

Alternatieve toetsmethoden in studentgericht onderwijs¹

D. Sluijsmans & F. Dochy

“The best way to discover how students think, or to diagnose where they are having difficulties in learning--which, aside from accountability and placement, is the main reason for testing--is to give them as much range as possible to express themselves fully (Archbald & Newmann, 1988), and to assess their learning in its natural context, as they make active use of the skill” (Gardner, 1994).

Vele universiteiten en hogescholen zijn momenteel doende innovatieprogramma's voor hun onderwijs te implementeren. Het kernbegrip daarbij is studentgericht onderwijs: leren leren, probleemgestuurd leren en casusgeoriënteerd leren. Daarnaast is de laatste vijf jaar op nationaal en internationaal niveau een brede belangstelling waar te nemen voor de zogenaamde alternatieve toetsmethoden. Deze verzamelnaam heeft betrekking op de toetsen die gericht zijn op het meten van kennis en vaardigheden die gepaard gaan met complexe authentieke handelingen en denkprocessen. In dit artikel wordt een overzicht gegeven van de betekenis van alternatieve toetsmethoden in het hoger onderwijs, wordt het achterliggend ideeëngoed beschreven en wordt betoogd dat studentgerichte onderwijsvormen het gebruik van dergelijke toetsmethoden noodzakelijk maken.

Inleiding

Recente veranderingen in het hoger onderwijs richten zich in grote mate op het realiseren van de “krachtige leeromgeving” (De Corte, 1990). Met name die omgeving waarin student cooperatief leren, uitgaande van concrete authentieke problemen en gericht op vooral het toepassen van kennis en vaardigheden. Het werkveld heeft meer stem gekregen in de inhoud van het onderwijs (Pilot, 1997) en het is voor velen duidelijk geworden dat in de informatiemaatschappij van morgen afgestudeerden naast de basiskennis van de wetenschap vooral vaardigheden moeten bezitten als probleemoplossen, analyseren, synthetiseren, coachen, leiding kunnen geven, presenteren, kritisch evalueren, etc. Uit de organisatie van het onderwijs en het leren zal moeten blijken dat studenten niet alleen goede probleemoplossers worden die kunnen werken in multidisciplinaire teams, maar ook dat zij kunnen reflecteren op de kwaliteit van het eigen werk en op het werk van anderen. De combinatie van deze vaardigheden maken dat de afgestudeerde als een *reflective practitioner* aan de slag kan (Schön, 1987). Vormen van leren leren, probleemgestuurd onderwijs, competentiegericht onderwijs, combinaties van werken en leren of coöperatief leren worden gebruikt om dit niveau te bereiken.

Bij een dergelijke verandering van onderwijskundige aanpak kan men echter op het niveau van de toetsing niet blijven steken bij kennisgerichte toetsen alleen. Erger nog, een overschakeling naar leren leren of probleemgestuurd onderwijs met een status quo bij de toetsing zal onherroepelijk leiden tot verwarring bij studenten en tot een verdwijnen van de meerwaarde van de innovatie. De toetsing blijft steeds de meest sturende factor in het onderwijs voor het leergedrag van studenten. Met andere woorden, ook voor de beoordeling zal een meer externe oriëntatie gezocht moeten worden. Uit het tentamengebouw zal moeten blijken dat door het werkveld gewenste kennis en vaardigheden beoordeeld zijn op voldoende niveau (Moerkerke & Terlouw, 1998).

De context voor alternatieve assessmentvormen

¹ Dit artikel is eerder verschenen in het Tijdschrift voor Hoger Onderwijs, december 1998

Het ontwerp en de implementatie van studentgerichte krachtige leeromgevingen waarbinnen het leren leren en het verwerven van vaardigheden en competenties centraal staat, heeft consequenties voor de wijze waarop wordt beoordeeld. De veranderingen in de taken van de student en de taken van de onderwijsinstelling/docent geven aan dat de beoordeling niet meer alleen in de handen hoeft te liggen van de docent. Studenten worden opgevat als zelfstandige, autonome en initiatiefrijke individuen die in grote mate zelf hun leerproces sturen (Sluijsmans, 1997). Kennis en vaardigheden worden bij voorkeur niet overgedragen via reproductieve technieken, maar op actieve wijze door studenten worden verworven. Veel kennis is situationeel bepaald. Door de beroepspraktijk zo goed mogelijk in de leeromgeving te weerspiegelen kan een betekenisvolle context voor kennisverwerving ontstaan (Brown, Collins & Duguid, 1989). De student moet leren kritisch te kijken naar eigen handelen en de kennis over onderwijs. De beoordeling zal in een krachtige leeromgeving niet alleen gericht zijn op het "afrekenen" van de student, maar zal ook een rol moeten hebben in het structureren en monitoren van het leerproces. De mogelijke toetsvormen die hierbij kunnen helpen worden gevat onder de term alternatieve assessment. Deze in plaats van assessment vraagt eerder om een analyse van de sterkten en zwakten van de kennis en vaardigheden, dan om toetsen die tot een simpele eendimensionale score leiden. Toetsen dienen niet alleen certificerend te zijn, maar ook sturend voor het leerproces. Anderzijds is het onmiskenbaar dat men bij onderwijsinnovaties veelal veranderingen doorvoert en pas later gaat denken aan de toetsing. Gezien de toetsing de meest sturend factor is voor studenten, is het mogelijk dat de vernieuwing aan kracht zal verliezen omdat studenten niet geneigd zijn inspanningen te leveren die niet worden beloond. Een voorkomende valkuil is dat het onderwijs wordt ingericht als krachtige leeromgeving (vanuit probleemtaken, cooperatief leren, gericht op probleemoplossen en andere academische vaardigheden), maar dat traditionele toetsing van in hoofdzaak feitenkennis blijft domineren. In voorkomend geval vertonen studenten de neiging de onderwijsgroepen niet bij te wonen, geen inspanning te leveren om authentieke problemen op te lossen, maar richten ze zich op het van buiten leren van de volgens hen relevante literatuur. Studenten zijn in dat opzicht economisch denkende mensen.

In het huidige onderwijs zijn veel verschillende soorten toetsen te onderscheiden, veelal gericht op het toetsen van kennis. Voorbeelden zijn de toets met open vragen, het essay-examen, de multiple-choice-toets, en combinaties hiervan. Alternatieve assessment is meer dan dat: studenten worden beoordeeld op basis van hun actieve prestatie om kennis te gebruiken op een creatieve manier om zinvolle problemen op te lossen, het leerproces te bevorderen en kritisch te evalueren. Deze problemen zijn reële problemen, authentieke representaties van problemen in de werkelijkheid. Nieuwe assessmentvormen wijken veelal op één of meerdere aspecten af van het traditionele beoordelen. In ieder geval op het waarom van het toetsen. Voorts leiden innovaties ten aanzien van wie, wat, hoe en wanneer getoetst worden tot andere beoordelingsvormen.

Vormen van alternatieve assessment

Hoe kan kennis die studenten zelf hebben verworven en vaardigheden als probleemoplossen, leren leren, informatie kunnen zoeken, discussiëren, vergaderingen leiden, notuleren, synthetiseren, analyseren, kritische evalueren, enzovoort worden beoordeeld? In dit artikel wordt een descriptief overzicht gegeven van beoordelingsvormen die gebruikt worden om één of meerdere van deze aspecten te toetsen. Deze zijn geordend volgens de manier waarop ze afwijking vertonen van het traditioneel toetsen. Het zal de lezer duidelijk zijn dat het vaak gaat om varianten van toetsen of om nieuwere beoordelingsvormen die veelal gecombineerd worden gebruikt. Het doel is het overzicht van deze vormen te systematiseren door het accent te leggen op hun specifieke kenmerken waarin ze verschillen. Onderzoek naar het gebruik van dergelijke nieuwe toetsvormen is

elders gepubliceerd (Birenbaum & Dochy, 1995; Falchikov & Boud, 1989; Sluijsmans, Dochy & Moerkerke, 1998).

Doel van assessment

Assessment beoogt de student kennis te geven van de leerresultaten: het voorbije leerproces wordt opnieuw bekeken en het toekomstige leerproces wordt (bij)gestuurd. Tevens verwerft de student een bewijs van competentie (certificering), dat de student van een professionele functie in de maatschappij verzekert. Boud (1990) onderkent de kloof tussen datgene dat van de student wordt vereist in het hoger onderwijs en wat gebeurt in het professionele leven. Hij benadrukt de behoefte aan onderzoek naar assessment om te kijken of de doelen van deze assessment overeenkomen met de doelen van het hoger onderwijs. Het is nu alom erkend dat het hoofddoel van hoger onderwijs is studenten te helpen ontwikkelen naar de reeds eerder genoemde "reflective practitioners", die kritisch kunnen denken over hun eigen professioneel handelen. Het is door deze overtuiging dat bijvoorbeeld probleemgestuurde onderwijsprojecten opgestart worden in het hoger onderwijs gebruik makend van alternatieve assessmentvormen. Door deze vormen van assessment, zoals deze in het volgende aan bod zullen komen, worden gegevens verkregen over de voortgang in ontwikkeling bij studenten. Door alternatieve assessment krijgen deze gegevens mede een formatief karakter en dragen ze bij tot de ontwikkeling van reflectieve en probleemoplossende vaardigheden.

De rollen in het beoordelingsproces: wie beoordeelt

In het klassieke onderwijs wordt weinig de vraag gesteld wie de beoordeling uitvoert en wie de criteria bepaalt. Dat is veelal de instelling of docent. Nieuwere vormen van onderwijs (leren leren, competentiegericht onderwijs, probleemgestuurd onderwijs) stellen deze vraag wel. Daarbij kan de student zelf, de medestudenten of de docent en student samen verantwoordelijk zijn voor het beoordelen. Een positief antwoord op deze vragen zou respectievelijk leiden tot, self, peer-, of co-assessment.

Self-assessment

Bij self-assessment zullen studenten zichzelf toetsen op hun verworvenheden of vorderingen. Dat kan betekenen dat zij een organisatorische en/of sturende krijgen. Dit heeft onder andere als consequentie dat studenten zichzelf moeten kunnen beoordelen. Onderzoek toont aan dat studenten dat ook moeten leren (Boud & Falchikov, 1989). Self-assessment wordt ingezet om de actieve betrokkenheid van studenten bij hun eigen leerproces te verhogen en het wordt vooral gebruikt als formatief toetsmoment om reflectie te bewerkstelligen (Boud, 1995). Naast en tijdens het beoordelingsproces speelt reflectie een belangrijke rol. Dit reflectieproces heeft tot doel tot een verandering te leiden in denken, houding en gedrag.

Peer-assessment

Wanneer studenten elkaar beoordelen is sprake van peer-assessment (Falchikov, 1995). Bij deze vorm van beoordelen gaat het niet altijd om de uiteindelijke beoordeling, maar ook om de wijze waarop de peer-assessment een bijdrage levert aan de kwaliteit van het leerproces. Peer-assessment komt veelal voor in onderwijs- of opleidingssituaties waar het werken in groep een centrale rol speelt. Leren in kleine groepen (6-15) kan gebruikt worden om "dieper leren" aan te moedigen, om meer autonomie aan de studenten te geven door een deel van de verantwoordelijkheid voor doceren en leren naar de studenten te schuiven.

Co-assessment

Bij een co-assessment tenslotte hebben studenten en docent samen een rol in het beoordelingsproces. Met de docent maakt de student een selectie in de beoordelingscriteria.

In een co-assessment hebben de studenten de gelegenheid elkaar te beoordelen, maar wordt het uiteindelijke oordeel uitgesproken door de docent (Hall, 1995). Co-assessment heeft de volgende drie bedoelingen. Op de eerste plaats biedt co-assessment de student de gelegenheid de rol van de docent in te nemen. Op de tweede plaats verwerft de student inzichten in het beoordelingsproces, waardoor hij beter in staat is anderen te beoordelen. Ten slotte wordt de student meer vaardig in het beoordelen van zichzelf. Co-assessment is een dialoog tussen docenten en studenten. Deze samenwerking kan plaatsvinden op verschillende niveaus en in verschillende fasen. Het kan louter formatief zijn of ook deel uitmaken van het summatieve beoordelingsproces. Daarbij kunnen twee modellen worden onderscheiden. In het ene model liggen de criteria op voorhand vast. Dit wordt duidelijk gemaakt aan de staf. Er zijn dus geen aanpassingen mogelijk, noch door studenten, noch door docenten. In het andere model dragen studenten zelf criteria aan. Indien de studenten zelf moeilijk criteria kunnen vinden presenteert de docent een lijst van criteria en laat de studenten elk item bediscussiëren, verklaren, argumenteren voor en tegen, voorstellen doen, toevoegingen, enzovoort. Indien de studenten zelf kritisch kunnen denken in termen van criteria laat de docent de studenten een lijst met criteria aanleggen. Daarna bediscussiëren de studenten deze lijsten totdat er een uiteindelijke lijst is.

Self-, peer-, en co-assessment zijn verzameltermen voor toetsbenaderingen met verschillende verschijningsvormen. De gedachte is steeds dat studenten worden betrokken bij de beoordeling, om zo de kwaliteit van het leren te bevorderen. In de praktijk blijkt dat ook het geval te zijn (Sluismans, Dochy, & Moerkerke, 1998). Klassieke vormen van assessment liggen meestal in handen van de docent, die de criteria opstelt, producten van de studenten beoordeelt en feedback geeft. Aan de andere kant van het continuüm staat de student die zijn eigen doelen stelt en zijn werk zelf beoordeelt. Deze twee “modellen” representeren de extremen van een continuüm dat de graad van autonomie van de student voor zijn leren bepaalt (figuur 1).

Docentgestuurd
Studentgestuurd

traditionele	co-	peer-	self-
vormen van	assessment	assessment	assessment
assessment			

Figuur 1: Docentgestuurde assessment versus studentgestuurde assessment

Duidelijk is dat co-assessment de eerste stap is naar nieuwe vormen van assessment. De betrokkenheid van de studenten in het assessment-proces is groter en vergemakkelijkt het leren. Het wordt sneller aanvaard dan peer en self-assessment, gezien de meerderheid van de docenten gewend is aan autonomie met betrekking tot het beoordelingsproces. De docenten blijven betrokken en hebben de doorslaggevend stem bij het beoordelen.

Self-, peer- en co-assessment maken deel uit van het veranderingsproces naar een meer studentgecentreerde benadering. Deze verschuiving eist een verandering van nadruk op de norm naar nadruk op het criterium van beoordeling, van de puur summatieve naar de formatieve en summatieve beoordeling, van externe naar interne beoordeling en van produktbeoordeling naar produkt- en procesbeoordeling (Somervell, 1993).

Vormen als self-, peer- en co-assessment hebben een aantal voordelen. Op de eerste plaats is er een hoge betrokkenheid van studenten: wanneer een student zelf mee de resultaten van de assessment en de ontwikkeling ervan heeft bepaald, zal hij deze eerder accepteren.

Tevens is het een leermoment voor de student door de feedback die men krijgt en het zelf mee beoordelen. Ook versterken deze vormen van beoordelen de motivatie: de student krijgt verantwoordelijkheid (men ondergaat niet slechts een programma) (Somervell, 1993). Er is echter ook een nadeel. Studenten zijn niet gewend kritiek te geven op hun vrienden en voelen zich nog onervaren bij het beoordelen van anderen. Men is bang om onfair beoordeeld te worden. Dit nadeel is echter tijdelijk van aard en kan door adequate training worden weggewerkt (Williams, 1992). Studenten moeten leren om risico's te nemen, kritisch te reflecteren en anderen te durven confronteren met feedback.

Samenvattend kenmerken self-, peer- en co-assessment zich door het volgende:

- het gaat hier niet meer om een eenzijdige beoordeling, de studenten worden actief bij het beoordelingsproces betrokken;
- ook de vaardigheid probleemoplossen en het proces van leren leren wordt beoordeeld;
- na een tijdje vertonen de studenten een authentieke nauwgezetheid en volledigheid in de manier waarop ze zichzelf beoordelen;
- "Meeting criteria improves learning"; studenten zijn actieve participanten in hun eigen leerproces. Ze presteren beter als ze weten naar welke doelen ze toewerken, wanneer ze de gelegenheid krijgen om de relevantie van bepaalde criteria te onderzoeken en wanneer ze begrijpen hoe hun eigen kunnen vergeleken wordt met een set van gevestigde criteria;
- ontwikkeling van interpersoonlijke vaardigheden: de studenten leren via discussies, gesprekken, e.d. hun mening uiten en voor zichzelf en eventuele medestudenten (bij de combinatie met *self- and peer-assessment*) opkomen;
- studenten krijgen inzicht in groepsdynamiek;
- studenten worden voorbereid op hun toekomstig professioneel functioneren..

Bij het gebruiken van deze assessmentvormen kan het voorkomen dat studenten criteria voorstellen die de docent niet haalbaar acht. Oplossingen hiervoor zijn dat de docent de criteria bediscussieerd en beredeneerd zodat studenten overtuigd worden van de rationaliteit ervan of dat de docent de criteria door de studenten laat gebruiken en hen zelf laat ontdekken of deze nu relevant zijn of niet. De uiteindelijke beslissing hierover ligt bij de docent.

Tenslotte bestaat in het begin van het gebruik van co-assessment een lichte tendens zichzelf minder goed te beoordelen (Boud, 1981). Studenten hebben jarenlang enkel ervaring opgedaan met een eerder eenzijdig beoordelingssysteem, wat voor sommigen een negatief zelfbeeld opleverde. Dit verdwijnt echter naarmate nieuwe assessmentvormen meer en meer ingeburgerd raken. Tevens vraagt collaborative assessment in het begin meer tijd, maar de opbrengst (leeropbrengst, rendement, validiteit beoordeling) is groter.

In probleemgestuurd onderwijs worden peer- en co-assessment gebruikt om onder meer vaardigheden van studenten tijdens de onderwijsgroepsbijeenkomsten te beoordelen. Daarnaast hebben deze beoordelingen het effect dat studenten de beoordeling fair vinden, hun motivatie wordt vergroot en dat uitval wordt voorkomen (Sluijsmans, Dochy & Moerkerke, 1998).

Performance assessment: wat wordt beoordeeld

Nieuwe beoordelingsvormen meten niet alleen kennis, maar richten zich evenzeer op het toetsen van vaardigheden. In deze bijdrage wordt het accent gelegd op toetsen waarin vaardigheden centraal staan. Datgene wat beoordeeld gaat worden kan dan worden geplaatst onder de noemer performance assessment. Performance assessment doet een beroep op studenten om specifieke vaardigheden te demonstreren oftewel de kennis en vaardigheden die zij beheersen toe te passen (Keeves, 1994). Performance assessment is

ontstaan als een reactie op de vele gestandaardiseerde multiple-choice examens. Studenten uit sociaal lagere klassen werden bij deze toetsen benadeeld omdat deze toetsen geen beroep doen op de gehele context waarin die studenten leven. Een tweede kritiek is dat de traditionele multiple-choice examens te weinig beroep doen op metacognitieve vaardigheden (Moerkerke & Dochy, 1995).

Performance assessment wordt gedefinieerd als een systematische poging om de toepassingsvaardigheid van de lerende te meten. Met andere woorden, is de lerende in staat om reeds geziene kennis te gebruiken bij het oplossen van nieuwe problemen of bij het vervolledigen van specifieke taken? Werkelijke of gesimuleerde assessment-oefeningen worden gebruikt om oorspronkelijke antwoorden aan het licht te brengen, te observeren en te scoren (Gipps, 1994, p.99). Van de studenten worden duidelijk meer dan alleen maar basisvaardigheden verwacht. Zo wordt hen bijvoorbeeld gevraagd een essay te schrijven, in een groep een wetenschappelijk experiment uit te voeren, het antwoord op een wiskundig probleem schriftelijk te beargumenteren of een portfolio bij te houden van hun prestaties (O'Neil, 1992, p.15). Performance assessment doet een beroep op vaardigheden van studenten, meer specifiek op de uitvoering ervan. De beoordeling is veelal geïndividualiseerd en gecontextualiseerd en dient niet op één tijdstip afgenomen te worden. Dit kan een langere periode omvatten. De beoordelingscriteria dienen gekend te worden. Wanneer studenten weten welke de criteria zijn, wat ze eigenlijk moeten bereiken gaan ze ook gemotiveerder en beter presteren.

Performance assessment levert een belangrijke bijdrage aan het creëren van een volledig beeld van het bereik van een bepaalde student (Gipps, 1994; O'Neil, 1992). Aangezien instructie en assessment sterk (zouden moeten) samenhangen, zal performance assessment tevens aanleiding geven tot een waardevollere instructie. Performance assessment wordt immers meer gebruikt ter beoordeling van het proces dan van het product. Na deze beoordeling, die tot doel heeft foutieve redeneringen, strategieën, enzovoort op te sporen, ontvangen de studenten hieromtrent feedback. Daarbij zijn docenten meer dan anders betrokken bij het beoordelingsproces. Zo zijn het vaak de docenten die de taken ontwikkelen, de criteria voor scoring formuleren en deze scoring uitvoeren (al hoeft dit laatste niet het geval te zijn, vergelijk co-, peer- en self- assessment). Op deze wijze draagt performance assessment bij tot hun professionele ontwikkeling.

Als nadelig kan worden beschouwd dat de invoering van performance assessment gepaard gaat met verhoogde kosten. Zo is het onder andere wenselijk docenten te trainen bij het afnemen van dergelijke assessment-taken en het scoren ervan. Ook duurt de afname van performance assessment ontegensprekelijk langer dan bij gewone gestandaardiseerde toetsen: de taken moeten zorgvuldig worden ontwikkeld. Hetzelfde geldt voor de scoring, er is de veranderde instructie die van de docent een grondige voorbereiding vereist (Dochy & Moerkerke, 1997). Tenslotte is bij performance assessment een aantal psychometrische potentiële valkuilen te formuleren. Deze waarschuwingen zijn vooral belangrijk wanneer performance assessment wordt gebruikt voor "accountability"-doeleinden (Mello, 1993). Met dit laatste wordt verwezen naar eventuele sancties die verbonden zijn aan de score op de beoordelingen. Het is belangrijk de generaliseerbaarheid van de beoordelingen aan te tonen, alsook de vergelijkbaarheid tussen studenten en scoorders en op die manier de consistentie ervan. Vooral wat deze laatste eis betreft kan een aantal hulpmiddelen worden geformuleerd: de afname van de toetsen gebeurt in gelijke groepen, iedereen ontvangt dezelfde (gestandaardiseerde) instructie, de scoringscriteria zijn over de groepen heen gelijk én vooraf voor iedereen duidelijk. Vooral deze laatste hulpmiddelen wijzen op het belang van goed overleg tussen de docenten. De hulpmiddelen kunnen echter niet verhinderen dat velen gestandaardiseerde toetsen blijven prefereren in geval van "accountability"-doeleinden.

Voor het ontwerp van performance assessments hebben Wiggins (1992) en Moerkerke (1996) basisvereisten geformuleerd waaraan een dergelijk ontwerp moet voldoen (zie tabel 1). De vereisten hebben betrekking op de taken, de criteria of standaarden en de scoring. Zij worden concreet vertaald binnen de afzonderlijke fasen van het ontwerpproces.

Tabel 1: Ontwerp van performance assessments

a.	Kiezen wat getoetst wordt	<ul style="list-style-type: none"> • kies voor curriculumonderdelen die binnen conventionele toetsen zelden of nooit aan bod komen • denk aan de getoetste kennis als zijnde een middel om een bepaald product te realiseren
b.	De taken ontwerpen	<ul style="list-style-type: none"> • probeer betekenisvolle taken te formuleren • contextualiseer de taken • ontwerp prestaties, geen driloefeningen • verfijn de opgestelde taken door ze vanuit de - in de volgende fase opgestelde - scoringscriteria nogmaals te herlezen
c.	Het afnemen van de toets	<ul style="list-style-type: none"> • tracht de toetsafname zo authentiek mogelijk te realiseren • omschrijf duidelijk hoe de toets door de docent wordt geïnstrueerd • taken dienen zoveel mogelijk voor zichzelf te spreken
d.	Overwegingen voor het scoren	<ul style="list-style-type: none"> • score wat belangrijk is, niet wat het gemakkelijkst te scoren is • controleer de scoringsstandaarden op voldoening aan toekomstige eisen

Een beoordeling geven van de mate waarin studenten beschikken over de vooropgestelde vaardigheden in het oplossen van vakinhoudelijke problemen is geen eenvoudige zaak. De taken die het individu aanpakt hebben immers een complex patroon van interacties tussen een cognitieve, een motorische, een sociaal-psychologische en een handigheidscomponent in zich. Een beoordeling van het gedrag en/of de producten naar aanleiding van deze taken veronderstelt een doelgerichte beslissingsprocedure die ook gerechtvaardigd moet kunnen worden. Uit onderzoek blijkt dat er voor kennisdomeinen in het hoger onderwijs nauwelijks methoden voor de constructie van dergelijke opgaven zijn ontwikkeld of dat ze te sterk beperkend zijn (Boud, 1990; Moerkerke, 1996).

Wijzen van beoordelen

Het hoe van het beoordelen is thans ook aan verandering onderhevig. Er is meer dan het gebruiken van lijsten van items of itembanken, al is deze methode nog steeds erg goed bruikbaar bij vooral op kennis en inzicht gerichte beoordelingen. Binnen de stroming van de alternatieve toetsing wordt veel aandacht besteed aan: authentieke probleemtoetsen (als open-boek toetsen), casusgerichte toetsen, portfolio's, en assessment centers (Birenbaum, 1992, 1994; Birenbaum & Dochy, 1995). In deze bijdrage wordt dieper ingegaan op het portfolio als beoordelingsinstrument en op casusgerichte beoordeling, met name de Overall toets (OAT) (Segers, 1995; 1996). Voor andere vormen als het assessment center wordt verwezen naar Bruijns en Pieké (1997), De Rijke en Beemer (1995) of Birenbaum en Dochy (1995).

Portfolio assessment

Het idee en het gebruik van portfolio's in onderwijs is afgeleid van het kunst- domein: ze dienden om de verwezenlijkingen van de artiest te kunnen aantonen. Het gebruik van portfolio's in een educatieve omgeving houdt in dat de student beschikt over een verzamelmap die zichtbare informatie geeft over de persoonlijke prestaties en ervaringen van die student. De student wordt ertoe aangezet te reflecteren op zijn werk, activiteiten,

keuzes en doelen. In die zin is een portfolio veel ruimer dan een verzameling van geselecteerd schriftelijk werk.

Bij portfolio assessment worden de studenten actief betrokken bij de assessment: portfolio's worden gemaakt door en met de student, ze worden de student niet opgelegd. Het is in deze een doelgericht aangelegde verzameling, die de moeite, de vooruitgang en de verwezenlijkingen van de student kunnen tentoonstellen, waarbij de studenten zelf participeren in het selecteren van inhoud, van criteria voor selectie, van criteria ter beoordeling van arbeid. Daarbij moet duidelijk bewijs zijn van reflectie op dit alles door de student. De studenten leren door middel van reflectie, denken, analyseren, synthetiseren en schrijven in hun portfolio's, door de feedback die ze krijgen. Bovendien zijn de leervaardigheden en -strategieën die ze zullen ontwikkelen voor het gebruik van de portfolio (data-verzameling, data-analyse, organiseren, interpreteren, reflecteren) wezenlijke vaardigheden van het life-long learning idee (Courts & McInerney, 1993).

De inhoud van het portfolio moet nauwkeurig geselecteerd worden om als bewijs te dienen dat er is geleerd. Het geheel moet aantonen wat de student weet en kan doen in een particulier domein. Ook de gemaakte vooruitgang moet eruit blijken. Studenten moeten goed op de hoogte zijn van wat in het portfolio moet zitten. Hierover worden richtlijnen verspreid en besproken. De criteria en standaarden waaraan de inhoud moet voldoen kan in overleg met de docent tot stand komen. Als de portfolio af is en klaar voor assessment schrijft de student een introductie waarin het doel staat, het onderwerp en het thema. Dit bevat ook een algehele samenvatting (van de zelfreflecties) op de portfolio als geheel.

Het portfolio is een "tool for learning" dat instructie en assessment integreert. Het wordt gebruikt voor een longitudinaal assessment van proces en produkt. Het bevordert de dialoog en de samenwerking tussen student en docent, waar de twee samenwerken bij het totstandkomen van de vorderingen van de student ten aanzien van het onderwerp. Het portfolio geeft een breed beeld van de leerresultaten van de student over een bepaalde inhoud, en levert bijgevolg een methode voor: inhoudelijke portrettering, het verhaal van de student als lerende te vertellen, het aanduiden van de sterke en zwakke kanten van de student. Ook in probleemgestuurd onderwijs komt het voor dat na elk blok de onderwijsgroep gevraagd wordt een produkt- of procesportfolio voor te leggen.

Portfolio-assessment kan worden beschouwd als authentic assessment. De authentieke assessment is gegroeid uit de overtuiging dat de resultaten niet effectief gemeten kunnen worden door het toetsen van geïsoleerde componenten, omdat de materie te complex is. Het is eerder zo dat deze assessment technieken tot doel hebben delen van de realiteit, probleem-oplossingssituaties te voorzien, in welke studenten globale toepassingen van relevante componenten van vaardigheden en kennis kunnen demonstreren. Als dusdanig is het meer een kenmerk van bepaalde assessments dan een aparte methode. Authentic assessment vereist dat studenten actief construeren en denken vanuit werkelijkheidsgetrouwe situaties of simulaties. Het ontlokt hogere orde denken en is veelal geïntegreerd met de instructie. Authentic assessment is in ieder geval sterk geïntegreerd in de praktijk, bijvoorbeeld via simulaties, video's van de praktijk of een casus met materiaal uit de beroepspraktijk (Meyer, 1992).

Overall toetsing (OAT): casusgerichte assessment

De Overall-toetsvorm werd ontwikkeld door Segers (1992; 1998) en wordt thans in diverse opleiding gebruikt waaronder de faculteit "Economische Wetenschappen en Bedrijfskunde" aan de Universiteit Maastricht. Hier wordt kort het programma en het gebruikte beoordelingsstelsel, dat een combinatie is van traditionele kennistoetsen en casusgerichte toetsen geïllustreerd. Het programma in het eerste jaar bestaat uit vier blokken van elk acht weken rond het centrale thema: "De markt als mechanisme van Coördinatie en Allocatie." Daaruit worden de volgende vier subthema's afgeleid: Bedrijven en Markten, Micro-economisch beleid en het Markt-, Fluctuatie-, Groei- en Stabiliseringsbeleid en tussentijdse Financiële Markten. Deze vier blokken worden doorlopen in groepen van 10 tot 12 studenten

begeleid door een "tutor". De groepen komen twee maal per week samen in sessies van elk twee uur om te werken aan de oplossing van een probleem, dat dus vanzelfsprekend volgens het hoofdprincipe van probleemgestuurd leren, een sturende rol heeft in het leerproces. Daarnaast worden sessies georganiseerd voor het beheersen van de nodige kennis op het gebied van wiskunde, statistiek, economie, informatica en schriftelijke communicatie en een afsluitend elektronisch economie game.

Hoe wordt nu in een dergelijk programma de beoordeling georganiseerd? In dit geval is er sprake van een combinatie van twee toetsen van verschillende aard. Op de eerste plaats is er de "kennistoets" die na elk blok wordt afgenomen. Deze toets is vooral gericht op het meten van een georganiseerd kennisbestand (feiten, symbolen, principes, concepten en modellen) uit een verschillende disciplines. Daarnaast is er de "Overall Toets" die tweemaal, namelijk na het tweede en het vierde blok, wordt afgenomen. Hier wordt nagegaan in welke mate studenten problemen kunnen herkennen, definiëren, analyseren en kunnen bijdragen aan de oplossing van het probleem door het gebruik van concepten, modellen en theorieën uit verschillende disciplines. Daarvoor krijgen de studenten na twee blokken telkens twee weken collegevrij voor zelfstudie op basis van een handleiding. Deze handleiding bevat op de eerste plaats de voornaamste doelstellingen van de "Overall Toets", ten tweede een uitgewerkt voorbeeld met bijbehorende toetsitems, verder wat praktische informatie en tenslotte 8 tot 10 artikelen van verschillende aard. Deze artikelen bevatten praktijkrelevante problemen. Na deze twee weken wordt de toets afgenomen die grosso modo nagaat of de studenten de kennis die ze hebben verworven in de onderwijsgroepsbijeenkomsten kunnen toepassen op de nieuw aangeboden problemen. Samengevat kan worden gezegd dat de "kennistoets" vooral declaratieve kennis toetst en de "Overall Toets" vooral procedurele kennis en probleemoplossende vaardigheden toetst (Segers, 1996; 1997; 1998).

Voorkennistoetsen en voortgangstoetsen

Tenslotte zijn een aantal toetsen te onderkennen die van het traditioneel examen afwijken door het moment waarop ze worden afgenomen. Uiteraard is er een tendens om tussentijdse formatieve toetsen af te nemen, maar anderzijds is er in het leren behoefte om meer aan summatieve procesbeoordeling te doen. Daarnaast noodzaakt de gerichtheid op rendement ook te kijken naar de mogelijkheden die voorkennis- en intaketoetsen bieden. Summatieve toetsen zijn traditiegetrouw eindtoetsen. In het onderstaande zal enige aandacht worden besteed aan voorkennis- en voortgangstoetsen.

Voorkennistoetsen

Voorkennis is een belangrijke studentvariabele bij een leerproces. Een sleutel tot de ontwikkeling van een geïntegreerde en generatieve kennisbasis is het voortbouwen op de voorkennis van de lerende (Dochy, 1997a,b). Assessment van voorkennis is niet alleen een vrij precieze predictor van leren en succes, maar voorziet bijkomend een nuttige basis voor instructie en begeleiding van het leerproces.

Er kan een onderscheid gemaakt worden tussen domein-specifieke en domein-overstijgende kennis, allebei essentieel voor het leren. Verschillen in voorkennis blijken immers de grootste bron te zijn voor intra- individuele en inter- individuele verschillen in cognitieve prestaties. Verder blijkt domein- specifieke kennis te kunnen compenseren voor lage intellectuele mogelijkheden, terwijl hoge intellectuele capaciteiten niet kunnen compenseren voor een laag niveau van voorkennis (Weinert & Schneider, 1995).

Om voorkennis te toetsen gebruikt men de zogenoemde voorkennistoetsen waaruit kennisprofielen in de vorm van grafieken van groeps- of individuele scores op de toets naar voor komen. Deze kennisprofielen kunnen op verschillende dimensies uitgezet worden, zoals bijvoorbeeld gedragsniveaus of inhoudsniveaus.

Er kunnen verschillende soorten voorkennistoetsen worden onderscheiden:

Op beginniveau:

- *Optimaal vereiste voorkennistoetsen*: over welke kennis en kunde moet de student beschikken om in optimale condities het geplande onderwijs te volgen? Een dergelijke toets is vrijwel noodzakelijk als het bijvoorbeeld gaat om een vak in het hoger onderwijs dat zijn equivalent in het secundair onderwijs niet kent of bij nieuwe onderwijsvormen als probleemgestuurd onderwijs of casusgestuurd onderwijs.
- *Vakspecifieke voorkennistoetsen*: deze toets bestaat veelal uit ongeveer 40 vragen die kennis en inzicht nagaan die specifiek voor een bepaald vak nodig zijn. Bijvoorbeeld kennis Nederlands op het secundair onderwijsniveau voor instromende studenten aan de universiteit.

Op eindniveau:

- *Domein-specifieke voorkennistoetsen*: meten kennis en inzicht van studenten op een bepaald vakgebied. Het betreft hier het toetsen van wat nog moet komen om inzicht te krijgen in welke onderdelen studenten wellicht al meer, minder of niet beheersen.

Zo kan een domein-specifieke voorkennistoets verschillen blootleggen tussen studenten, daar waar een toets op beginners-niveau dit niet kan. Verder meet een toets op beginners-niveau slechts een deel van de voorkennis, waar deze domein-specifieke voorkennistoets veel meer meet.

Voorts biedt zulk een toets niet alleen meetresultaten, maar ook informatie en richtlijnen om met deficiënties om te gaan, ze vormen een uitgangspunt voor gedifferentieerde diagnostische en begeleidende benaderingen bij het leerproces en om de komende onderwijsperiode te flexibiliseren.

Deze domein- specifieke toetsen bevatten een groter aantal items (40 tot 100) waarmee gewerkt kan worden dan andere toetsen. Ook kunnen de sterktes en zwaktes van studenten geanalyseerd worden.

Aan de Open Universiteit Nederland worden thans diverse vormen van voorkennis- en intaketoetsen gebruikt. Voor meer informatie kan de lezer terecht op <http://www.ou.nl/info-alg-intake/index.htm>.

Voortgangstoetsen

Naast voorkennistoetsen kan tevens getoetst worden tijdens het leerproces met behulp van voortgangstoetsen. Het is belangrijk onderscheid te maken tussen de voortgangstoetsen zoals zij door Wijnen (1995) gebruikt werden (op basis van einddoelen van een opleiding), en de vorm van voortgangsbeoordeling zoals die anders wordt gebruikt (op basis van doelen van curriculumonderdelen of curriculumdoelen) (Dochy, 1992; Dochy, 1997a,b). Bij de Wijnen-vorm is het doel van de einddoeltoetsen om tijdens een studieprogramma regelmatig een beoordeling door te voeren, waardoor een realistisch beeld van de actuele kennis en een beeld van de voortgang van de studenten kan worden opgebouwd. Daartoe worden heel veel items opgesteld die alle betrekking hebben op de einddoelstellingen van het programma, gebaseerd op het beroepsprofiel. De kwaliteitszorg over deze items is belangrijk en bovendien wordt de "itembank" best door verschillende evaluatoren/docenten gemaakt. Deze einddoeltoetsen zijn in principe programma-onafhankelijk. Studenten uit twee verschillende programma's met dezelfde einddoelstellingen (of beroepsprofiel) kunnen dus met dezelfde toets beoordeeld worden. Deze toetsen worden het best bij meerjarige programma's gebruikt, met een tussenperiode van drie maanden tussen de toetsen.

Nadeel van deze toetsvorm is ondermeerdat het oorspronkelijk doel om studenten te verhinderen om stof van buiten te leren steeds moeilijker overeind blijft. Studenten bouwen door de jaren heen ook hun eigen itembank mee op en verhandelen deze aan medestudenten. Zoals bij vele toetsvormen kan ook hier het toetsgericht studeren voorkomen. Daarnaast kunnen studenten bij zo'n brede beoordeling zich erg specialiseren in onderdelen zonder de minimumdoelen te halen.

De voordelen van voortgangstoetsen zijn dat bij beide vormen meer kans is op het vormen van meer functionele kennis: in de einddoel-vorm doordat de kennis voortdurend aanwezig is in het studieprogramma (en dus niet enkel voor de examens wordt ingestudeerd), in de “voortgangs” vorm omdat de toets van het eerste deel toelaat om een goede basis te scheppen voor het tweede deel bij een progressief opgesteld programma. Voortgangstoetsen doorbreken de sterke sturing door het examen door het groot aantal items en effecten van bijsturing kunnen duidelijk gemaakt worden. Verschillende programma’s (of programma-onderdelen voor de tweede vorm) met dezelfde doelstellingen kunnen vergeleken worden en docenten zijn niet rechtstreeks verantwoordelijk voor het beoordelen van het eigen onderwijs en de beoordeling is dus objectiever. Tevens bestaat de mogelijkheid tot een betere aansluiting bij de voorkennis (via bijv. een voorafgaandelijke voortgangstoets).

Automatisering is mogelijk en studenten kunnen de toets interactief afleggen (deelnameplicht is dan voldoende). Tenslotte leiden voortgangstoetsen tot hoger rendement, hogere studiesnelheid en betere studieresultaten als de pacing (ritme) door de docent of instelling wordt opgelegd (Moerkerke, 1996).

Conclusie: nieuwe ontwikkelingen met betrekking tot assessment

In nieuwe vormen van studentgericht onderwijs zoals het leren leren, probleemoplossend leren en competentiegericht leren heeft de docent geen voortdurende controle, de studenten nemen immers een aantal van de taken van de docent over. Deze vormen van onderwijs, eerder krachtige leeromgevingen genoemd, pogen de meest recente resultaten van de onderwijsresearch in de praktijk om te zetten. Het succes van dit coöperatief en probleemgestuurd leren zit echter ook in de verruiming die wellicht geboden wordt in termen van leeropbrengst. Er wordt niet alleen een georganiseerd bestand van basiskennis opgedaan, er wordt evenzeer een arsenaal van essentiële academische vaardigheden ontwikkeld.

De besproken alternatieve toetsvormen zijn in zowel in onderwijspraktijk als in onderzoek noodzakelijke instrumenten gebleken die inspelen op het doel van de krachtige leeromgeving, met name het ontwikkelen van “reflective practitioners”. Uit het voorafgaande is ook te concluderen dat nieuwe toetsmethoden in het hoger onderwijs op verschillende wijzen kunnen worden geïmplementeerd in studentgericht onderwijs. Veelal worden combinaties gebruikt. Onderzoeken naar onder meer validiteit en betrouwbaarheid van deze toetsvormen zijn veelbelovend maar nog te schaars (Birenbaum & Dochy, 1995; Falchikov & Boud, 1989; Sluismans, Dochy & Moerkerke, 1998).

Implementatie van een krachtige leeromgevingen leidt tot grote veranderingen. Dat vergt aanpassingen zowel van de student als de docent. Assessment als “tool for learning” heeft een grote impact op het leren van studenten. De studenten krijgen mede verantwoordelijkheid voor het toetsen. De bruikbaarheid, haalbaarheid en effectiviteit van methoden als self- peer- en co-assessment leidt tot een nieuw toekomstperspectief. Pluspunten van nieuwe toetsvormen liggen onder meer in het bieden van meer mogelijkheden voor “skills development” en de kennis wordt op een meer integratieve wijze verwerkt. Studenten werken samen en worden gemotiveerd na te denken over het eigen leerproces. De zorgpunten bij het invoeren van dergelijke toetsvormen zitten wellicht in de organisatie en de tijdsinvestering. Implementatie betekent een herontwerp van het toetsgebouw en docenten en studenten hebben behoefte aan training in deze nieuwe wijze van werken en denken. Inmiddels heeft onderzoek aangetoond dat het gebruik van deze nieuwe toetsvormen leidt tot een aantal verschillende positieve effecten, met name: een toename in het vertrouwen van studenten om te presteren, een stijgende bewustwording ten aanzien van de kwaliteit van eigen werk, een toenemende reflectie op eigen gedrag en resultaten, een verhoging van de kwaliteit van leeropbrengsten, een stijgende

verantwoordelijkheid van studenten voor het leren en een grotere onafhankelijkheid, een hogere satisfactie van studenten, en het rapporteren van een beter leerklmaat (Sluijsmans, Dochy & Moerkerke, 1998).

De toetsvormen die in dit artikel aan bod zijn gekomen sluiten elkaar niet uit. Ze kunnen in combinatie worden ingezet en toegepast. Zo kan binnen het portfolio heel goed een self-assessment worden opgenomen of kan peer-assessment gebruikt worden voor het toetsen van het proces en produkt van onderwijsgroepen waarna een OAT als afsluiting volgt. Van belang is dat het toetsontwerp goed past bij het curriculumontwerp en dat bij het ontwikkelen van een nieuwe module, cursus of onderwijseenheid het ontwerp van de toetsing zo meteen wordt betrokken.

Referenties

- Archbald, D.A., & Newmann, F.M. (1988). Beyond standardized testing: Assessing, authentic, academic achievement in the secondary school. Madison, Wisconsin: National Center on Effective Secondary Schools.
- Birenbaum, M. (1995). Assessment 2000: Toward a pluralistic approach to assessment. In M. Birenbaum & F. Dochy, *Alternatives in assessment of achievements, learning processes and prior knowledge* (3-30). Boston: Kluwer Academic.
- Birenbaum, M. & Dochy, F. (1995). *Alternatives in assessment of achievements, learning processes and prior knowledge*. Boston: Kluwer Academic.
- Birenbaum, M. (1994). Toward adaptive assessment - The student's angle. *Studies in educational evaluation*, (20), 239-255.
- Bloom, B.S., Hastings, J.T., & Madaus, G.F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.
- Boud, D. (1981). *Developing Student Autonomy in Learning*. London: Kogan Page.
- Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education*, 15, 1, 101-111.
- Boud, D. (1995). *Enhancing Learning through Self Assessment*. Kogan Page. London: Philadelphia.
- Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18, 1, 32-42.
- Courts, P.L. & McInerney, K.H. (1993). *Assessment in higher education. Politics, pedagogy, and portfolios*. London: Westport, Connecticut.
- Bruijns, V. & Pieké, E. (1997). Beoordelen van kwalificaties. *Tijdschrift voor Hoger Onderwijs*, 15, 3, 208-227.
- De Corte, E. (1990) *A State-of-the-art of research on learning and teaching*. Keynote lecture presented at the first European Conference on the First Year Experience in Higher Education, Aalborg University, Denmark, 23-25 April 1990.
- De Rijke, T.R., & Beemer, C.F.H. (1995). Zelf- assessment. In F.J.R.C. Dochy, & T.R. de Rijke (Red.), *Assessment centers: Nieuwe toepassingen in opleiding, onderwijs en HRM*. Utrecht: Lemma.
- Dochy, F.J.R.C. (1992). *Assessment of prior knowledge as a determinant of future learning: the use of knowledge state tests and knowledge profiles*. Utrecht/London: Lemma B.V./Jessica Kingsley Publishers.
- Dochy, F. (1997a). A line of argument for innovation in teaching and assessment starting from students' conceptions and misconceptions in learning processes. *European Journal of Agricultural Education and Extension*, 3, 4, 231-241.
- Dochy, F. (1997b). *Explanatory models and an overview of research results concerning the effects of teaching academic skills in universities*. Invited lecture at the meeting of the

- Deutsche Gesellschaft für Psychologie, Frankfurt am Main, J.W. Goethe Universität, september 30.
- Dochy, F., & Moerkerke, G. (1997). The present, the past and the future of achievement testing and performance assessment. *International Journal of Educational Research*, 27(5), 415 - 432.
- Falchikov, N. (1995). Peer Feedback Marking: Developing Peer Assessment. *Innovations in Education and Training International*, 32(2), 175-187.
- Falchikov, N. & Boud, D. (1989). Student Self-Assessment in Higher Education: A Meta-Analysis. *Review of Educational Research*, 59(4), 395-430.
- Gardner, R & Alexander, P.A. (1994). *Beliefs about text and about instruction with text*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gipps, G.V. (1994). *Beyond testing: Towards a theory of educational assessment*. London: The falmer press.
- Hall, K. (1995). *Co-assessment: participation of students with staff in the assessment process. A report of Work in Progress*. Hall, K. (1995). Invited paper at the Second European Electronic Conference On Assessment And Evaluation, February 6-10, EARLI-AE list European Academic & Research Network (EARN) (EARLI_AE on Listserv@nic.surfnet.nl).
- Harris, D. & Bell, C. (1990). *Evaluating and assessing for learning*. New York: Nichols Publishing Company.
- Imbos, T.J. (1989). Het gebruik van einddoeltoetsen bij aanvang van de studie. Gulpen: Alberts/Druko.
- Keeves, J.P. (1994). Methods of assessment in schools. In T. Husén & T.N. Postlethwaite (Eds.), *International Encyclopedia of Education* (pp. 362-370). Oxford/New York: Pergamon Press.
- Mello, J.A. (1993). Improving individual member accountability in small work group settings. *Journal of Management Education*, 17, 253-259.
- Meyer, C.A. (1992). What's the difference between authentic and performance assessment. *Educational Leadership*, 49, 39-40.
- Moerkerke, G. (1996). *Assessment for flexible Learning*. Utrecht: Lemma.
- Moerkerke, G., & Dochy, F.J.R.C. (1995). Het toetsen van vaardigheden. In G. Ten Dam, H. Van Hout, C. Terlouw, & J. Willems (Eds.), *Handboek onderwijskunde voor het hoger onderwijs* (pp. 214-236). Assen: Van Gorcum.
- Moerkerke, G. & Terlouw, C. (1998). Herontwerp van toetsing. *THEMA*, (1), 19-25.
- Nevo, D. (1995). *School-based evaluation: A dialogue for school improvement*. London, New York: Pergamon Press.
- O'Neil, J. (1992). Putting performance assessment to the test. *Educational Leadership*, 49, 14-19.

- Pilot, A. (1997). *Curricula voor een op informatie gerichte maatschappij*. Paper gepresenteerd op het Consortium COO HBO en SUNCOO conferentie, 11 en 12 december 1997 te Utrecht.
- Schön, D.A. (1987). *Educating the Reflective Practitioner: Towards a New Design for Teaching and Learning in the Professions*. San Francisco: Jossey-Bass.
- Segers, M.S.R., Tempelaar, D., Keizer, P., Schijns, J., Vaessen, E., Van Mourik, A. (1992). *De OverAll Toets: een tweede experiment met een nieuwe toetsvorm*. [The OverAll test: A second experiment]. Maastricht: University of Maastricht.
- Segers, M.S.R. (1995). Problem-Solving and Assessment. The Maastricht experiences. In: W.H. Gijsselaers, D.T. Tempelaar, P.K., Keizer, J.M. Blommaert, E.M. Bernard & H. Kasper (eds.). *Educational innovation in economics and business administration. The case of problem-based learning*. Boston/Dordrecht/London: Kluwer Academic Publishers.
- Segers, M.S.R. (1996). Assessment in a Problem-Based Economics Curriculum. In: M. Birenbaum, & F.J.R.C. Dochy (eds.). *Alternatives in assessment of achievements, learning processes and prior knowledge*. Boston/Dordrecht/London: Kluwer Academic Publishers.
- Segers, M.S.R. (1997). An Alternative for assessing problem-solving skills: the OverAll test. *Studies in Educational Evaluation*, 23, 4, 373-398.
- Segers, M.S.R. (1998). Het toetsen van probleemoplossende vaardigheden. Ervaringen met de OverAll Toets. *Tijdschrift voor hoger onderwijs* (in press).
- Sluismans, D. (1997). *Dichter bij studenten. Een evaluatie van de Post-academische opleiding "Adaptief onderwijs" voor leidinggevendenden in de PABO*. Doctoraalscriptie Vakgroep Onderwijskunde: Nijmegen.
- Sluismans, D., Dochy, F., & Moerkerke, G. (1998). Creating A Learning Environment By Using Self-, Peer- And Co-Assessment. *Learning Environments Research (accepted for publication)*.
- Somervell, H. (1993). Issues in Assessment, Enterprise en Higher Education. *Assessment & Evaluation in Higher Education*, 18, 229-233.
- Weinert, F.E., Schneider, W. (Eds.) (1995). Memory performance and competencies: issues in growth and development. Mahwah, N.J.: Lawrence Erlbaum Publishers.
- Wiggings, G. (1992). Creating tests worth taking. *Educational Leadership*, 49, 26-33.
- Williams, E. (1992). Student Attitudes towards Approaches to Learning and Assessment. *Assessment and Evaluation in Higher Education*, 17(1), 45-58.
- Wijnen, W. (1995). *Medisch onderwijs in de praktijk*. Assen: Van Gorcum.