

'Assessing teacher effectiveness'

Vierde konferentie van de 'International Association for Educational Assessment' (IAEE), Baden, Oostenrijk, 22-25 mei 1978.

De IAEE is een vereniging waarvan landelijke en regionale instituten voor test- en toetsontwikkeling de kern zijn. Naast instituten kunnen ook individuele personen lid worden van de IAEE. De vereniging is vooral een internationale studiegroep en het is dan ook voor de hand liggend dat zij aan de organisatie van de jaarlijkse konferenties ruime aandacht schenkt. Op die konferenties worden ervaringen uitgewisseld, gezamenlijke projecten gepland en geëvalueerd, maar vooral mondiaal in de aandacht staande topics op het werkveld van de instituten besproken. Zo werd in 1976 de toelating tot het hoger onderwijs tot konferentiethema gekozen en waren in 1977 de eindexamens onderwerp van bespreking. De conferentie van 1978, met deelnemers uit 22 landen, die gehouden werd in Baden (bij Wenen), had de "vaststelling van effectiviteit van leerkrachten" tot onderwerp. Hoewel voor de meeste instituten dit topic niet het belangrijkste aandachtsveld is, hebben ze er meestal toch direct of indirect iets mee te maken. Naar onze indruk is in ontwikkelingslanden de bemoeienis van de centrale overheid met de opleiding van leerkrachten via de instituten voor test- en toetsontwikkeling ('examination boards') groter dan in het westen. Men kan zich zelfs afvragen of die invloed wel altijd wordt aangewend met onderwijskundige oogmerken. Hoe dan ook, politieke- en arbeidsmarktoverwegingen speelden een rol op de achtergrond van een conferentie over methodieken en middelen voor de meting van 'teacher effectiveness'.

Het centrale thema werd op de conferentie aan de orde gesteld door middel van lezingen, panel- en groepsdiscussies over een drietal specifieke onderwerpen. Opvallend was de grote gevarieerdheid waarmee het onderwerp "leerkracht-effectiviteit" door de diverse inleiders werd ingevuld. Hierbij speelt waarschijnlijk ook de diversiteit in de functie van het onderwijs in de deelnemende landen een rol. Het is duidelijk dat een ontwikkelingsland bijvoorbeeld minder geneigd zal zijn de schaarse middelen te investeren in fijnkorrelig onderzoek naar het precieze effect van bepaalde leerkrachtkenmerken, maar meer geïnteresseerd zal zijn in het ontwikkelen van een selectie-instrument voor het selecteren van potentieel "effectieve" docenten.

Het is uiteraard ondoenlijk hier alle gepresenteerde papers de revue te laten passeren. Hetgeen hieronder besproken wordt is noodzakelijkerwijs een (vnl. door onze onderzoeksinteresse bepaalde) selectie.

De conferentie werd ingeleid door Yoloye (Nigeria). In zijn lezing zette hij uiteen dat, voorafgaande aan de beantwoording van de effectiviteitsvraag, een drietal sub-vragen gesteld moet worden: 'What is the teacher expected to do; What is he trained to do; What is he equipped to do?' Gekonstateerde diskrepancies in de antwoorden tussen deze sub-vragen kunnen (in elk geval in veel ontwikkelingslanden, maar zeker niet alleen daar!) verhelderend werken bij het zoeken naar oorzaken van effectiviteitsproblemen. Aan de hand van het voorbeeld van de onderwijzer in de rurale gemeenschap in een ontwikkelingsland lichtte hij vervolgens toe op welke wijze de verwachtingen die de gemeenschap van een leerkracht heeft in betrekkelijk korte tijd drastisch kunnen veranderen. Als hulpmiddel bij de beoordeling van het leerkrachtgedrag pleitte hij voor het opstellen van profielen van "ideaal" observeerbaar leerkrachtgedrag, vnl. geïnspireerd door (modifikaties van) het interactie-analyse-schema, waarbij overigens niet werd aangegeven op welke wijze deze profielen tot stand moesten komen. Vervolgens werden enkele resultaten gepresenteerd uit onderzoek dat aan de universiteit van Ibadan is uitgewerkt. Als voornaamste oorzaak van de geringe toepasbaarheid van de onderzoeksresultaten op dit gebied werd het ontbreken van een 'theory of instruction' genoemd.

De tweede inleidende bijdrage, waarin getracht werd de effectiviteitsproblematiek in algemene zin te belichten, was afkomstig uit België (Schlusmans). De drie deelvragen waarin zij het effectiviteitsprobleem uiteenlegde zijn dezelfde vragen die ook in de "historie" van het effectiviteitsonderzoek voortdurend terugkeren: (1) Van welke criteria moet men uitgaan bij het beoordelen van de effectiviteit van de leerkracht. (2) Welke karakteristieken zijn kenmerkend voor de effectieve leerkracht. (3) Hoe meten we deze karakteristieken. Interessant in deze bijdrage was de rapportage van een literatuurstudie waarin bij een groot aantal publikaties was nagegaan welke leerkrachtkenmerken worden genoemd als kenmer-

kend voor de "effektieve" leerkracht, en waarbij de publikaties in een aantal categorieën werden ingedeeld. Als de meest relevant geachte karakteristieken kwamen naar voren: 'interest in the child, willingness to understand others, kindness, mastery of subject matter, patience, intelligence, respect for others'. Geen enkele karakteristiek werd overigens door meer dan 50% van de studies relevant genoemd.

De overige papers van het kongres kunnen, zij het met ruime overlap, in enkele categorieën worden ingedeeld. Zo had één aantal voordrachten betrekking op het onderzoek naar *leerkrachtenkarakteristieken*, zowel op het terrein van persoonlijkheidsvariabelen als op het terrein van gedragsvariabelen in de onderwijsleersituatie. In een aantal gevallen werden deze karakteristieken gerelateerd aan effectiviteitscriteria. Enkele papers hadden betrekking op de problematiek van het *selecteren van leerkrachten* op zodanige wijze dat daardoor een maximaal rendement voor het onderwijssysteem ontstaat. Ten slotte hield een aantal papers zich bezig met de vraag of en op welke wijze men de (a.s.) *leerkracht assessmentstechnieken* moet bijbrengen.

In de papers waarin werd ingegaan op de persoonlijkheids- en stijl-karakteristieken van leerkrachten werd zeker geen volledig overzicht gegeven van de ontwikkelingen op dit terrein. Bertini (Italië) ging in op de relatie tussen de stijlfactor veldafhankelijkheid en onderwijsgedrag. Hij volgde daarbij Witkin en zijn onderzoeksgroep. De belangrijkste onderzoeker uit het gezelschap was ongetwijfeld Frederick J. McDonald, die als projectleider optrad van de 'Beginning Teacher Evaluation Study' in Californië (zie voor het samenvattende eindverslag van de tweede fase van dit onderzoeksprogramma o.m. ERIC-document ED 142 592). Hij presenteerde enige resultaten uit het onderzoek waarin bij basisschoolleerkrachten werd aangehouden dat bepaalde patronen van onderwijsgedrag (geobserveerd met het observatiesysteem APPLE) significant samenhangen met leerlingenprestaties. Een andere conclusie was dat bepaalde onderwijspatronen bij bepaalde vakken wel, maar bij andere vakken niet effectief zijn.

M.b.t. de selectie-problematiek was met name de bijdrage van Itai Zak (Israël) interessant ('Self-selection and the image of the teaching profession'). Zak gaf aan dat er nauwelijks selectie plaatsvindt die gebaseerd is op een duidelijk beeld van de kwaliteiten die nodig zijn om de rol van leraar te vervullen. Selectie voor het leraarsberoep lijkt meer een soort natuurlijke selectie volgens een bepaald beroepsbeeld te zijn, dan een weloverwogen en gerichte selectie. In zijn paper deed hij verslag van een studie waarin men ervan is uitgegaan dat diegenen zichzelf zullen selecteren voor een lerarenopleiding, wiens persoonlijkheidskenmerken overeenkomen met eigenschappen die de maatschappij van een leraar verwacht. Deze beweringen worden gestaafd door middel van vergelijkingen van profielen van verschillende groepen toekomstige leraren met profielen van leeftijdgenoten als normatieve groep. De groepen toekomstige leraren blijken duidelijk overeenkomstige eigenschappen te bezitten. Uit andere onderzoeken blijkt dat deze eigenschappen vaak bij leraren voorkomen ('submissive, conforming, prudent, sober,

sensitive, conservative').

In een aantal papers werd door enkele Engelse instituten die zich met o.m. inservice-training bezighouden (City and Guilds of London Institute) gerapporteerd over recente ontwikkelingen op het gebied van het trainen van leerkrachten in beoordelingstechnieken.

De onderwerpen van de discussiegroepen waren zodanig gekozen dat daarin veel van de problematiek uit de papers in meer toegespitste vorm besproken kon worden. In de discussie over het, met name voor de ontwikkelingslanden relevante, thema 'The teacher as a change agent outside the school' kwam de belangrijke rol aan de orde die de leerkracht kan spelen bij het bevorderen van maatschappelijke veranderingen. Tot nu toe is weinig onderzoek gedaan naar dit aspect van 'teacher effectiveness'. Wel lijkt het duidelijk dat de rol van de docent in deze verschillend is in westerse- of ontwikkelingslanden, en binnen een land in stedelijke- of plattelandsgebieden. De discussie richtte zich vooral op de docent in een rurale gemeenschap. Onderwerpen die aan bod kwamen waren: de aard van de gewenste veranderingen die door activiteiten van de docent buiten de school zouden moeten plaatsvinden (bijv. innovatie, coöperatie), de verschillende manieren waarop een docent die veranderingen teweeg kan brengen (o.a. als raadgever, als organisator), de voorwaarden waaraan binnen de gemeenschap moet zijn voldaan zodat de docent de mogelijkheid heeft om op te treden als change agent (administratieve infra-structuur etc.), de mogelijkheden en moeilijkheden die juist gelden voor iemand die in de functie van docent als change agent optreedt.

Een tweede discussiegroep had als onderwerp: 'Implications of assessment for teacher training'. Deze titel kan op twee manieren worden geïnterpreteerd: (1) hoe, door wie en op welke criteria moet de toekomstige docent beoordeeld worden, en een interpretatie die door de groep is overgenomen maar die minder aansloot bij het algemene thema van de conferentie, nl. (2) hoe bereidt men de toekomstige docent tijdens de trainingsperiode voor op de taak om zijn of haar eigen studenten te beoordelen. Aan de hand van een overzicht van de bestaande praktijken in de verschillende landen om docenten te leren beoordelen werd duidelijk, dat relatief weinig aandacht wordt besteed aan zaken als evaluatie en beoordeling. Als het onderwerp al ter sprake komt vormt het een onderdeel van bijvoorbeeld een meer algemene cursus 'leerlingbegeleiding'. Uit de discussie kwamen o.m. de volgende onderwerpen naar voren: Men hechtte belang aan een evenwicht tussen pre- en in-service training op dit gebied en aan de manier waarop de school zelf bij in-service training een rol kan spelen. De rol van organisaties als ETS en CITO werd meer gezien in het leveren van bijdragen aan toenemende competentie van docenten in beoordeling en evaluatie dan aan het leveren van steeds betrouwbare tests. De docent moet in die mate op de hoogte zijn van beoordelingstechnieken en -materialen dat hij een verantwoorde keuze kan maken uit wat hem door professionele instituten wordt aangeboden. Men gaf de noodzaak aan om op trainingsinstituten in-service cursussen voor docenten te ontwikkelen waar zij kennis kunnen nemen van de

principes en processen van beoordeling, echter zonder daarbij te proberen van de docenten test-konstruktoren te maken.

De derde discussiegroep hield zich bezig met aspecten aangaande het thema 'Identifying and rewarding excellence in teaching'. De groep heeft getracht probleemvelden te identificeren die de aandacht verdienen en onderwerpen aan te dragen voor verder onderzoek. Men achtte het allereerst noodzakelijk te definiëren wat wordt bedoeld met 'excellence' en criteria aan te geven om deze te kunnen meten en evalueren; men kan excellence alleen herkennen na evaluatie en dan eerder na evaluatie van het lesgeven zelf (proces) dan in termen van de uitkomsten. Als belangrijk onderdeel van de evaluatie werd zelf-evaluatie aangegeven, waarbij de nadruk zou moeten liggen op het meten van progressie in de vaardigheden van de docent. Hiervoor zouden zelf-diagnostische instrumenten moeten worden ontwikkeld. Andere discussiepunten waren;

- beloning van goede docenten binnen het onderwijsveld i.p.v. promotie naar functies buiten het directe onderwijsveld.
- aandacht voor het formuleren van algemene doelstellingen waaraan de docent zijn gedrag kan meten.
- ontwikkeling van criteria voor de evaluatie van docenten op nationaal niveau.
- beoordeling moet geschieden door onderwijsgevenden zelf en niet door outsiders. Deze laatste kunnen wel technische middelen aandragen.

Enkele van de aspecten die in de discussiegroepen aan de orde gesteld zijn, zullen door verschillende leden van de IAEA nader worden bestudeerd en uitgewerkt.

Nog enkele opmerkingen over vanuit de IAEA geïnitieerde internationale projecten.

Het eerste project heeft als titel 'A cross-cultural factor-referenced study of human abilities and personality'. Dit project is ontstaan vanuit de behoefte aan een internationaal vergelijkbaar instrument ter bepaling van de cognitieve vermogens van leerlingen. Het is de bedoeling de m.b.v. dit instrument gemeten factoren uiteindelijk te relateren aan criteria als succes in het hoger onderwijs. Momenteel verkeert het project nog in een voorbereidend stadium waarin men probeert de tests zodanig te construeren dat in uiteenlopende kulturen toch dezelfde constructs worden gemeten.

Een tweede project, 'The multinational aptitude test' is aan het eerste project verwant, maar is veel sterker gericht op de concrete inhouden van het voortgezet onderwijs. Het doel van dit project is het ontwikkelen van een instrument (in verschillende talen) waarmee de doorstroming van studenten van het ene naar het andere land zou kunnen worden vergemakkelijkt.

Als derde kan het 'Instructional item banking project' genoemd worden, waarin internationaal wordt samengewerkt bij de constructie van toetsitems.

Als samenvattende indruk van de conferentie kan gesteld worden dat een zo grote gevarieerdheid aan benaderingen van de problematiek aanwezig was dat het zicht op het geheel soms verloren dreigde te gaan. Als oorzaken hiervoor kunnen genoemd worden: het conferentiethema ('teacher effectiveness' is, niet in de laatste plaats door de spraakverwarring die in het verleden rond dit begrip heeft geheerst, een thema dat bijzonder moeilijk af te bakenen is), de kulturele achtergrond van de deelnemers (westers, niet-westers), de werk-achtergrond van de deelnemers (beleidsmakers en -adviseurs, onderzoekers, toetsontwikkelaars). Een pluspunt bij een dergelijke stand van zaken is uiteraard het feit dat men voortdurend gedwongen wordt eigen standpunten en wijzen van benaderen te relativiseren. De vraag in hoeverre de conferentie zal leiden tot wezenlijke veranderingen in de werkwijze van de deelnemende instituten is moeilijk te beantwoorden. Er werd weliswaar een grote variatie aan technieken gepresenteerd en besproken (overigens zonder veel aandacht voor de geavanceerde) maar dit leidde in het algemeen meer tot kennismaken dan tot beoordelen. Dat de IAEA het heeft aangedurfd de leerkracht-effectiviteitsproblematiek als conferentiethema te kiezen verdient overigens alle lof.

Een conferentieverlag zal worden gepubliceerd door Pergamon Press. Het secretariaat van de IAEA is: IAEA, Box 2829, Princeton, New Jersey 08540, U.S.A.

De volgende IAEA-conferentie zal in 1979 in Princeton (USA) worden gehouden; het thema zal dan luiden: The role of assessment in evaluation of educational programs and systems.

H. Dekkers
A. Verhoeven
N. Verloop