twee keer per jaar een week lang geobserveerd en geïnterviewd. Bij de leerlingen van alle deelnemers (en bij leerlingen in een controlegroep) was een wiskundetoets en een attitudevragenlijst afgenomen. De onderzoekers constateerden op alle onderzochte aspecten significante veranderingen in de gewenste richting. Uitzondering hierop vormde de attitude van de leerlingen van de deelnemers, die niet verschilde van de leerlingen in de controlegroep. De onderzoekers concluderen dat vernieuwingen in het wiskunde-onderwijs mogelijk zijn, maar dat het veel tijd kost: pas na het eerste jaar werden veranderingen zichtbaar.

Ondanks de hoopgevende resultaten van deze studie, bleek het effect van de diverse programma's voor de professionele ontwikkeling

van leraren nogal wisselend te zijn. Bestaande opvattingen van leraren over onderwijzen blijken vaak zeer hardnekkig te zijn. Deze opvattingen passen vaak meer bij een transmissie-model van het onderwijsleerproces dan bij een constructivistische benadering.

In de papers lag overwegend de nadruk op de kennis en opvattingen van docenten als aangrijpingspunt voor veranderingen. Het in de praktijk brengen van nieuwe modellen van onderwijzen, vraagt echter ook om de verwerving van nieuwe vaardigheden. Op de conferentie was nauwelijks aandacht voor gerichte vaardigheidstraining als component van een programma voor professionele ondersteuning. Evenmin werd aandacht geschonken aan de rol die curriculummaterialen kunnen spelen in een dergelijk traject. Wellicht dat een meer integrale benadering de effectiviteit en efficiency van dergelijke ondersteuningsprogramma's kan vergroten.

**Schooleffectiviteitsonderzoek** (H. de Vos, Universiteit Twente)

Alhoewel er op de AERA in 1995 zo'n 1400 papersessies waren, werden er maar drie gekenmerkt als schooleffectiviteit-sessies. Eén ging over de relatie van schooleffectiviteitsonderzoek en schoolverbetering, één over onderwijs indicatoren en de laatste ging over voortgang van schooleffectiviteitsonderzoek. De sessies kenmerkten zich door hun internationale karakter: de presentaties werden niet alle door Amerikanen verzorgd.

Aangezien mijn interesse voornamelijk methodologisch was getint, zal ik met name de presentaties eruit lichten, die aan mijn interesse tegemoet komen.

Een geheel Nederlandse sessie was verroosterd op donderdagmorgen: The evaluation of the Dutch Educational Priority Program. Het publiek bestond grotendeels uit Nederlanders. Het onderwerp was, voor zover dat nog niet duidelijk was uit de titel, de evaluatie van zes jaren OVB. De participanten aan deze sessie waren Van Tilborg en Walraven van SVO, Mulder van het ITS en Van der Werf van het GION. Alhoewel blijkt dat de prestaties op de OVB-scholen verbeterd zijn, is het moeilijk te zeggen of dit ook door het OVB is bewerk-

stelligd. De oorzaak zou ook kunnen liggen in het feit dat de leerlingen zes jaar langer in Nederland wonen. Welk deel van de verbetering wordt verklaard door factoren binnen de school, factoren buiten de school en OVB is nog onduidelijk. De discussiant Slavin, van de John Hopkins University, was erg enthousiast over de OVB, mede omdat het Nederlandse onderwijsbeleid ten aanzien van specifieke doelgroepen soms lijnrecht tegenover de uitkomsten van het gedecentraliseerde onderwijs-financieringsstelsel van de Amerikaanse staten staat: (relatief) arme staten besteden minder geld aan de compensatie van achterstanden van gedepriveerde groepen leerlingen.

Ook donderdagochtend (om 8.15 uur!) was er een sessie over internationaal gebruik van onderwijs-indicatoren: International uses of quality indicators. De toevoeging "internationaal" was zeker op zijn plaats; bijdragen aan deze sessie kwamen uit Canada, Engeland, Frankrijk, Nederland en de Verenigde Staten. Onderwijsindicatoren worden gebruikt om de kwaliteit van onderwijs en scholen te beschrijven. De papers belichtten zienswijzen omtrent ontwikkeling en optimaal gebruik van deze indicatoren. In Canada, aldus McEwen, bepalen de provincies het onderwijs en dus ook het gebruik van indicatoren. Zij stelde een aantal verbeteringen voor op provinciaal en nationaal niveau, zoals het gebruik van indicatoren voor vergelijkingen tussen provincies (een soort benchmarking?) en voor het optimaliseren van schoolbegeleiding. Uit Engeland kwam een bijdrage van Fitz-Gibbon. Zij vroeg zich af of het systeem van indicatoren dat wordt gebruikt in de UK, wel een beeld geeft van de kwaliteit van onderwijs. Het indicatorensysteem wordt met name bepaald door examenresultaten en resultaten van onderwijsinspectie. De kwaliteit van het indicatorensysteem wordt momenteel verhoogd door de toevoeging van 'value-added' maten. Creemers (Nederland) ging in op de bijdrage van schooleffectiviteitsonderzoek aan de ontwikkeling en het gebruik van onderwijsindicatoren. Hij stelde dat onderzoek, zoals IEA (International Association for the Evaluation of Educational Achievement) en ISERP (The International School Effectiveness Research Program), de ontwikkeling ondersteunen. Het is echter wel zaak, aldus

Creemers, dat indicatoren internationaal vergelijkbaar zijn, dus onafhankelijk van het onderwijsstelsel. Bonnet uit Frankrijk hield een pleidooi voor het monitoren van het onderwijs met behulp van indicatoren. Hij waarschuwde dat moet worden opgepast dat indicatoren geen politieke instrumenten worden.

Op de vrijdagmiddag was er een sessie over internationale voortgang in schooleffectiviteitsonderzoek georganiseerd door ISERP, onder voorzitterschap van Reynolds. Tymms (UK) ging in zijn presentatie 'Monitoring effectiveness: Examples of the evaluation of complex systems' ver door te beweren dat schooleffectiviteitsonderzoek er nooit in zal slagen zodanige kennis te vergaren dat scholen in het algemeen kunnen worden verbeterd, omdat wat er in scholen en klassen gebeurt onvoorspelbaar, complex en chaotisch is. Hij hield een pleidooi voor gebruik van onderwijsindicatoren, die meer inzicht kunnen geven in wat er nu eigenlijk gebeurt. In de paperpresentatie over 'Two year quantitative findings from ISERP' (o.a. Creemers) werd ingegaan op het gegeven dat de leerlingvariabelen (SES, IQ, ethniciteit) het grootste deel van de wiskundeprestaties van leerlingen verklaren in internationaal vergelijkend onderzoek (net als in het meeste onderzoek). Alhoewel in de meeste landen ook scholenfactoren van invloed zijn op de leerlingprestaties is dat in Nederland (en ook in Taiwan) niet het geval.

### **Measurement and Research Methodology**

(H. J. Vos, Universiteit Twente)

Het aandachtsgebied 'Measurement and Research Methodology' was ruim vertegenwoordigd in de vorm van thematische papersessies, (invited) symposia waarvan 36 in samenwerking met de National Council on Measurement in Education (NCME), posterpresentaties en pre-conference workshops.

Een van de belangrijkste trends betrof het adaptief toetsen van vaardigheden. Computerized Adaptive Testing (CAT) wordt steeds meer gebruikt in de praktijk en lijkt de toetsvorm van de toekomst te worden. Het belangrijkste onderdeel van deze toetsvorm in vergelijking met conventionele potlood-en-papier (P&P) toetsen is dat volstaan kan worden met kortere toet-

sen bij gelijkblijvende of zelfs hogere nauwkeurigheid van de geschatte vaardigheid. Er is echter nog veel psychometrisch onderzoek te doen op dit gebied, zowel voor 'fine-tuning' van de theorie als voor het oplossen van praktische problemen bij de implementatie. Aan beide aspecten van CAT werd in verschillende papersessies en symposia ruimschoots aandacht besteed.

Veerkamp en Chang stelden beiden nieuwe criteria voor om items te selecteren in een adaptieve toetsprocedure gebruikmakend van item response theorie (IRT) modellen. Het belangrijkste bezwaar bij de meeste huidige criteria is dat ze geen rekening houden met de onzekerheid van de geschatte vaardigheid. Terwijl Veerkamp een gewogen gemiddelde van de gebruikelijke informatiefunctie voorstelde, stelde Chang daarentegen een 'global information approach' voor om de gesignaleerde bezwaren bij de huidige item selectiecriteria te omzeilen.

In de presentatie van Rocklin, Vispoel en Wang over de mogelijkheid tot manipulatie van het CAT-algoritme als item review mogelijk is door de 'test-taker' (Self-Adapted Testing) kwam naar voren dat het erg riskant is voor de test-taker om op deze manier te trachten onterecht een hogere score te behalen. In theorie kan een test-taker door er voor te zorgen alleen erg gemakkelijke items voorgelegd te krijgen en die vervolgens allemaal goed te beantwoorden kunstmatig een erg hoge score verkrijgen. Als echter enkele vragen fout beantwoord worden zal zijn of haar score te laag uitvallen.

Lunz en DeVille deden verslag van een studie waarin toetsen geconstrueerd via de computer met een adaptieve selectieprocedure werden vergeleken met toetsen die handmatig werden geconstrueerd en op papier werden afgenomen. Hierbij werd met name gekeken naar de validiteit van de geselecteerde items, de volgorde van de items, de statistische eigenschappen van de toetsen en in hoeverre de geselecteerde items voldeden aan de toetsspecificaties. Uit deze studie bleek dat beide methoden voor toetsconstructie vergelijkbaar zijn voor wat betreft validiteit en statistische eigenschappen en dat beide methodes voldeden aan de toetsspecificaties.

Davey en Parshall keken naar exposure

rates van items (percentage test-takers dat een bepaald item voorgelegd krijgt) in adaptieve toetsen. Er werd gekeken hoe hoog deze exposure rates zijn als er niet voor gecontroleerd wordt. Verschillende methodes van item exposure control werden voorgesteld en met elkaar vergeleken.

Schnipke onderzocht hoe itemresponstijden gebruikt kunnen worden om te achterhalen hoe een persoon een item heeft trachten te beantwoorden, door snel te gokken of door serieus naar een oplossing te zoeken. Op deze manier is het mogelijk gokgedrag uit de test te filteren en daardoor de vaardigheid beter te schatten.

Tenslotte presenteerden Way, Smith en Lewis hun paper over het gebruik van Bayesiaanse besliskunde bij 'computerized mastery testing' (CMT). Testlets (d.w.z. een klein aantal samenhangende items) werden sequentieel afgenomen in hun onderzoek. Na iedere aangeboden testlet werd een slaag, zak, of continueer beslissing genomen. Van duidelijke masters of nonmasters kon na slechts enkele testlets reeds worden besloten of ze waren geslaagd of gezakt. Test-takers wier prestaties echter in de buurt van de grensscore lagen moesten de meeste testlets beantwoorden.

Zoals gebruikelijk de afgelopen jaren gingen veel papers over IRT-modellen en statistische problemen hierbij die nog om een oplossing vragen. Het symposium 'Statistical issues in item response theory models', georganiseerd door Berger, was hiervan een goed voorbeeld.

Verhelst rapporteerde over het door het CITO ontwikkelde gegeneraliseerde One-Parameter Logistic Model (OPLM). Hij liet zien dat bij het OPLM-model enerzijds de elegante statistische eigenschappen van het Rasch-model bewaard blijven terwijl anderzijds de grotere flexibiliteit van het 2-parameter logistische model (Birnbaum model) optimaal kan worden benut.

Glas toonde aan dat in een klasse van IRT-modellen, welke het Rasch model, het lineair logistische test model (LLTM), het (gegeneraliseerde) partial credit model, het OPLM-model en modellen met multidimensionele persoonsparameters bevatte, zowel Conditional Maximum Likelihood (CML) als Marginal Maximum Likelihood (MML) schattingsprocedures kunnen worden toegepast. Tevens werd een taxonomie van gegeneraliseerde

Pearson tests beschreven welke power hebben tegen specifieke modelschendingen in deze klasse van modellen. De benadering werd geïllustreerd aan de hand van twee voorbeelden om item bias (DIF) te detecteren.

Sijtsma deed verslag over nonparametrische IRT-modellen voor dichotome en polytome data, welke een ordinale ordening van personen en/of items veronderstellen. Hij vergeleek nonparametrische en parametrische IRT-modellen met elkaar en ging hierbij met name in op de vraag in welke situaties het gebruik van nonparametrische IRT-modellen de voorkeur genieten (o.a. wanneer de assumpties bij logistisch parametrische modellen te restrictief zijn).

Van der Linden en Luecht presenteerden een optimalisatiemodel (0-1 lineair programmeringsmodel) dat toetsconstructeurs in staat stelt om de vorm van een vooraf gekozen scoreverdeling op een toets voor een populatie met een bekende vaardigheidsverdeling zo goed mogelijk te controleren.

Tenslotte toonde Kelderman in een meer fundamenteel paper aan dat het Rasch model kan worden afgeleid uit een gegeneraliseerd lineair model voor de uitwisselbaarheid van metingen.

Een laatste opvallende trend was de hernieuwde interesse voor het probleem van het vaststellen van grensscores. Na een hausse in de jaren zeventig en tachtig als gevolg van mastery learning was tijdens de afgelopen AERA/NCME dit thema weer prominent aanwezig op de onderzoeksagenda. Het symposium 'Applications of decision theory to practical testing problems', georganiseerd door Sawyer met Van der Linden als moderator en vijf bijdragen omvattend, ging met name in op de rol die Bayesiaanse besliskunde kan spelen bij het vaststellen van grensscores.

Ben-Shakhar, Kiderman en Beller deden verslag van een vergelijkende studie naar twee verschillende beslissingstheoretische toelatingsprocedures voor 'Liberal Arts' op universitair niveau. De ene toelatingsprocedure hield wel rekening met eerder behaalde SAT (Scholastic Aptitude Test)-scores, terwijl de andere procedure hier geen rekening mee hield. De resultaten toonden aan dat er geen winst te behalen valt bij toevoeging van de SAT-scores als het de bedoeling is om studenten te recruter



die voldoen aan minimale kwalificaties. Wanneer het echter de bedoeling is om potentieel excellente studenten te selecteren bleek toevoeging van de additionele informatie van de SAT-scores aanzienlijke winst op te leveren.

Van der Linden presenteerde een beslissingstheoretisch model voor het veel in de praktijk voorkomende probleem van het optimaal toewijzen van studenten aan twee verschillende hiërarchisch geordende cursussen ('course placement'). Op basis van een aptitude test moest worden besloten of studenten ofwel werden toegewezen aan de eerste cursus ofwel rechtstreeks werden toegelaten tot de tweede cursus. De eerste cursus werd hierbij afgesloten met een beheersingstoets welke bij een voldoende resultaat toegang gaf tot de tweede cursus.

Vos rapporteerde over het simultaan optimaliseren van combinaties van onderwijskundige beslissingen. De belangrijkste toepassingen hiervan liggen op het terrein van het formuleren van optimale regels voor voortgangsbepalingen in geïndividualiseerde instructiesystemen. In een empirisch geïndividualiseerd instructiesysteem, bestaande uit een gecombineerde selectie-plaatsings-beheersingsbeslissing, bleek dat bij de simultane benadering een aanzienlijke winst in verwachte utiliteit te behalen viel vergeleken bij het afzonderlijk optimaliseren van iedere beslissereg. Tevens werd aangetoond dat onder tamelijk milde voorwaarden optimale beslisseregels noodzakelijkerwijs compensatorisch van aard waren.

Terwijl de drie bovenstaande bijdragen alle uitgingen van bekend veronderstelde utiliteitsfuncties hielden de laatste twee bijdragen zich bezig met het empirisch specificeren van utiliteitsfuncties. Als zodanig vormden ze een goede aanvulling op de eerdere papers, omdat het voor een beslissingstheoretische benadering essentieel is te kunnen beschikken over een realistische utiliteitsstructuur.

Mellenbergh, Van der Gaag en Van den Brink lieten zien dat empirische utiliteitsfuncties voor cultuureerlijke selectie bevredigend konden worden benaderd door lineaire- en normaal-ogief utiliteitsfuncties.

Tenslotte toonde Sawyer aan dat het empirisch specificeren van utiliteitsfuncties voor 'course placement' tot aanzienlijke verschillen

in uitkomsten kan leiden afhankelijk van het feit of een deterministische of stochastische (Von Neumann-Morgenstern) techniek werd gebruikt.

Aan deze kroniek werkten mee: *M. Mulder, J.M. Pieters, M.G. Roes, J.N. Streumer, H.J. Vos en H. de Vos.*

De eindredactie werd verzorgd door *J.N. Streumer.*