

[januari 1971, onbekend voor welke gelegenheid dit is gemaakt]

CENTRUM VOOR ONDERZOEK VAN HET WETENSCHAPPELIJK ONDERWIJS

BEOORDELEN IN HET ONDERWIJS -- DISCUSSIONOTA

In verband met de weinige tijd beschikbaar, zal op de technische aspecten van de diverse beoordelingsprocedures niet ingegaan kunnen worden. Het feit dat er technische problemen zijn in verband met bijna iedere vorm van beoordelen, zal in de discussie zeker ter sprake komen. Voor de meer praktische en technische problemen zij verwezen naar het handboek voor de toetsconstructie van De Groot, Van Naerssen e.a. (1970), of van Robert L. Ebel: *Measuring Educational Achievement*, (1965, Prentice Hall).

Beoordelen, of dat nu in onderwijs-situaties of elders gebeurt, blijkt bij enige nadere analyse een allerm minst vanzelfsprekende aangelegenheid te zijn, noch wat het doel en gebruik van beoordelingen betreft, noch waar het gaat om het probleem op welke wijze de beoordelingen zo "eerlijk" mogelijk gemaakt kunnen worden. Dit geldt evenzeer voor het gebruik van meerkeuze studietoetsen. Het zou te ver voeren een systematische beschouwing over de beoordelingsproblematiek te geven, daarom het volgende voorstel:

Wij presenteren hier een reeks stellingen m.b.t. gebruik en rechtvaardiging van het gebruik van verschillende beoordelingsmethoden. Voor zover nodig zullen deze stellingen van een korte toelichting voorzien worden. Op basis hiervan kan een discussie in kleine groepen plaatsvinden, waarbij denkbaar is dat ofwel de groep zelf tot een stellingname komt op bepaalde controversiële stellingen die gepresenteerd zijn, ofwel er een discussie plaatsvindt tussen de groep en ondergetekende(n).

Doelstelling hierbij is onzerzijds om mensen die al een groot deel van hun leven als beoordeelden te maken gehad hebben met beoordelingsselectie-methoden en in de toekomst wellicht zelf zullen beoordelen er meer bewust te maken dat een en ander niet zo vanzelfsprekend is.

1. Het meest bekend is de beoordeling die deel uitmaakt van een beslissingsprocedure als: toelating, selectie, overgang, eindexamen. Met betrekking tot deze soort beoordelingen willen wij het volgende stellen, ons baserend op (exacte) psychometrische principes of huidige onderwijskundige inzichten (voor bestrijding vatbaar):

a.) Deze "selectieve" beoordeling is de belangrijkste vorm van beoordeling die in de huidige onderwijspraktijk voorkomt.

Vele andere soorten beoordeling zijn hieraan ondergeschikt. Vorm en inhoud van het gegeven onderwijs worden er in hoge mate door bepaald. Het studiegedrag van de leerling is er grotendeels op gericht. Dit is een hoogst ongewenste situatie.

b) De verantwoordelijkheid voor de studieprestaties wordt primair gelegd op de schouders van de leerling, niet op die van de docent.

Dit resulteert bijv. in het verschijnsel dat men spreekt van "schoolzwakke kinderen", waar het juist zou zijn om van "kindzwakke scholen" te spreken (S.Soutendijk); het is evident dat de ene of de andere definitie een totaal verschillende aanpak van de betreffende onderwijsproblemen impliceert.

c) Een rechtvaardiging van de praktijk van voortdurende periodieke selectie bestaat noch op psychologische gronden (ontwikkelingspsychologie, persoonlijkheidsleer gepaard aan leerpsychologie), noch op psychometrische gronden (de wiskundig-statistische basis voor een dergelijke selectieprocedure is moeilijk of geheel niet te geven). Voor het eerste zij bijv. verwezen naar het werk van Guilford over intelligentie (1967, McGraw-Hill), voor het laatste naar Cronbach & Gleser: Psychological Tests and Personal Decision (1965, Univ. of Illinois Press).

d) Dit soort beoordelingen heeft de neiging globaal te zijn, bijv. het combineren van cijfers voor verschillende vakken om zodoende tot een overgangsbepaling te komen.

Ook als toelichting op de "psychologie" in de vorige stelling, het volgende: zoiets als bijv. "de intelligentie" van iemand (of zo men wil "begaafdheid") is niet iets dat in één score te vangen is, er zijn vele vormen van specifieke intelligentie of begaafdheid, die tamelijk weinig onderling samenhangen.

Bloom vond in een onderzoek, dat, wanneer we als definitie van be

gaaftheid geven het "behoren tot de beste 10%", ongeveer 3/5 van een

bepaalde groep studenten tot de beste 10%, dus tot de "begaafden",

behoorden op één of meer primaire vaardigheden die getoetst werden.

Op een algemene intelligentietest zou slechts 1/10 tot de "begaafden" gerekend worden.

Ons onderwijs houdt onvoldoende rekening met de unieke begaafdheden van studenten en leerlingen.

Tenslotte, deze hantering van beoordelingen heeft geresulteerd

in een studie-toets problematiek waarin buitensporig veel aandacht aan "betrouwbaarheid" van de toets, en aan "constante normen", "objectieve" beoordelingsmethoden etc. is besteed. Dit heeft weliswaar vele zeer slimme en ingewikkelde methoden opgeleverd, maar deze zijn van veel minder belang wanneer we beoordeling gaan gebruiken voor andere doeleinden dan selectie-beslissingen.

2. Met het voorgaande wil niet gezegd zijn dat er geen legitieme selectie situaties of selectie methoden ten behoeve van het onderwijs zouden zijn. Waar stelling tegen wordt genomen is, dat vrijwel alle beoordelingen in het onderwijs gericht zijn op het nemen van selectie beslissingen. Op deze wijze worden andere functies die beoordelingen kunnen vervullen, veronachtzaamd, zoals bijvoorbeeld beoordeling als onderdeel van het leerproces.

a) Snelle feedback aan de leerling over zijn verrichte handelingen is een noodzakelijke voorwaarde wil er van leren sprake zijn. Deze onmiddellijke feedback kan in diverse vormen gegeven worden: aanwijzingen door de leraar, verschaffen van de juiste antwoorden bij het maken van opgaven, discussie-methoden en andere vormen van groepswork waar de eigen activiteiten voortdurend onderwerp van bespreking in de groep vormen, etc. De meerkeuze studietoetsen hangen nauw samen met de ontwikkeling van geprogrammeerde instructie, een onderwijsvorm die bij uitstek gericht is op het voortdurend geven van feedback aan de student m.b.t. zijn vorderingen.

Wanneer genoemde vormen van onderwijs worden afgesloten door selecterende toetsen, ontstaat de situatie dat de bevordering van het

leerproces door feedback doorkruist kan worden door de bekende proef

werk-, tentamen- of examengerichtheid. De student is dan nl. minder

gericht op goede onderwijs-prestaties dan op goede toets-prestaties,

twee verschillende zaken!

3. Naast beoordeling met snelle feedback als leermiddel is de beoordeling met feedback aan de student over zijn vorderingen of tekortkomingen van belang voor het leerproces op langere termijn. Deze vorm van feedback levert de student concrete aanwijzingen op welke onderdelen van een vak hij in het bijzonder zijn aandacht zal moeten richten. Deze soort van beoordeling kan een sterk motiverende invloed hebben op de

inzet van de student.

a) Het louter geven van cijfers is een onderwijskundig gezien irrelevante praktijk. Een cijfer heeft geen of ongewenste praktische consequenties voor de handelwijze van leerling of docent, behalve voorzover het de selectiekansen betreft.

Dit impliceert bijvoorbeeld, dat een proefwerkbespreking geen luxe is, of een tegemoetkoming aan de wensen van de leerling, maar juist tot één van de kernpunten van het onderwijs behoort. De bespreking dient daarbij zo veel mogelijk op de individuele leerling gericht te zijn. Het gaat er om dat de leerling informatie krijgt over welke vaardigheden hij beheerst. en aan welke aspecten meer aandacht gegeven moet worden. Meerkeuze studie-toetsen kunnen hier gebruikt worden, immers hier ligt een pracht kans om met gebruikmaking van geschikte computer programma's iedere leerling op zeer korte termijn een zeer gedetailleerd inzicht te geven in zijn vorderingen op basis van zijn toetsprestaties. Met andere woorden we willen pleiten voor een meer op het individu toegesneden onderwijs, waarbij de meest moderne toets-technieken ingeschakeld worden om de individuele leerling zoveel inogelijk dát onderwijs te geven dát hij, qua vorderingen en tempo, het best kan verwerken. Een en ander sluit natuurlijk niet uit dat een zelfde procedure ook voor kleine groepen gehanteerd kan worden.

4. Niet alleen feedback naar de student is van belang, feedback naar de docent is van belang voor een juiste programmering en opzet van het onderwijs.

a) Eén van de belangrijkste functies van beoordelingen zou moeten zijn het verkrijgen van informatie over het functioneren van het onderwijs, meer in het bijzonder in welke richting verbeteringen ingevoerd moeten worden.

Hier is niet van belang of school A op een bepaalde studietoets lagere prestaties haalt dan school B. Ten eerste kan dit aan de beperktheid van de toets liggen, ten tweede, en daar gaat het ons om, geeft een dergelijk resultaat geen enkele aanwijzing over de oorzaak, nl. welke specifieke tekorten er in het onderwijs aan school A optreden. Of op welke wijze de kenmerken van de leerlingen, in termen van specifieke vaardigheden en begaafdheden, afwijken van die van school B. (Een uitstekend artikel over deze problematiek van de hand van de bekende onderwijspsycholoog Lee J.Cronbach, getiteld Evaluation for Course Improvement, is te vinden in N.E.Gronlund (ed.) Readings in Measurement and Evaluation, 1968, McMillan Cy. Genoemd boek bevat zeer veel informatie en

kritische artikelen over beoordeling in het onderwijs.
Cronbach:

"Onderwijsbeoordeling (evaluatie) moet informatie geven over de veranderingen die resulteren uit het gegeven onderwijs, en moet aanwijzingen geven over onderwijs-onderdelen die herziening behoeven. Onder het effect van onderwijs moeten we meer algemene resultaten vatten dan alleen de inhoud van een bepaalde cursus zelf: attitudes, beroepskeuzen, algemene inzichten en intellectuele vaardigheden, en het vermogen tot (zelfstandige) verdere studie op een bepaald gebied. Analyse van de verrichtingen op afzonderlijke items of soorten van problemen geeft meer informatie dan analyse van samengestelde scores. Het is niet noodzakelijk of gewenst om dezelfde toets aan alle leerlingen te geven; beter is zoveel mogelijk vragen te geven door de vragen te verdelen over kleinere groepen uit het totale aantal leerlingen. Kostbare technieken als mondelinge ondervragingen en schriftelijke werken kunnen dan op kleine steekproeven uit de leerlingen gedaan worden, wat vooral daar van belang is waar het op deze wijze toetsen van iedereen te kostbaar is of te veel docententijd zou vergen".

5. Tenslotte feedback aan de docent als steun voor de programmering en individualisering van het onderwijs. De docent moet zo volledig mogelijke informatie hebben over de stand van de vorderingen van zijn leerlingen, om het te geven verdere onderricht daar zo goed mogelijk op aan te kunnen laten sluiten.

Niet alleen gaat het hierbij om de inhoudelijke kanten van het onderwijs, maar ook factoren als de psychologische ontwikkeling van de leerling. Een voor de hand liggend voorbeeld is de schoolrijpheid van de 5 tot 7-jarigen.

Hoewel nog niet voldoende kennis over dergelijke ontwikkelingsstadia en de daarbij het best passende vormen en inhouden van onderwijs, beschikbaar is, is het belang hiervan evident. Wanneer de nederlandse onderwijswereld verlost is van de monomanie van de selecterende beoordeling, zal voldoende aandacht hiervoor vrij kunnen komen. Ook de democratisering van het onderwijs, in de zin van gelijke kansen voor iedereen" is hier in het geding. Dit- eist feedback op grote schaal, om de collectieve verantwoordelijkheid van alle bij het onderwijs betrokkenen, richting te kunnen geven naar initiatieven die de starheid van ons nederlandse onderwijs kunnen doorbreken.

Vrijwel niets is tot nog toe expliciet gezegd over onderwijsdoelstellingen, impliciet echter veel. Bij iedere vorm van beoordelen spelen de doelstellingen die in het

onderwijs gehanteerd worden een hoofdrol, immers, zij vormen de toetssteen voor de kwaliteit van het onderwijs. Worden de doelstellingen niet geformuleerd, dan is ook niet duidelijk naar welke maatstaven beoordeeld moet worden, wat kan resulteren in verschijnselen als "de wet van Posthumus" (het vrijwel constante percentage afvallers bij de diverse soorten van selectieve beoordeling in het onderwijs). Het doelstellingenprobleem is echter dermate complex dat we het beter buiten deze discussie over beoordeling kunnen houden, zonder daarmee het fundamentele belang ervan te vergeten.

januari 1971. B.Wilbrink